

NEUROSCIENCE PERSPECTIVES ON SECURITY: TECHNOLOGY, DETECTION, AND DECISION MAKING

EDITED BY: Elena Rusconi, Kenneth C. Scott-Brown and
Andrea Szymkowiak

PUBLISHED IN: Frontiers in Human Neuroscience





frontiers

Frontiers Copyright Statement

© Copyright 2007-2015 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-600-5

DOI 10.3389/978-2-88919-600-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

NEUROSCIENCE PERSPECTIVES ON SECURITY: TECHNOLOGY, DETECTION, AND DECISION MAKING

Topic Editors:

Elena Rusconi, University College London, UK

Kenneth C. Scott-Brown, University of Abertay Dundee, UK

Andrea Szymkowiak, University of Abertay Dundee, UK

In security science, efficient operation depends typically on the interaction between technology, human and machine detection and human and machine decision making. A perfect example of this interplay is ‘gatekeeping’, which is aimed to prevent the passage of people and objects that represent known threats from one end to the other end of an access point. Gatekeeping is most often achieved via visual inspections, mass screening, random sample probing and/or more targeted controls on attempted passages at points of entry. Points of entry may be physical (e.g. national borders) or virtual (e.g. connection log-ons). Who and what are defined as security threats and the resources available to gatekeepers determine the type of checks and technologies that are put in place to ensure appropriate access control. More often than not, the net performance of technology-aided screening and authentication systems ultimately depends on the characteristics of human operators. Assessing cognitive, affective, behavioural, perceptual and brain processes that may affect gatekeepers while undertaking this task is fundamental. On the other hand, assessing the same processes in those individuals who try to breach access to secure systems (e.g. hackers), and try to cheat controls (e.g. smugglers) is equally fundamental and challenging. From a security standpoint it is vital to be able to anticipate, focus on and correctly interpret the signals connected with such attempts to breach access and/or elude controls, in order to be proactive and to enact appropriate responses. Knowing cognitive, behavioral, social and neural constraints that may affect the security enterprise will undoubtedly result in a more effective deployment of existing human and technological resources. Studying how inter-observer variability, human factors and biology may affect the security agenda, and the usability of existing security technologies, is of great economic and policy interest. In addition, brain sciences may suggest the possibility of novel methods of surveillance and intelligence gathering.

These are just a few examples of typical security issues that may be fruitfully tackled from a neuroscientific and interdisciplinary perspective. The objective of our Research Topic was to document across relevant disciplines some of the most recent developments, ideas, methods and empirical findings that have the potential to expand our knowledge of the human factors involved in the security process. To this end we welcomed empirical contributions using different methodologies such as those applied in human cognitive neuroscience, biometrics and ethology. We also accepted original theoretical contributions, in the form of review articles, perspectives or opinion papers on this topic. The submissions brought together researchers from different backgrounds to discuss topics with scientific, applicative and social relevance.

Citation: Rusconi, E., Scott-Brown, K. C., Szymkowiak, A., eds. (2015). Neuroscience Perspectives on Security: Technology, Detection, and Decision Making. Lausanne: Frontiers Media. doi: 10.3389/978-2-88919-600-5

Table of Contents

- 04 Neuroscience perspectives on security**
Elena Rusconi, Kenneth C. Scott-Brown and Andrea Szymkowiak
- 06 Keystroke dynamics in the pre-touchscreen era**
Nasir Ahmad, Andrea Szymkowiak and Paul A. Campbell
- 16 Right inferior frontal gyrus activation as a neural marker of successful lying**
Oshin Vartanian, Peter J. Kwantes, David R. Mandel, Fethi Bouak, Ann Nakashima, Ingrid Smith and Quan Lam
- 22 Prospects of functional magnetic resonance imaging as lie detector**
Elena Rusconi and Timothy Mitchener-Nissen
- 34 Addressing social resistance in emerging security technologies**
Timothy Mitchener-Nissen
- 38 Can laptops be left inside passenger bags if motion imaging is used in X-ray security screening?**
Marcia Mendes, Adrian Schwaninger and Stefan Michel
- 48 Suspiciousness perception in dynamic scenes: a comparison of CCTV operators and novices**
Christina J. Howard, Tom Troscianko, Iain D. Gilchrist, Ardhendu Behera and David C. Hogg
- 57 Looking for trouble: a description of oculomotor search strategies during live CCTV operation**
Matthew J. Stainer, Kenneth C. Scott-Brown and Benjamin W. Tatler
- 66 Cues derived from facial appearance in security-related contexts: a biological and socio-cognitive framework**
Christopher D. Watkins
- 70 The reasoning criminal vs. Homer Simpson: conceptual challenges for crime science**
Noémie Bouhana
- 76 A biological security motivation system for potential threats: are there implications for policy-making?**
Erik Z. Woody and Henry Szechtman
- 81 Sensing, assessing, and augmenting threat detection: behavioral, neuroimaging, and brain stimulation evidence for the critical role of attention**
Raja Parasuraman and Scott Galster
- 91 Non-invasive brain stimulation can induce paradoxical facilitation. Are these neuroenhancements transferable and meaningful to security services?**
Jean Levasseur-Moreau, Jerome Brunelin and Shirley Fecteau
- 104 Why non-invasive brain stimulation should not be used in military and security services**
Bernhard Sehm and Patrick Ragert
- 107 Is it ethical and safe to use non-invasive brain stimulation as a cognitive and motor enhancer device for military services? A reply to Sehm and Ragert (2013)**
Jerome Brunelin, Jean Levasseur-Moreau and Shirley Fecteau



Neuroscience perspectives on security

Elena Rusconi^{1,2,3*}, Kenneth C. Scott-Brown² and Andrea Szymkowiak⁴

¹ Department of Security and Crime Science, University College London, London, UK

² Division of Psychology, Abertay University, Dundee, UK

³ Department of Neurosciences, University of Parma, Parma, Italy

⁴ School of Science, Engineering and Technology, Abertay University, Dundee, UK

*Correspondence: elena.rusconi@gmail.com

Edited and reviewed by:

Hauke R. Heekeren, Freie Universität Berlin, Germany

Keywords: security, deception detection, threat detection, crime science, neuroenhancement, applied neuroscience, applied psychology, military

Security issues have been under the spotlight on a daily basis since the 9/11 terrorist attacks to the Twin Towers, which—aired on live TV—were witnessed by millions of people around the globe. This has been accompanied by the increased availability (and leakage) of security information on the Internet, the increase in public awareness over related issues, and the surge of ethical debates on the possible ethical and legal consequences of “security states”; security has taken priority in political agendas, academic debates, and research funding—the security industry is thriving. Against this background, more and more academics are exploring ways to contribute to the debate, and to inform and influence security decision making. This is both a challenging and a rewarding enterprise and neuroscience promises game-changing innovations.

Security science however is a multidisciplinary field, where physics and engineering, computer science and biology, psychology and medicine, pharmacology and neuroscience, philosophy and jurisprudence, sociology and ethology can all bring valuable contributions to the table. Accordingly, in this Research Topic we have hosted relevant contributions from neuroscience and psychology experts but also dipped into other disciplines such as engineering, physics, computer science, crime science, jurisprudence, and sociology of science. We would like to thank all of the authors and the reviewers for their excellent contributions and their effort in spanning disciplinary boundaries. It is not easy to strike the right balance between expertise and accessibility, to explore a little further outside of our comfort niche and convey meaning to a multifaceted type of readership, such as the one that can be reached via open access and via Frontiers in Human Neuroscience in particular. We hope that our Research Topic will provide a useful contribution to the dialog among disciplines on security-related issues and also a successful example of how the—often artificial—disciplinary boundaries can be challenged.

Almost every aspect of security is inextricably connected with technology. One of the aspects where accelerated advancements have been witnessed in recent years is the incorporation of psychological and physiological measures via new technologies. Reviews of the area of biometrics, traditionally described as the identification of individuals (or their emotional states) using physiological and behavioral characteristics, such as finger prints, iris or retinal patterns, facial features, handwriting or typing on a keyboard (see Ahmad et al., 2013), to name but a few,

and the uses of sophisticated imaging techniques, such as fMRI to detect indicators of deception (see Rusconi and Mitchener-Nissen, 2013; Vartanian et al., 2013), provide two representative examples of this. The ethical and legal aspects of the use of such technologies are widespread. On the one hand, the data gathered with such technology are challenging to process and interpret, so this bears the question as to how clearly experts can present their evidence to a jury in the context of criminal justice systems; on the other hand, findings based on the use of the technology are still far from being fully reliable, research based on laboratory experiments restricts the ecological validity of such measures, and the complexity and sensitivity of the technology makes it difficult to run trials outside the laboratory or even envisage real-world applications. Another question pertains to how transparent individuals and their internal (e.g., emotional, intentional, deceptive, etc.) states can ever be made, even with a fine-grained analysis of human behavior or characteristics, as individuals become aware of advancements in technologies to assess these. Drawing on the concept of measures and countermeasures—can human suspicious behavior and intent be camouflaged so well it is not traceable by the latest neuroscientific detection systems? It is not yet clear to which extent the sophistication of technology and human perception to assess human mental and behavioral activity is juxtaposed with the sophistication of individuals to evade these security measures. Further, fully successful detection systems would have human rights, policy making and social acceptance implications, a critical issue that has been clearly recognized (see Mitchener-Nissen, 2013; Rusconi and Mitchener-Nissen, 2013).

While the above methods investigate physiological or behavioral indices with technological means and algorithms, the use of human operators during incident or threat detection is still irreplaceable and critical to the security discourse (see Howard et al., 2013; Mendes et al., 2013; Stainer et al., 2013). This bears the question on how secure we actually are as both technology and humans are fallible in their decision making. It is, however, generally assumed that the output of visualization techniques such as CCTV and transmission x-rays can be appropriately assessed by trained individuals. CCTV operators are presented with large volumes of constantly updating visual information, and the navigation through this temporal and spatial data feed is very demanding. In transmission x-rays, the difficulty of complex

image interpretation lies mostly in the superposition of several two-dimensional projections and the unusual views by which objects are seen in static images. To gather information about human performance in security image interpretation, diametrically opposite approaches can be adopted—from a classical hypothesis-driven experimental method to an *in situ* observational method reminiscent of a cognitive-ethological approach (Howard et al., 2013; Stainer et al., 2013). While technological improvements are being pursued to increase the efficiency of the screening process from an engineering and physics standpoint, these efforts may be hindered by the intrinsic limitations of the human visual perception system (see Mendes et al., 2013). Notably, to the extent that decisions are made by people, the assessment of potentially dangerous situations in a social environment is subjected to the limitations of the cognitive system that can be swayed or driven by appearances, biases, and previous experience (see Watkins, 2013; Woody and Szechtman, 2013). Of course, the same constraints will also apply to the decisions made by those individuals who actively engage in criminal activities (i.e., those who create breaches in security rather than help maintain it)—an awareness that seems yet to have been fully incorporated in evidence-based crime science (Bouhana, 2013).

Brain manipulation techniques such as Transcranial Magnetic Stimulation and transcranial Direct Current Stimulation may help overcome some of the intrinsic limitations of human security operators with their potential to augment human performance in a range of tasks (Levasseur-Moreau et al., 2013; Parasuraman and Galster, 2013). Although the state of the art may not be mature enough to allow for direct translations into the security field, it is of paramount importance that neuroscientists engage as early as possible with professionals from other disciplines to formulate critical appraisals of the larger-picture implications of any of the envisaged uses (Brunelin et al., 2013; Sehm and Ragert's, 2013). Arguably, rather than hinder or slow down scientific progress, these early multidisciplinary appraisals and interactions will help secure more public support and more resources for neuroscience research.

REFERENCES

- Ahmad, N., Szymkowiak, A., and Campbell, P. (2013). Keystroke dynamics in the pre-touchscreen era. *Front. Hum. Neurosci.* 7:835. doi: 10.3389/fnhum.2013.00835
- Bouhana, N. (2013). The reasoning criminal vs. Homer Simpson: conceptual challenges for crime science. *Front. Hum. Neurosci.* 7:682. doi: 10.3389/fnhum.2013.00682
- Brunelin, J., Levasseur-Moreau, J., and Fecteau, S. (2013). Is it ethical and safe to use non-invasive brain stimulation as a cognitive and motor enhancer device for military services? A reply to Sehm and Ragert (2013). *Front. Hum. Neurosci.* 7:874. doi: 10.3389/fnhum.2013.00874
- Howard, C. J., Troscianko, T., Gilchrist, I. D., Behera, A., and Hogg, D. C. (2013). Suspiciousness perception in dynamic scenes: a comparison of CCTV operators and novices. *Front. Hum. Neurosci.* 7:441. doi: 10.3389/fnhum.2013.00441
- Levasseur-Moreau, J., Brunelin, J., and Fecteau, S. (2013). Non-invasive brain stimulation can induce paradoxical facilitation. Are these neuroenhancements transferable and meaningful to security services? *Front. Hum. Neurosci.* 7:449. doi: 10.3389/fnhum.2013.00449
- Mendes, M., Schwaninger, A., and Michel, S. (2013). Can laptops be left inside passenger bags if motion imaging is used in X-ray security screening? *Front. Hum. Neurosci.* 7:654. doi: 10.3389/fnhum.2013.00654
- Mitchener-Nissen, T. (2013). Addressing social resistance in emerging security technologies. *Front. Hum. Neurosci.* 7:483. doi: 10.3389/fnhum.2013.00483
- Parasuraman, R., and Galster, S. (2013). Sensing, assessing, and augmenting threat detection: behavioral, neuroimaging, and brain stimulation evidence for the critical role of attention. *Front. Hum. Neurosci.* 7:273. doi: 10.3389/fnhum.2013.00273
- Rusconi, E., and Mitchener-Nissen, T. (2013). Prospects of functional magnetic resonance imaging as lie detector. *Front. Hum. Neurosci.* 7:594. doi: 10.3389/fnhum.2013.00594
- Sehm, B., and Ragert, P. (2013). Why non-invasive brain stimulation should not be used in military and security services. *Front. Hum. Neurosci.* 7:553. doi: 10.3389/fnhum.2013.00553
- Stainer, M. J., Scott-Brown, K. C., and Tatler, B. (2013). Looking for trouble: a description of oculomotor search strategies during live CCTV operation. *Front. Hum. Neurosci.* 7:615. doi: 10.3389/fnhum.2013.00615
- Vartanian, O., Kwantes, P., Mandel, D. R., Bouak, F., Nakashima, A., Smith, I., et al. (2013). Right inferior frontal gyrus activation as a neural marker of successful lying. *Front. Hum. Neurosci.* 7:616. doi: 10.3389/fnhum.2013.00616
- Watkins, C. D. (2013). Cues derived from facial appearance in security-related contexts: a biological and socio-cognitive framework. *Front. Hum. Neurosci.* 7:204. doi: 10.3389/fnhum.2013.00204
- Woody, E. Z., and Szechtman, H. (2013). A biological security motivation system for potential threats: are there implications for policy-making? *Front. Hum. Neurosci.* 7:556. doi: 10.3389/fnhum.2013.00556

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 August 2014; accepted: 23 November 2014; published online: 09 December 2014.

Citation: Rusconi E, Scott-Brown KC and Szymkowiak A (2014) Neuroscience perspectives on security. *Front. Hum. Neurosci.* 8:996. doi: 10.3389/fnhum.2014.00996

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Rusconi, Scott-Brown and Szymkowiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Keystroke dynamics in the pre-touchscreen era

Nasir Ahmad^{1,2}, Andrea Szymkowiak³ and Paul A. Campbell^{1,2} *

¹ CICA Group (Concepts & Innovation in Cavitation and Sonoptic Sciences), Carnegie Physics Laboratory, University of Dundee, Dundee, UK

² Division of Molecular Medicine, College of Life Sciences, University of Dundee, Dundee, UK

³ School of Science, Engineering and Technology, University of Abertay Dundee, Dundee, UK

Edited by:

Elena Rusconi, University College
London, UK

Reviewed by:

Giuseppe Sartori, University of
Padova, Italy

Ishbel Duncan, University of
St Andrews, UK

*Correspondence:

Paul A. Campbell, CICA Group
(Concepts & Innovation in Cavitation
and Sonoptic Sciences), Carnegie
Physics Laboratory, University of
Dundee, Dundee DD1 4HN, UK
e-mail: p.a.campbell@dundee.ac.uk

Biometric authentication seeks to measure an individual's unique physiological attributes for the purpose of identity verification. Conventionally, this task has been realized via analyses of fingerprints or signature iris patterns. However, whilst such methods effectively offer a superior security protocol compared with password-based approaches for example, their substantial infrastructure costs, and intrusive nature, make them undesirable and indeed impractical for many scenarios. An alternative approach seeks to develop similarly robust screening protocols through analysis of typing patterns, formally known as keystroke dynamics. Here, keystroke analysis methodologies can utilize multiple variables, and a range of mathematical techniques, in order to extract individuals' typing signatures. Such variables may include measurement of the period between key presses, and/or releases, or even key-strike pressures. Statistical methods, neural networks, and fuzzy logic have often formed the basis for quantitative analysis on the data gathered, typically from conventional computer keyboards. Extension to more recent technologies such as numerical keypads and touch-screen devices is in its infancy, but obviously important as such devices grow in popularity. Here, we review the state of knowledge pertaining to authentication via conventional keyboards with a view toward indicating how this platform of knowledge can be exploited and extended into the newly emergent type-based technological contexts.

Keywords: keystroke analysis, pre-touchscreen, security, authentication, identity

OVERVIEW: AUTHENTICATION

With the magnitude of online and computer-based systems and services increasing rapidly over recent decades, the need for enhanced computer security has become a significant concern. Accurate authentication of user identity is of paramount importance, and the following techniques are most often used toward that objective (Wood, 1977):

- A unique ("hidden") electronic key is employed, known only by the user, most commonly a password or PIN (personal identification number), which serves to access the system.
- A physical security measure (e.g., formal identification/swipe card) is used to identify the user. Systems often use such objects in conjunction with a hidden key.
- A biometric authentication system may be used such that a user's unique physical or behavioral traits are inspected for verification of his/her identity by comparison with a validated database record.

Currently, systems most commonly in use prompt for a hidden password alongside an identifying username. These systems often recommend that the password used should be a completely unique, complex, and long entry that is not used for any other purpose. In reality, most users find remembering different sets of long alpha-numeric sequences for each and every service impractical, and tend to reuse the same password for more than one service. Alternatively, users might record their passwords, either electronically or on paper. Moreover, in order to assist password recall, users will often create a password or PIN which is in some way related

to a personal aspect of their lives (e.g., birthdays and names). Recording and repetition of passwords obviously compromises the hidden requirement for their unique key's security, opening the way for intruder access. Furthermore, passwords which are based upon the personal details of a user's life can be susceptible to dictionary or "brute force" hacks, as well as educated guesses made by an informed imposter.

Systems with physical security measures also represent security issues, as the physical nature of the tokens/keys makes them prone to theft, or the data within them may be simply cloned, again compromising the target service. The implementation of further security layers is therefore a critical goal of biometric authentication.

Biometric authentication and identification are methods whereby unique physiological attributes or characteristic traits of individuals are used to verify their identity. Analyses of an individual's fingerprint or unique iris pattern are two of the most widely used security techniques in this field. Although these methods are a great deal more secure than a single password, their significant setup costs and the intrusive nature of scanning makes them impractical for many purposes. Regardless of cost, it must be recognized that such systems remain fallible, but at the moment still prevail as the most accurate route to authentication available.

Analysis of keystroke dynamics is an alternative approach to biometrics authentication. This technique makes use of the natural pattern and manner in which a user types at a keyboard to verify their identity. In moving toward the establishment of a validated record, a user must initially be enrolled within a system,

whereupon the user's typing pattern is recorded and stored within the system. This record can then be consulted/compared when the user attempts to gain access to his/her system. This type of authentication would be implemented within a login system such that a user's entry of their username and password is analyzed – thereby adding a new layer of security to the existing systems.

A BRIEF HISTORY OF KEYSTROKE DYNAMICS

Of the early documented research and analysis into keystroke dynamics authentication, the insightful and thorough article by Gaines et al. (1980) is particularly illuminating. Their research showed that the field was effectively initiated during the initial manual phase of telegraphy, where operators had been observed to have a unique “fist” (tapping style) by which their colleagues could often identify them. By extrapolation of that principle, they hypothesized that a similar signature could arise during regular typing and a preliminary analysis was conducted, investigating the relevance and effectiveness of a system of identification of individuals, based upon their unique keystroke signatures. While Gaines et al. (1980) concluded that such a system *could* be effective as a tool for authentication, they acknowledged that the findings were only based on a small sample, using the data from seven touch typists, their task having been to type three distinct sections of text, some 4 months apart. Moreover, not every typist was available for each repeat session. Despite this small number of subjects, the researchers were able to observe and differentiate between their differing typing styles.

The study by Gaines et al. (1980) popularized the use of digraph data, i.e., data associated with two successively typed letters (viz. *in*, *io*, *no*, *on*, and *ul*) – a method that paved the way for many subsequent keystroke analysis groups to forge a first path into the field and which has remained popular with analysts. Following this preliminary assessment of the viability of keystroke analysis, other researchers pursued different routes for user identification and authentication, with an emphasis on the reduction of two error rates, i.e., false acceptance rate (FAR) and false rejection rate (FRR). FAR involves the mistaken acceptance of imposters, i.e., false positives; FRR is the error associated with the false rejection of valid users, i.e., false negatives. By altering the threshold for acceptances (or rejections), FAR and FRR can be optimized to generate a measure of equal error rate (EER), that is, when FAR is equal to FRR. The use of this measure allows for a comparison of the accuracies across studies that may use different authentication methods and subject numbers.

Subsequently, research arising in the 1980s and 1990s (e.g., Umphress and Williams, 1985; Young and Hammon, 1989; Bleha et al., 1990; Joyce and Gupta, 1990; Obaidat and Macchairolo, 1993, 1994; De Ru and Eloff, 1997; Lin, 1997; Monrose and Rubin, 1997; Obaidat and Sadoun, 1997, 1999; Robinson et al., 1998; Coltell et al., 1999; Monrose et al., 1999; Tapiador and Sigüenza, 1999) began to explore alternative methods of keystroke analysis, typically employing a range of novel mathematical analysis techniques, but also differing in the formal data collection method. Statistical techniques (Gaines et al., 1980; Umphress and Williams, 1985; Young and Hammon, 1989; Bleha et al., 1990; Joyce and Gupta, 1990; Bleha and Obaidat, 1991; Monrose and Rubin, 1997; Robinson et al., 1998; Coltell et al., 1999; Monrose et al., 1999;

Obaidat and Sadoun, 1999), neural networks (Obaidat and Macchairolo, 1993, 1994; Lin, 1997; Obaidat and Sadoun, 1997), and fuzzy logic (De Ru and Eloff, 1997; Tapiador and Sigüenza, 1999) have all been used in attempts to increase the accuracy and effectiveness of keystroke authentication. The data collected for use with these techniques were not only recorded directly by the computer being actively used, but also collected via a local network or server arrangement (Bleha et al., 1990; Bleha and Obaidat, 1991; Obaidat and Macchairolo, 1993; Tapiador and Sigüenza, 1999), showing that such keystroke authentication could be implemented in an online system.

A further innovation at this stage was that the keystroke analysis system could be implemented not only to authenticate users during login, but also to make that judgment more robust by recording/monitoring keystrokes during the downstream session – whilst they wrote documents/emails. If an intruder was detected, some action would be taken by the system to limit access. Formally, keystroke analysis completed only at log-in became known as *Static Analysis* while that undertaken during the entire user session is known as *Continuous Analysis*.

Research in the most immediate past (Changshui and Yanhua, 2000; Cho et al., 2000; Haider et al., 2000; Monrose and Rubin, 2000; Wong et al., 2001; Bergadano et al., 2002; D'Souza, 2002; Henderson et al., 2002; Mantyjarvi et al., 2002; Eltahir et al., 2003, 2004, 2008; Jansen, 2003; Nonaka and Kurihara, 2004; Peacock et al., 2004; Araújo et al., 2005; Chang, 2005; Lee and Cho, 2005; Rodrigues et al., 2005; Curtin et al., 2006; Hosseinzadeh et al., 2006; Lv and Wang, 2006; Clarke and Furnell, 2007; Hocquet et al., 2007; Loy et al., 2007; Grabham and White, 2008; Lv et al., 2008; Saevanee and Bhatarakosol, 2008, 2009; Campisi et al., 2009; Hwang et al., 2009a,b; Killourhy and Maxion, 2009; Revett, 2009; Nguyen et al., 2010; Chang et al., 2011, 2012; Giot et al., 2011; Karnan et al., 2011; Teh et al., 2011; Xi et al., 2011) has incorporated newly developed mathematical and data recording techniques – again employing statistical techniques and neural networks, but also attempting to fuse data from multiple parallel sensors. The types and differences between the various mathematical techniques are discussed in the next section. Other than new analysis techniques, novel types of data were also considered and analyzed. For example, existing keyboards were modified to generate a measure of the pressure with which a user presses a single key (Henderson et al., 2002; Eltahir et al., 2003, 2004; Nonaka and Kurihara, 2004; Lv and Wang, 2006; Hocquet et al., 2007; Loy et al., 2007) – the aim of which was, again, to increase the veracity of user analyzed identity. Such pressure measurements proved useful in building a more accurate template of users' unique keystroke patterns. Keyboard modification was generally achieved by addition of an analog electronic component sensitive to pressure, or some indirect measure of pressure (e.g., piezo-resistive film) was either placed between the keyboard and the surface upon which it sat, or alternatively, beneath a number of active keys. On-board microphones (Nguyen et al., 2010) could also be employed to take an indirect measure of pressure, based upon the characteristic acoustic signature arising.

The increased demand for security in other areas of modern technology has also led to keystroke dynamics research having been carried out on mobile phones, e.g., button-based (Clarke and

Furnell, 2007; Campisi et al., 2009; Hwang et al., 2009a) and touch-based devices (Mantyjarvi et al., 2002; Saevanee and Bhatarakosol, 2008, 2009); numerical keypad systems (Mantyjarvi et al., 2002; Grabham and White, 2008), and also web-based systems (Bleha et al., 1990; Bleha and Obaidat, 1991; Obaidat and Macchairolo, 1993; Tapiador and Sigüenza, 1999; Cho et al., 2000; Curtin et al., 2006). The applications of these systems are discussed alongside a measure of the accuracy of each method in subsequent sections.

It should be noted that research undertaken in this field tends to make use of different sets of data: studies generally have different numbers of subjects, and employ different sets of “test text” as authentication samplers. For example, some studies require the subjects to type out a username and password combination (of relatively short length) whilst other studies request the input of a large section of text. The difference in methodologies provides a challenge for making direct comparisons among papers using the stated error rates alone. Furthermore, the papers described below make use of different classes of keystroke latency. In principle, four types can be used: the timing for a key to go “Down–Up” (hold time), “Down–Down,” “Up–Down,” and “Up–Up.” Different combinations of these four latencies have been exploited by different groups and a specific choice may affect the indicative error rates arising.

KEYSTROKE DYNAMICS FOR SECURITY

MATHEMATICAL ANALYSIS TECHNIQUES

The mathematical approaches to keystroke analysis can be divided into the following groups, all of which are discussed below:

- Statistical techniques
- Artificial neural networks
- Fuzzy logic
- Other

Statistical techniques

Statistical analysis of keystroke dynamics is perhaps the most researched avenue within the field. Initially, basic statistical features such as the mean and standard deviation of keystroke timings were utilized, however, these were quickly expanded upon to ascertain the detection of anomalies and irregularities of timings.

t-test analyses were prevalent in the earliest reports. This method of analysis required the mean values of two samples to be taken and compared, in order to determine whether the two samples emanated from the same original source (typist). In the case of keystroke dynamics, the *t*-test analysis was used not only to compare the mean, but also the standard deviation of inter-key latencies (Gaines et al., 1980; Umphress and Williams, 1985).

In the work by Gaines et al. (1980), a group of repeated digraphs was analysed using this method, and with subjects typing comparatively large amounts of text, this technique proved effective. It should be noted, however, that with regular password strings, the digraphs are not repeated sufficiently often for this technique, in the form alluded to, to be appropriate for a login-based analysis of keystroke dynamics. However, although this technique may not be directly applicable to [short] login keystroke analysis, the accuracy rates, as mentioned above, proved encouraging, and certainly provided the initial indications that statistical analyses could be

sufficiently accurate for authentication in the context of computer systems.

Nowadays, techniques exploiting the features of statistical analysis often combine the mean and standard deviations for keystroke latencies as reference data (Joyce and Gupta, 1990; Robinson et al., 1998; Araújo et al., 2005). The data are collected when users are initially registered into a system, whereupon it is required to enter their authentication string (e.g., password) multiple times. The latency times are combined to create a “vector.”

Many reports have used a variation on this technique, combining it with an intrinsic threshold so that when a user attempts to access the system, the latencies of the entered authentication string are compared against the reference signature. If the differences between the two are within the threshold, the user is accepted. For example, Araújo et al. (2005) used four keystroke features, each with 10 character long password strings. Using the mean and standard deviation, a template for each keystroke feature for each element was made and stored. Interestingly, this approach was tested not only by valid users and imposters, but also “observer” imposters, such that these subjects were allowed to view the valid subject’s typing style. In the event, Araújo et al. (2005) were able to achieve an error rating of $FRR = 1.45\%$ and $FAR = 1.89\%$, an impressively high outcome for this style of statistical analysis.

Other studies have also made use of Bayesian analysis (Bleha et al., 1990; Bleha and Obaidat, 1991) in an attempt to achieve a lower rate of misclassification. This technique treats the pattern vector as a multivariate probability density function, and the analysis, when combined with a minimum distance classifier, was used extensively in attempts to gain accuracy. Minimum distance classifiers define the difference between two samples as an index of similarity. This can be beneficial in keystroke dynamics in that setting a threshold for this minimum distance allows a user to be authenticated in a keystroke analysis system within a threshold unique to their own variation in keystroke signature.

Other statistical analysis techniques include methods of distance classification and probability measures (weighted and non-weighted; Monroe and Rubin, 1997; Robinson et al., 1998). Auto-regressive (AR) and AR moving-average (ARMA) models were considered with and without measures of pressure (Changshui and Yanhua, 2000; Eltahir et al., 2004). Hidden Markov models (HMMs) have been implemented (Chang, 2005) with a similarity histogram, and, by attempting to recognize patterns, produce promising results. Gaussian mixture models (GMMs) have also been tested and found to attain low (under 3%) error rates (Hosseinzadeh et al., 2006). Moreover, combined multiple techniques have had their distinctive advantages, such as the fusion of a statistical method, a measure of disorder between feature vectors and time discretization (Hocquet et al., 2007). Teh et al. (2011) completed a multi-layer fusion of a Gaussian probability density function (GPD) and a directional similarity measure (DSM) attaining an EER of circa 1% with a “Multiple Layer Multiple Expert” fusion technique employing AND voting rules. This approach generally yields better error rates than many other fused analytical procedures, for instance, those making use of statistical and fuzzy logic approaches.

Artificial neural networks

Artificial neural networks (ANNs) are mathematical (or computational) models that imitate, and are inspired by, the function and processes in a biological neural network. The system is built using artificial neurons with well defined connection prescriptions. These ANNs can be utilized to extract complex connections and patterns in data.

In the context of keystroke dynamics, the input to the ANN is largely the timing between successive keystrokes. These keystroke timings are then computed through the network comparator to pre-collected and validated data, in order to determine whether the user is authentic. “Back-propagation” neural networks are usually implemented, which are feed-forward networks employing multiple layers between the input and output nodes.

The initial use of ANNs was to aid in user identification using keystroke dynamics (Obaidat and Macchiarolo, 1993, 1994) and several of the first wave of studies to implement such a neural network approach simply used users’ keystroke latencies as the basis for discrimination. It was found that a hybrid “sum-of-products” network gave the least error: this type of network consists of a simple back-propagation setup between the input and hidden layers followed by a sum-of-products connection between the hidden and output layers. This sum-of-products technique acts in such a way that the output of one node is the weighted sum of the inputs from multiple other nodes. The majority of the ANN systems use some variation on this technique, although in many cases, the ultimate analysis is completed by different types and complexities of the system. ANNs deliver reasonably high accuracy, 97.8% Obaidat and Macchiarolo (1993) and 96.2% for the same technique in Obaidat and Macchiarolo (1994) with a short neural network training time (~1 min training). However, in this case it is important to note that the system was typically used for identification only, i.e., the user keystrokes were matched against a database to find the closest match. Thus, accuracy was not an indication of how well the system was able to identify imposters.

Several research groups subsequently began to investigate the application of ANNs in verification, as a competing technique to statistical analysis. Here, one of the most successful research studies into this area was undertaken by Obaidat and Sadoun (1997), who tested both statistical and neural network approaches to keystroke dynamics verification and achieved zero percent error rates (EER) for learning vector quantization (LVQ), radial basis function networks (RBFN), and Fuzzy ARTMAP neural networks (i.e., a neural network architecture based on the synthesis of fuzzy logic and adaptive resonance theory). Whilst this result was extremely promising, it should be noted that the extent of data sampling required on participants was considerable and therefore poses limits on the implementation of such systems. Over the course of an 8-week experiment, 15 “valid” users provided 225 sequences, and 15 “invalid” users provided 15 samples each. The samples taken from invalid users were used to “train” the system, whereas in a realistic system, there would be no access to invalid user keystrokes for such training purposes (unless it would be an integral part of an intense enrolment procedure). Nevertheless, the strength of such studies is that they underscore the applicability and potential for neural network approaches

as part of the authentication/verification strategy. These same authors also discuss, and conclude, that the duration over which keys are held (hold time) is a better measure for keystroke signature than the time between key presses (inter-key time). However, the combination of *both* these timing sets serves to reduce errors.

Around the same period, Lin (1997) made first use of a dynamic multi-layered back-propagation neural network. This approach operated with distinct weightings being assigned to the keystroke latencies as they progressed through the system. These weightings were based on training sample data, and were constructed such that the root mean square error was reduced to an appropriate threshold. This study was able to validate users with a very low error – with FAR reaching lows of 0% and FRR = 1.1%. Although the error ratings were somewhat higher than those by Obaidat and Sadoun (1997), a much larger number of participants was tested (90 valid users and 61 invalid users) and intruder samples were not trained within the system, lending feasibility to its implementation.

More recently, Cho et al. (2000) developed a web-based neural network identity verification system and were able to attain very low error rates (average FRR error of 1% when FAR was 0%) using a multi-layer perceptron (MLP). Here, 25 valid users supplied 150–400 samples, with the last 75 being selected for testing. In parallel, 15 invalid users supplied five imposter attempts for each user, again resulting in 75 test signatures. The system was not required to be trained with the imposter signatures, however, the number of training signatures supplied by the user (75–325) would likely be too large, unless a continuous analysis were practical in the context of the application. A web-based system was also implemented using a Java applet that could be run within a web browser to connect to the server, illustrating that the system is available for electronic commerce applications.

The final notable approach within this category, k-NN, or k-nearest neighbor algorithms, has also been used with neural networks in order to accomplish pattern recognition. Wong et al. (2001) used a Euclidean distance measure for the nearest neighbor classification, however, the error rates achieved in this case were generally worse than those from the other studies employing neural networks.

Fuzzy logic

Fuzzy logic is a type of probabilistic logic that deals with reasoning that is approximate rather than fixed. For example, where other “crisp” logic systems have only two states (true/false, on/off) fuzzy logic makes use of the multi-valued interval between these states.

De Ru and Eloff (1997) made use of fuzzy logic as an analysis technique for keystroke dynamics. Here, the group used not only the time intervals between successive characters but also a measure of the typing *difficulty* of successive letters. This classification of difficulty was based upon the distance of the keys involved, and whether or not any of them were capitalized or had a range of whole number values. The time interval between two successive keystrokes was also identified using fuzzy logic, and subsequently binned within subsets: very short; short; moderately short; and somewhat short. By combining the timing and typing difficulty, a specific keystroke

combination was assigned to categories within some degree (e.g., 20% high and 40% short etc.). Using all of these variable and approximately 20 “fuzzy” rules, the group created a system of keystroke analysis which was able to function, but with some error.

Tapiador and Sigüenza (1999) created an Internet-based keystroke analysis system that made use of a username and password to create a fuzzy template. When a user then attempted to log-in, the sample was compared to the fuzzy template for authentication. Whilst the use of simple username and passwords aided the accuracy of their keystroke analysis system, it could, however, lead to intruders being more readily able to ascertain this access password. The authors did not provide any detailed information on error rates, however, and the statistically small sampling with only nine participants might limit the generalizability of this study.

Other

Although the majority of research in this field focussed on the application of digraph and inter-key latencies, some studies also approached the field with other techniques such as trigraph latency. Bergadano et al. (2002) made use of such trigraph latencies in a novel approach to keystroke dynamics analysis. They achieved a reasonably competitive error rate (4% FRR and 0.01% FAR). The use of 154 participants is statistically favorable compared with many other studies in this field, however, participants were tasked to enter a text consisting of 683 characters, which could be perceived as cumbersome or impractical for covert implementation. In this study, the data analysis was unique in that the group used mathematical techniques to arrange trigraphs in order of increasing typing time for each word. This created a “model” for that user such that when users subsequently attempted to access the system, their typing sample was compared to their specific model: if the distance between the two was sufficiently small, the user was accepted.

Many studies have made use of large sections of text when attempting to verify a user’s identity. In some cases, this was simply to ensure that there were sufficient data to facilitate reliable keystroke analysis, however, as already highlighted, such systems would be less useful for applications for user verification with log-in strings (username and password). However, they do underscore the applicability and accuracy of a system which monitors free text in a continuous mode where a user’s typing style throughout their active session is assessed. Curtin et al. (2006) studied the feasibility of a system monitoring large sections of text by extracting information such as the means and standard deviations of typing times for the eight most frequent letters in the alphabet (e, a, r, i, o, t, n, s), the means and standard deviations of the transition times between the most common letter pairs (in, th, ti, on, an, he, al, er, etc.), variables related to the number of presses of special keys (delete, enter, shift, arrow keys, etc.), the number of times the mouse keys were used (also double clicks), and the total time duration of the text input. A nearest neighbor classifier using Euclidean distance was then used to compare test data to training data for identification purposes. The classifier achieved accuracies greater than 90% for recognition, with accuracies up to 100% under certain conditions (large sections of text and small

participant size). This study showed the feasibility of this system, however, and importantly, did *not* test the system with imposter keystrokes to test detection in that context. Thus, this system could only be implemented to ensure that valid individuals were not making use of unauthorized machines, systems, or files.

Lee and Cho (2005) created a new system for classical keystroke dynamics that made use of valid and imposter training samples. Imposter samples become useful over time by the collection of data when imposters attempt to access a system, thus allowing for tightening of the signature of a user, so that the algorithm can more accurately identify valid and invalid users. After testing this system with six different analysis techniques, the one-class LVQ (1-LVQ) and support vector data description (SVDD) were found to be the most accurate, when inclusion of imposter samples were available. Although the inclusion of imposter samples in this case and others results in an increase in accuracy of the system, acquisition of such samples can be difficult. An imposter would first have to access the system knowing the password and be caught and identified as not being a valid user, whereas, if a valid user’s attempt was flagged as an imposter, the accuracy with which the valid user could be identified then might be reduced. Therefore, there remain significant issues with such systems at present.

VARIABLES AND EQUIPMENT

Pressure

After attaining fairly high accuracies with keystroke latency analysis, investigations into other variables which could be used to aid this accuracy were developed. The most applicable and investigated addition was that of keystroke pressure. Measures of pressure were achieved by making use of piezo-electric and piezo-resistive sensors interfaced with the computer system to which the active keyboard was connected (Eltahir et al., 2003, 2004; Nonaka and Kurihara, 2004). These sensors were either placed beneath specific (or all) keys (Eltahir et al., 2003, 2004) or upon the support sections of the keyboard (Nonaka and Kurihara, 2004).

For verification, details of the key-specific pressure waveform, or its associated temporal characteristics, were stored, and were then consulted when a user attempts log-in. The use of this additional pressure variable was seen to increase the accuracy with which the users were validated, albeit with varying degrees of success.

Nonaka and Kurihara (2004) made use of pressure waveforms by placing two pressure sensing strips as the keyboard support beneath the “W” and “O” keys. In this case they not only used the waveforms to attain pressure measures but also as a means to more accurately measure keystroke timings. To attain these accurate measures of keystroke timing, they reduced the pressure waveform to a set of transforms equivalent to maximal overlap discrete Haar wavelet transforms (MOHWT). The system was used with a small number of subjects, however, details of testing were not provided.

Eltahir et al. (2004) implemented an AR classifier for use with creating pressure templates for user validation. Eltahir et al. (2008) developed this method further and used an AR classifier with stochastic signal modeling for the analysis of the pressure aspect of the keystroke signature. This pressure template was used to verify user identities and was integrated into a program called

pressure-based biometric authentication system (PBAS). The system was created with a normal keyboard with embedded force sensors connected to a data acquisition system (filtering and amplification followed by a connection to a digital to analog PCI card in a PC). A measure of the Total Square Error (TSE) was used to discriminate between valid and invalid users. Here, the experiments were carried out with 23 participants and the group was able to achieve an EER of just over 3%.

Lv and Wang (2006) made use of pressure measurements for keystroke verification using three analysis methods. The three analysis methods consisted of a measure of global statistical features of the pressure wave (mean, standard deviation, difference between max and min, positive and negative energy centers), dynamic time warping of the waveform and traditional statistical keystroke analysis. These analyses were carried out after pre-processing of the waveforms using noise removal and normalization. The best error rates were achieved when each of the analysis techniques were weighted and applied. This resulted in an error rate of 1.41% EER, which was lower compared to the error rate when measures of pressure were removed, i.e., 2.04%. Thus, it is clear that pressure does indeed increase the accuracy of the verification, however, this small (0.63%) increase in accuracy should be evaluated based on the cost of the additional components required for pressure measurement, which are not available on typical keyboards.

Loy et al. (2007) used the ARTMAP-FD (FD – familiarity discrimination) neural network as a competing neural network analysis technique. In this case pressure was used by applying piezo-resistive force sensors beneath the keyboard matrix. After baseline subtraction, a fast Fourier transform (FFT) was used to transform the pressure time signals into frequency domain signals. Again, with the use of pressure, a reduction of 3.16% in EER was observed, however, the overall error was significantly higher than many other neural network and pressure-based applications (11.78% EER).

Other unique approaches that used pressure-based measurements were also implemented in systems such as by Nguyen et al. (2010). Here, a microphone was used to record the sounds produced by the keystrokes. The data from the microphone were then used to create a standardized “bio-matrix” detailing the keystroke timing and force, with data becoming extracted via an independent component analysis (ICA) routine. ICA extracted the data from the bio-matrix, and the Fast Artificial Neural Network library (FANN) was used for recognition and authentication. This technique proved to be competitive in terms of accuracy, achieving an FAR of 4.12% and an FRR of 5.55%. Furthermore, the use of a microphone represents a novel technique for acquiring pressure measurement, which could be much cheaper to implement than the alternative methods mentioned above. One obstacle to the use of microphones is that the results would be easily affected in the presence of noise – although it is fair to say that intelligent noise cancelation techniques are becoming main-stream even on civilian devices such as mobile phones.

Table 1 serves to summarize, in terms of input demand, analysis methods employed, and respective accuracy rates, for several key examples from the various typing biometric approaches used thus far.

Handheld devices and mobile phones

With the large increase in the use, access, and ownership of mobile phones, the protection of personal and sensitive information within such devices is an obvious concern and authentication using keystroke dynamics could be a suitable addition to the current security measures. Such handheld devices have a number of limitations in terms of security (Jansen, 2003):

- Due to the small size, devices are easily stolen or misplaced.
- User authentication is by default disabled.
- Authentication systems on such devices can be very limited and easily deceived.

Keystroke dynamics analysis on such handheld and mobile devices could be somewhat more limited than that of a computer. It is also important to remember that most users do not type as often on mobile phones as they do on computers and so the detection of unique signatures could be more difficult. Moreover, the preferred typing style (with thumbs or one finger only) may not be directly correlated with standard keyboard operation.

Clarke and Furnell (2007) investigated the use of keystroke dynamics in the application of mobile phones. They made use of the numerical keypad on a large number of mobile phones before touch screens were introduced, and tested a number of neural network-based analysis techniques: feed-forward MLP (FF MLP); radial basis function (RBF); and generalized regression neural networks, finding the FF MLP network to be the most stable and useful in this case.

When acquiring samples for a numerical system, two sample sizes were used of four and eleven numbers. These string lengths were chosen as common PINs used to lock phones for security are often four numbers in length, and phone numbers themselves can be of lengths up to eleven numbers. Alphabetic input classification was conducted using samples from participants who were asked to type thirty text messages consisting of mixtures of quotes, lines from movies and typical text messages. In the case of typing letters on such first generation devices, keys had to be pressed multiple times to acquire the correct letters. Impressively, the study by Clarke and Furnell (2007) combined not only keystroke analysis but also voice, facial, and fingerprint recognitions, attaining very high accuracies. However, such systems require more mobile capabilities (camera or fingerprint reader) and a significant level of processing on the mobile phone.

In this context, Saevanee and Bhatarakosol (2008) used a k-NN approach with data (hold and inter-key times) from a numerical touchpad and were able to achieve accuracies of 99.9% with pressure measurements alone. A similar study (Saevanee and Bhatarakosol, 2009) used a probabilistic neural network (PNN) and achieved comparable results. The significance of the result is, however, once again tempered by the low subject numbers involved (only 10 participants with sample sizes of 10 characters measured at 20 ms intervals), while the stated accuracy using PNN (99%) is higher than that of others using different analyses. For example, Campisi et al. (2009) conducted keystroke dynamics analysis on mobile phones with telephone keypads, achieving an EER of 13%. A statistical analysis technique was implemented making use of four key hold and latency times

Table 1 | Summary of salient typing demand, analysis mode, and accuracy rates for a spectrum of different keystroke biometric approaches.

Typing input demand	Method of analysis	Accuracy	Reference
10 character string input 10 times with 30 participants	Statistical (\bar{x} and ρ)	FAR = 1.89% FRR = 1.45%	Araújo et al. (2005)
Circa 40 character string input 10 times with 100 participants	Statistical (GPD fused with DSM)	EER \approx 1%	Teh et al. (2011)
Circa 30 character string input 10 times with eight participants	Statistical (GMM)	FAR = 2.1%; FRR = 2.4%EER < 3%	Hosseinzadeh et al. (2006)
Short phrase entry with six participants	Artificial Neural Net	Accuracy = 97.8%	Obaidat and Macchiarolo (1993)
15 valid and 15 invalid users \times 225 sequence	ANN + Fuzzy logic	EER = 0%	Obaidat and Sadoun (1997)
Short password entered three times with 90 valid and 61 imposter participants	Multilayer back propagated ANN	FAR = 1.1%; FRR = 0%*	Lin (1997)
7 character string input between 150 and 400 times with 25 participants	ANN using multilayer perceptron	FAR = 0%; FRR = 1%	Cho et al. (2000)
At least 8 character string input 25 times with 29 participants to study	Fuzzy logic	FAR = 2.79%; FRR = 7.379%	De Ru and Eloff (1997)
683 character string using 154 participants	Statistical - trigram-based	FRR = 4%; FAR = 0.01%	Bergadano et al. (2002)
Short ($n < 15$ characters) strings input 10 times with 23 participants	Auto-regressive classifier linked to pressure data	EER \approx 3%	Eltahir et al. (2008)
10 character password input to database enrolment with 50 samples (30 genuine and 20 forged)	Statistical; & Statistical augmented with pressure data	EER = 2.04%; EER = 1.41% (P-augmented)	Lv and Wang (2006)
8 character string with 10 timing- and 10 pressure vectors recorded	Artificial Neural Net augmented with pressure data	EER values of 16.5, 14.94, and 11.78% for respectively, pressure, latency, and pressure + latency	Loy et al. (2007)
Short string pairs input 15 times with 20 participants	Independent component analysis and fast-ANN augmented with acoustic record	FAR = 4.12%; FRR = 5.55%	Nguyen et al. (2010)

*With refined thresholding.

for the typing of six 10-character passwords which were each repeated 20 times. The stated EER achieved was relatively high in comparison to implementation on a full keyboard which, using statistical techniques, typically report EERs of under 5% (see above).

Hwang et al. (2009a) applied keystroke dynamics analysis to four number PINs for mobile phones. Twenty-five participants took part and two different approaches were investigated, “Natural Rhythm without Cue” and “Artificial Rhythms with Cues.” The best results were achieved when the participants were required to use artificial rhythms – which reduced the EER to around 4%. A follow-up study by this same group into artificial rhythms (Hwang et al., 2009b) further elucidated the effects of pauses with cues, and attained sub 2% error rates. Chang et al. (2011) conducted a similar study investigating the feasibility of “click rhythm” based systems using mouse clicks, with EERs below 8%.

Keypads

Naturally, when considering the use of new security measures, keypad systems are important due to their current use in cash withdrawal systems or for controlling access to secure areas. Mantjarvi et al. (2002) designed and made use of an unconventional keypad system. Their system implements an infrared receiver and transceiver system as a substitute for a button-based numerical input system. They then implemented an MLP and a k-NN algorithm to attempt keystroke verification. The accuracy achieved was affected by the implementation of this unique system, achieving classification results of 78–99% for k-NN, and 69–96% for MLP (the authors did not, however, provide details of the test data).

Using a similar setting, Rodrigues et al. (2005) used two analysis techniques to authenticate users using a numerical keypad, i.e., a statistical classifier and pattern recognition using a HMM. The statistical classifier exploited the means and standard deviations of

keystroke timings and these were compared to any samples being tested by a measure of distance. The HMM produced the lowest error rate of 3.6% (EER) and although this is comparable to some error rates achieved by HMMs with full keyboards, the use of only the numerical keypad reduces the number of keys being pressed, making this finding relevant for implementation in actual keypad systems.

Grabham and White (2008) conducted similar tests, using the variables of applied force and key-press duration, which were coupled with a component-wise verification scheme and which resulted in a higher EER (~10%) when using an actual ATM keypad with individual force sensing devices beneath every key. Importantly, the keypad was designed to look and operate identically to an orthodox keypad system to ensure validity of the approach with a real-world scenario.

NOVEL AND FUTURE APPLICATIONS

The field of keystroke dynamics has many other areas of use other than authentication. Lv et al. (2008) used pressure-based keystroke analysis for a completely novel application, where the pressure wave component was used as a technique for the detection of emotion. Fifty participants took part in their study, and were subjected to six different emotion inductions (neutral, anger, fear, happiness, sadness, and surprise) providing a total of 3000 samples and obtaining an accuracy of 93.4%. To induce the emotions, the subjects were asked to listen to and watch a short story for each emotion and immerse themselves in the situation when typing. Each individual emotional state was shown to produce a different pressure sequence. To analyze these different emotional states some initial pre-processing was needed (noise removal and normalization) and then three analysis techniques were fused together, including two pressure analysis approaches and one traditional keystroke approach. The two pressure analysis techniques included the analysis of Global Features of the pressure sequence and dynamic time warping as with Lv's study (2006). The analysis was shown to be effective for these particular six emotions and as such, emotional state detection could have uses for many fields.

Lv et al. (2008) report that this emotional recognition system was used for intelligent game control and other applications. Feedback from a computer system based upon a user's emotional state could be an interesting area of application, however, this research direction is still very much in its infancy. We suggest that the use of such an emotional recognition system could be relevant for controlling access to secure systems, in that emotional states such as anger or fear might be associated with critical states of the user that could potentially be monitored.

Other than the above analysis of emotional states for detection of different emotions, a similar analysis could also be applied for the detection of deception. Such a system could obtain a reference or baseline signature for a user and then, using keystroke data, attempt to identify when a user could be trying to deceive the system. For such categorization, a measure of the stress that the user is experiencing could be detected and analyzed. Investigation into such applications could use a greater number of variables than the typical keystroke analysis systems, as such measurements could increase the accuracy of the detection. Such analysis would most likely not be completed with an average keyboard, especially

when pressure is a measure and so a more technologically advanced keyboard design is required.

Future keystroke analysis authentication tools could take to the Internet as web-based security systems for aiding in the security of online accounts and systems. For such systems to work effectively, they need to be able to complete keystroke analysis not only on traditional keyboards but also on touch-based devices. Investigation into the relationship between keystroke signatures obtained with traditional keyboards and those captured with touch-based systems could prove extremely useful. With the number of touch-based systems and tablet computers increasing rapidly in the last few years, such research could help to create a universal signature that could be used across platforms without need for multiple input data to each sensor. This research could lead to the development of such web-based keystroke analyses tools being a great deal more flexible in their use and ability.

CONCLUSION

The application of keystroke dynamics to authentication has met with some compelling success, yet the standards continue to evolve in the drive toward optimal reliability. The accuracies achieved have reached heights of 99% with multiple techniques and with several data sets, proving that the use of such techniques would be valid and beneficial additions to current security systems. The analysis techniques used include statistical, neural network, and fuzzy logic approaches, and the inclusion of new parameter spaces such as pressure variables. The main variables against which the quality of the authentication systems have been measured are FAR, FRR, or EERs, which are ultimately the main indicators of the success of a biometric system. However, a comparison of different authentication methods based on these standard error rates is still challenging because of the heterogeneity of timing variables recorded (e.g., down-up, down-down, up-down, up-up times, digraphs, trigraphs, etc.). A comparison of different classifiers for user authentication appears to be only useful to the extent that they rely on the same variables.

Regarding the actual application of biometric systems, we conclude that ease of manner of enrolment should be a critical factor in determining the choice of a system, as this affects the practicality of the suggested biometric approach. For example, a number of the reviewed studies (Cho et al., 2000; Araújo et al., 2005; Lee and Cho, 2005) relied on imposter login attempts to refine the biometric system. The use of imposter data allows the specification of a more refined user profile and might be reasonable in the context of applications in which the user might expect to go through a specific enrolment procedure (e.g., access to secure military systems). However, relying on this approach is less practicable for systems that are used by standard, non-specialist users, as the ease with which individuals can be enrolled in a biometric authentication system becomes more relevant. A quick enrolment procedure using as few password and username characters would be preferable, however, few characters make the system more susceptible to classification errors. The balance of error rates and ease-of-use thus needs to be carefully determined, depending on the severity of the consequences of breaching a secure system.

Associated with this aspect is also the actual context of user enrolment. Enrolling via a server (e.g., Bleha et al., 1990;

Obaidat and Macchiarolo, 1993; Tapiador and Sigüenza, 1999), which could be an option for online banking, for example, shifts the responsibility of “proper” enrolment to the user. In a situation where the enrolment process is not controlled (e.g., accomplished in a structured environment and/or supervised by trained staff) the enrolment data might be “noisy,” thus increasing the likelihood of authentication errors. With emerging advances in authentication algorithms and technological developments, as well as sufficiently reliable systems, we would expect an increase in the actual implementation of such systems in the “real-world.” This also implies that the user-friendliness of such systems becomes more important for determining the success of the biometric application.

Other than the use of keystroke dynamics analysis with traditional keyboards, similar investigations have been carried out with other input devices such as touch screens and keypads. These used similar analysis techniques and were able to achieve accuracies close to those with full keyboards showing the applicability of this field to a range of devices and systems. Coinciding with an emerging interest in affective computing (Picard, 2000), keystroke analysis has also been implemented for other purposes, such as the detection of emotions. However, more research is needed in this avenue in order to achieve the maturity and reliability that traditional orthodox methodologies have achieved.

AUTHOR CONTRIBUTIONS

Nasir Ahmad conducted the associated lab-work that informed this paper, and wrote the draft, both under guidance from Paul A. Campbell. Andrea Szymkowiak and Paul A. Campbell corrected and updated the manuscript.

ACKNOWLEDGMENTS

The authors wish to thank the Royal Society’s Industry Research Fellowship scheme (IF09010).

REFERENCES

- Araújo, L. C., Sucupira L. H. Jr., Lizarraga, M. G., Ling, L. L., and Yabu-Uti, J. B. T. (2005). User authentication through typing biometrics features. *IEEE Trans. Signal Process.* 53, 851–855. doi: 10.1109/TSP.2004.839903
- Bergadano, F., Gunetti, D., and Picardi, C. (2002). User authentication through keystroke dynamics. *ACM Trans. Inf. Syst. Secur. (TISSEC)* 5, 367–397. doi: 10.1145/581271.581272
- Bleha, S., Slivinsky, C., and Hussien, B. (1990). Computer-access security systems using keystroke dynamics. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 1217–1222. doi: 10.1109/34.62613
- Bleha, S. A., and Obaidat, M. S. (1991). Dimensionality reduction and feature extraction applications in identifying computer users. *IEEE Trans. Syst. Man Cybern.* 21, 452–456. doi: 10.1109/21.87093
- Campisi, P., Maiorana, E., Lo Bosco, M., and Neri, A. (2009). User authentication using keystroke dynamics for cellular phones. *IET Signal Process.* 3, 333–341. doi: 10.1049/iet-spr.2008.0171
- Chang, W. (2005). “Improving hidden Markov models with a similarity histogram for typing pattern biometrics,” in *Proceedings of the IEEE International Conference on Information Reuse and Integration*, Las Vegas, 487–493.
- Chang, T. Y., Tsai, C. J., and Lin, J. H. (2012). A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices. *J. Syst. Softw.* 85, 1157–1165. doi: 10.1016/j.jss.2011.12.044
- Chang, T. Y., Tsai, C. J., Yang, Y. J., and Cheng, P. C. (2011). “User authentication using rhythm click characteristics for non-keyboard devices,” in *Proceedings of the 2011 International Conference on Asia Agriculture and Animal IPCBEE*, Vol. 13, Hong Kong, 167–171.
- Changshui, Z., and Yanhua, S. (2000). AR model for keystroke verification. *IEEE Int. Conf. Syst. Man Cybern.* 4, 2887–2890.
- Cho, S., Han, C., Han, D. H., and Kim, H. I. (2000). Web-based keystroke dynamics identity verification using neural network. *J. Organ. Comput. Electron. Commer.* 10, 295–307. doi: 10.1207/S15327744OCE1004_07
- Clarke, N. L., and Furnell, S. M. (2007). Authenticating mobile phone users using keystroke analysis. *Int. J. Inf. Secur.* 6, 1–14. doi: 10.1007/s10207-006-0006-6
- Coltell, O., Badfa, J. M., and Torres, G. (1999). “Biometric identification system based on keyboard filtering,” in *Proceedings of the IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology*, Madrid, 203–209.
- Curtin, M., Tappert, C., Villani, M., Ngo, G., Simone, J., Fort, H. S., et al. (2006). “Keystroke biometric recognition on long-text input: a feasibility study,” in *Proceeding International Workshop Scientific Computing and Computational Statistics (IWSCCS 2006)*, Hong Kong.
- De Ru, W. G., and Eloff, J. H. (1997). Enhanced password authentication through fuzzy logic. *IEEE Expert* 12, 38–45. doi: 10.1109/64.642960
- D’Souza, D. (2002). *Typing Dynamics Biometric Authentication*. Available at: <http://innovexpo.itce.uq.edu.au/2002/projects/s373901/thesis.PDF> (accessed July 30, 2013).
- Eltahir, W. E., Lai, W. K., Salami, M. J. E., and Ismail, A. F. (2003). “Hardware design, development and evaluation of a pressure-based typing biometric authentication system,” in *Proceedings of the Eighth Australian and New Zealand Intelligent Information Systems Conference (ANZIIS)*, Sydney, 49–54.
- Eltahir, W. E., Salami, M. J. E., Ismail, A. F., and Lai, W. K. (2004). “Dynamic keystroke analysis using AR model,” in *Proceedings of the IEEE International Conference on Industrial Technology (IEEE-ICIT04)*, Hammamet, 1555–1560.
- Eltahir, W. E., Salami, M. J. E., Ismail, A. F., and Lai, W. K. (2008). Design and evaluation of a pressure-based typing biometric authentication system. *EURASIP J. Inf. Secur.* 2008, Article ID 345047. doi: 10.1155/2008/345047
- Gaines, R. S., Lisowski, W., Press, S. J., and Shapiro, N. (1980). *Authentication by Keystroke Timing: Some Preliminary Results*. Technical Report R-2526-NSF. Santa Monica, CA: Rand Corporation.
- Giot, R., El-Abed, M., Hemery, B., and Rosenberger, C. (2011). Unconstrained keystroke dynamics authentication with shared secret. *Comput. Secur.* 30, 427–445. doi: 10.1016/j.cose.2011.03.004
- Grabham, N. J., and White, N. M. (2008). “Use of a novel keypad biometric for enhanced user identity verification,” in *IEEE International Conference on Instrumentation and Measurement Technology*, Victoria, 12–16.
- Haider, S., Abbas, A., and Zaidi, A. K. (2000). A multi-technique approach for user identification through keystroke dynamics. *IEEE Int. Conf. Syst. Man Cybern.* 2, 1336–1341. doi:10.1109/ICSMC.2000.886039
- Henderson, N. J., White, N. M., Veldhuis, R. N. J., Hartel, P. H., and Slump, C. H. (2002). “Sensing pressure for authentication,” in *Proceedings of 3rd IEEE Benelux Signal Processing Symposium (SPS-2002)*, Leuven.
- Hocquet, S., Ramel, J. Y., and Cardot, H. (2007). User classification for keystroke dynamics authentication. *Adv. Biom. LNCS* 4642, 531–539. doi: 10.1007/978-3-540-74549-5_56
- Hosseinzadeh, D., Krishnan, S., and Khademi, A. (2006). “Keystroke identification based on Gaussian mixture models,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, 3, III.1144–III.1147.
- Hwang, S. S., Cho, S., and Park, S. (2009a). Keystroke dynamics-based authentication for mobile devices. *Comput. Secur.* 28, 85–93. doi: 10.1016/j.cose.2008.10.002
- Hwang, S. S., Lee, H. J., and Cho, S. (2009b). Improving authentication accuracy using artificial rhythms and cues for keystroke dynamics-based authentication. *Expert Syst. Appl.* 36, 10649–10656. doi: 10.1016/j.eswa.2009.02.075
- Jansen, W. (2003). “Authenticating users on handheld devices,” in *Proceedings of the Canadian Information Technology Security Symposium*, Ottawa.
- Joyce, R., and Gupta, G. (1990). Identity authentication based on keystroke latencies. *Commun. ACM* 33, 168–176. doi: 10.1145/75577.75582
- Karnan, M., Akila, M., and Krishnaraj, N. (2011). Biometric personal authentication using keystroke dynamics. *Rev. Appl. Soft Comput.* 11, 1565–1573. doi: 10.1016/j.asoc.2010.08.003
- Killourhy, K. S., and Maxion, R. A. (2009). “Comparing anomaly-detection algorithms for keystroke dynamics,” in *IEEE/IFIP International Conference on Dependable Systems & Networks*, Estoril, 125–134.

- Lee, H. J., and Cho, S. (2005). "Retraining a novelty detector with impostor patterns for keystroke dynamics-based authentication," in *Advances in Biometrics*, Vol. 3832, *Lecture Notes in Computer Science* (Berlin: Springer), 633–639.
- Lin, D. T. (1997). Computer-access authentication with neural network based keystroke identity verification. *IEEE Int. Conf. Neural Netw.* 1, 174–178.
- Loy, C. C., Lai, W. K., and Lim, C. P. (2007). "Keystroke patterns classification using the ARTMAP-FD neural network," in *IEEE Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Kaohsiung, 1, 61–64.
- Lv, H. R., Lin, Z. L., Yin, W. J., and Dong, J. (2008). "Emotion recognition based on pressure sensor keyboards," in *IEEE International Conference on Multimedia and Expo*, Toronto, 1089–1092.
- Lv, H. R., and Wang, W. Y. (2006). Biologic verification based on pressure sensor keyboards and classifier fusion techniques. *IEEE Trans. Consum. Electron.* 52, 1057–1063. doi: 10.1109/TCE.2006.1706507
- Mantjarvi, J., Koivumaki, J., and Vuori, P. (2002). Keystroke recognition for virtual keyboards. *Proc. IEEE Int. Conf. Multimed. Expo 2*, 429–432. doi: 10.1109/ICME.2002.1035630
- Monrose, F., Reiter, M. K., and Wetzel, S. (1999). "Password hardening based on keystroke dynamics," in *Proceedings of the 6th ACM Conference on Computer and Communications Security*, New York, NY, 73–82.
- Monrose, F., and Rubin, A. (1997). "Authentication via keystroke dynamics," in *Proceedings of the 4th ACM Conference on Computer and Communications Security*, New York, NY 48–56.
- Monrose, F., and Rubin, A. D. (2000). Keystroke dynamics as a biometric for authentication. *Future Gener. Comput. Syst.* 16, 351–359. doi: 10.1016/S0167-739X(99)00059-X
- Nguyen, T. T., Le, T. H., and Le, B. H. (2010). "Keystroke dynamics extraction by independent component analysis and bio-matrix for user authentication," in *PRICAI 2010: Trends in Artificial Intelligence*, Vol. 6230, *Lecture Notes in Computer Science* (Berlin: Springer), 477–486.
- Nonaka, H., and Kurihara, M. (2004). "Sensing pressure for authentication system using keystroke dynamics," in *International Conference on Computational Intelligence*, Istanbul, 19–22.
- Obaidat, M. S., and Macchiarolo, D. T. (1993). An online neural network system for computer access security. *IEEE Trans. Ind. Electron.* 40, 235–242. doi: 10.1109/41.222645
- Obaidat, M. S., and Macchiarolo, D. T. (1994). A multilayer neural network system for computer access security. *IEEE Trans. Syst. Man Cybern.* 24, 806–813. doi: 10.1109/21.293498
- Obaidat, M. S., and Sadoun, B. (1997). Verification of computer users using keystroke dynamics. *IEEE Trans. Syst. Man Cybern. B Cybern.* 27, 261–269. doi: 10.1109/3477.558812
- Obaidat, M. S., and Sadoun, B. (1999). "Keystroke dynamics based authentication," in *Biometrics. Personal Identification in Networked Society*, eds A. Jain, R. Bolle, and S. Pankanti (Norwell, MA: Kluwer Academic Publishers), 213–229.
- Peacock, A., Ke, X., and Wilkerson, M. (2004). Typing patterns: a key to user identification. *IEEE Secur. Priv.* 2, 40–47. doi: 10.1109/MSP.2004.89
- Picard, R. W. (2000). *Affective Computing*. Cambridge: MIT Press.
- Revett, K. (2009). A bioinformatics based approach to user authentication via keystroke dynamics. *Int. J. Control Autom. Syst.* 7, 7–15. doi: 10.1007/s12555-009-0102-2
- Robinson, J. A., Liang, V. W., Chambers, J. M., and MacKenzie, C. L. (1998). Computer user verification using login string keystroke dynamics. *IEEE Trans. Syst. Man Cybern. A Syst. Hum.* 28, 236–241. doi: 10.1109/3468.661150
- Rodrigues, R. N., Yared, G. F., Costa, C. R. D. N., Yabu-uti, J. B., Violaro, F., and Ling, L. L. (2005). Biometric access control through numerical keyboards based on keystroke dynamics. *Adv. Biom. LNCS* 3832, 640–646. doi: 10.1007/11608288_85
- Saevanee, H., and Bhatarakosol, P. (2008). "User authentication using combination of behavioural biometrics over the touchpad acting like touch screen of mobile device," in *IEEE International Conference on Computer and Electrical Engineering*, Washington, 82–86.
- Saevanee, H., and Bhatarakosol, P. (2009). "Authenticating user using keystroke dynamics and finger pressure" in *Proceedings of the 6th IEEE Consumer Communications and Networking Conference*, Las Vegas, 1–2.
- Tapiador, M., and Sigüenza, J. A. (1999). "Fuzzy keystroke biometrics on web security," in *IEEE Workshop on Automatic Identification Advanced Technologies*, Morristown, 133–136.
- Teh, P. S., Teoh, A. B. J., Tee, C., and Ong, T. S. (2011). A multiple layer fusion approach on keystroke dynamics. *Pattern Anal. Appl.* 14, 23–36. doi: 10.1007/s10044-009-0167-9
- Umphress, D., and Williams, G. (1985). Identity verification through keyboard characteristics. *Int. J. Man Mach. Stud.* 23, 263–273. doi: 10.1016/S0020-7373(85)80036-5
- Wong, F. W. M. H., Supian, A. S. M., Ismail, A. F., Kin, L. W., and Soon, O. C. (2001). "Enhanced user authentication through typing biometrics with artificial neural networks and k-nearest neighbor algorithm," in *Proceedings of the IEEE Conference on Signals Systems and Computers*, Pacific Grove, CA, 2, 911–915.
- Wood, H. M. (1977). *The Use of Passwords for Controlled Access to Computer Resources* (Vol. 500, No. 9). US Department of Commerce, National Bureau of Standards, Gaithersburg, 500–509.
- Xi, K., Tang, Y., and Hu, J. (2011). Correlation keystroke verification scheme for user access control in cloud computing environment. *Comput. J.* 54, 1632–1644. doi: 10.1093/comjnl/bxr064
- Young, J. R., and Hammon, R. W. (1989). Method and apparatus for verifying an individual's identity, Patent No. 4805222, U.S. Patent and Trademark Office, Washington, 1989.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 August 2013; accepted: 18 November 2013; published online: 19 December 2013.

Citation: Ahmad N, Szymkowiak A and Campbell PA (2013) Keystroke dynamics in the pre-touchscreen era. *Front. Hum. Neurosci.* 7:835. doi: 10.3389/fnhum.2013.00835

This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Ahmad, Szymkowiak and Campbell. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Right inferior frontal gyrus activation as a neural marker of successful lying

Oshin Vartanian^{1,2*}, Peter J. Kwantes^{1,3}, David R. Mandel^{1,4}, Fethi Bouak¹, Ann Nakashima¹, Ingrid Smith¹ and Quan Lam¹

¹ Defence Research and Development Canada, Toronto, ON, Canada

² Department of Psychology, University of Toronto–Scarborough, Toronto, ON, Canada

³ School of Psychology, University of Queensland, Brisbane, QLD, Australia

⁴ Department of Psychology, York University, Toronto, ON, Canada

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

Alberto Priori, Università di Milano, Italy

Nobuhito Abe, Kyoto University, Japan

*Correspondence:

Oshin Vartanian, DRDC Toronto, 1133 Sheppard Avenue West, Toronto, ON M3K 2C9, Canada
e-mail: oshin.vartanian@drdc-rddc.gc.ca

There is evidence to suggest that successful lying necessitates cognitive effort. We tested this hypothesis by instructing participants to lie or tell the truth under conditions of high and low working memory (WM) load. The task required participants to register a response on 80 trials of identical structure within a 2 (WM Load: high, low) × 2 (Instruction: truth or lie) repeated-measures design. Participants were less accurate and responded more slowly when WM load was high, and also when they lied. High WM load activated the fronto-parietal WM network including dorsolateral prefrontal cortex (PFC), middle frontal gyrus, precuneus, and intraparietal cortex. Lying activated areas previously shown to underlie deception, including middle and superior frontal gyrus and precuneus. Critically, successful lying in the high vs. low WM load condition was associated with longer response latency, and it activated the right inferior frontal gyrus—a key brain region regulating inhibition. The same pattern of activation in the inferior frontal gyrus was absent when participants told the truth. These findings demonstrate that lying under high cognitive load places a burden on inhibition, and that the right inferior frontal gyrus may provide a neural marker for successful lying.

Keywords: deception, lying, inhibition

A substantial body of behavioral evidence—collected both in the psychological laboratory as well as during police interviews—suggests that lying requires effort (Vrij et al., 2006, 2010). Given this observation, one potential strategy for catching a liar or detecting a lie would be to increase a suspect's cognitive load. To the extent that limited cognitive resources—including working memory (WM) and executive functions—are depleted, so is their availability to aid a liar to maintain a lie (see Vrij et al., 2010). Indeed, previous studies have demonstrated that a number of methodologies known to increase cognitive load are effective in helping to detect lies, including requiring subjects to maintain continuous eye contact (Beattie, 1981), asking questions that are irrelevant to some focal event (Quas et al., 2007), and instructing suspects to recall events in reverse order (Vrij et al., 2008). The present experiment was designed to test the hypothesis that a *direct* manipulation of WM load based on a variation of Sternberg's (1966) classic short-term memory paradigm will achieve the same result. Specifically, it will be more effortful to lie successfully when WM load is high compared to when WM load is low—as measured by an increase in response time (RT). Compared to previous approaches, the most salient feature of this technique is that the manipulation of WM load is non-verbal, and it can be implemented with ease on a trial-by-trial basis.

However, it has also been shown that the exertion of effort while lying could have multiple sources such as, the formulation of a lie, lie activation, self-monitoring of behavior, monitoring of

the interviewer's behavior, truth suppression, and the implementation of reminders to lie (Vrij et al., 2010). This means that in addition to measures of cognitive effort, additional metrics are necessary to identify the source of the effort. One type of evidence that can be gainfully employed for this purpose is brain activation data, although the utility of brain imaging data depends on the specificity of the cognitive processes associated with the activated regions (Poldrack, 2006). In the context of the present study, the cognitive process that we were particularly interested in was inhibition, and its widely accepted role in truth suppression (e.g., Langleben et al., 2002). To determine whether the added burden on inhibition contributes to the increased effort while lying, we turned to data collected in the functional magnetic resonance imaging (fMRI) scanner. At the neural level, inhibition has been shown to be reliably correlated with activation in the inferior frontal gyrus (Aron et al., 2004), bolstered further by neuropsychological evidence demonstrating that persons with damage to this region are impaired at inhibition tasks (Aron et al., 2003). This evidence points to the inferior frontal gyrus as a likely candidate region for regulating inhibition during lying.

Vartanian et al. (2012) recently demonstrated that lying is correlated with increased activation in the WM network. They found that the inferior frontal gyrus was activated more in successful liars than in less-skilled ones. Based on scores taken from one condition of the task in which high variability in performance was observed, an independent samples *t*-test between good and

poor liars revealed a significant difference in activation exclusive to the right inferior frontal gyrus (BA 44). Furthermore, a regression in which lying accuracy was regressed onto activation in the right inferior frontal gyrus demonstrated that activation in the right inferior frontal gyrus was a reliable predictor of lying accuracy, accounting for 29% of the observed variance in performance. The result suggests that individual differences in people's ability to suppress the truth (as measured by activity in the right inferior frontal gyrus) is an important predictor of lying skill.

Extending from the approach employed by Vartanian et al. (2012), and based on Vrij et al.'s (2010) conjecture about how taxing cognitive load might help identify liars, we examined how the inferior frontal gyrus might be activated when participants lied successfully under high and low WM load, and compared it to when they were instructed to tell the truth under the same WM load conditions. Under instructions to tell the truth, participants do not need to suppress a truthful response. Without the need to suppress a response, we did not expect an increase in WM load to have a strong impact on the activation of the inferior frontal gyrus. On the other hand, to the extent that depleting limited WM's resources increases inhibitory workload (Vrij et al., 2010), we predicted that the inferior frontal gyrus would experience substantially higher activation when participants lied under a high WM load than when lies were committed under low WM load.

MATERIALS AND METHOD

PARTICIPANTS

This research proposal was approved by DRDC's Human Research Ethics Committee and Sunnybrook Health Sciences Centre's Research Ethics Board. The participants were 15 neurologically healthy right-handed volunteers (1/3 female, age range 19–48 years) with normal or corrected-to-normal vision.

STIMULI AND PROCEDURE

The task required participants to register a response on 80 trials of identical structure within a 2 (WM load, high or low) \times 2 (Instruction: truth or lie) repeated-measures design (Figure 1). Trials involving the instruction to lie were distributed equally among match and no-match stimuli, resulting in 20 trials in each of the four conditions. The trial structure involved a modification of Sternberg's (1966) classic short-term memory paradigm, wherein participants are presented with a sequence of symbols (e.g., letters or digits) that must be encoded into memory, followed after a delay with the presentation of a test stimulus (i.e., a letter or a digit). The participant's task is to decide whether

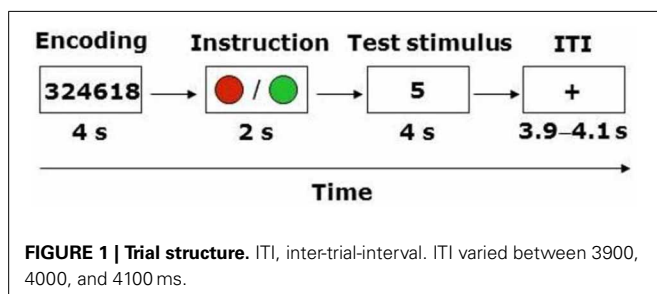
the test stimulus matches one of the symbols in the sequence presented earlier. The standard finding from that literature indicates that there is a linear relationship between the mean reaction time (RT) to make this decision and the length of the sequence. In the present experiment, each trial began with the presentation of a four- or six-digit string for 4 s. Participants were instructed to encode this string into memory. At the end of the trial the participant was presented with a test stimulus (i.e., a digit), and had to decide whether it matched one of the digits in the string within a 4 s response window. The variation in the length of the sequence (i.e., four vs. six digits) represented our WM manipulation. Notably, however, immediately following the presentation of the digit string participants were presented with a cue for 2 s that instructed them to either report truthfully or to lie about whether the test stimulus matched one of the digits in the string. In the truth condition the cue appeared as a green circle, whereas in the lie condition it appeared as a red circle.

Thus, the total duration of each trial was 10 s, and successive trials were interspersed with a fixation point with variable inter-trial interval (ITI, 3900, 4000, or 4100 ms averaged at 4 s across all trials). Participants recorded their responses using an MRI-compatible keypad that had separate keys labeled "match" and "mismatch." Match and mismatch responses were registered using the index and middle finger of the same hand. The hand used to enter responses as well as the keys (for match and mismatch) were counterbalanced across participants. In the scanner, the 80 trials were presented in a single run. The order of trials was randomized for each participant. The duration of the task was 18 min and 40 s (80 trials \times 14 s). Prior to entry into the scanner participants completed 10 practice trials to familiarize themselves with the task.

fMRI ACQUISITION AND ANALYSIS

A 3-Tesla MR scanner with an 8-channel head coil (Discovery MR750, 22.0 software, GE Healthcare, Waukesha, WI) was used to acquire T1 anatomical volume images ($0.86 \times 0.86 \times 1.0$ mm voxels). For functional imaging, T2*-weighted gradient echo spiral-in/out acquisitions were used to produce 26 contiguous 5 mm thick axial slices [repetition time (TR) = 2000 ms; echo time (TE) = 30 ms; flip angle (FA) = 70°; field of view (FOV) = 200 mm; 64×64 matrix; voxel dimensions = $3.1 \times 3.1 \times 5.0$ mm], positioned to cover the whole brain. The spiral sequence was acquired sequentially. The first five volumes were discarded to allow for T1 equilibration effects, leaving 560 volumes for analysis.

Data were analyzed using Statistical Parametric Mapping (SPM8). Head movement was less than 2 mm in all cases. All functional volumes were spatially realigned to the first volume. A mean image created from realigned volumes was spatially normalized to the Montreal Neurological Institute's echo-planar imaging (MNI EPI) brain template using non-linear basis functions. The derived spatial transformation was applied to the realigned T2* volumes, and spatially smoothed with an 8 mm full-width half-maximum (FWHM) isotropic Gaussian kernel. Time series across each voxel were high-pass filtered with a cut-off of 128 s, using cosine functions to remove section-specific



low frequency drifts in the BOLD signal. Condition effects at each voxel were estimated according to the GLM and regionally specific effects compared using linear contrasts. The BOLD signal was modeled as a box-car, convolved with a canonical hemodynamic response function. Each contrast produced a statistical parametric map consisting of voxels where the z -statistic was significant at $p < 0.001$. Reported activations survived voxel-level intensity threshold of $p < 0.001$ (uncorrected for multiple comparisons) at the voxel level and $p < 0.05$ (uncorrected for multiple comparisons) at the cluster level using a random-effects model.

RESULTS

BEHAVIORAL

Mean correct RT and percent correct for each condition are shown in **Table 1**. Mean correct RT across all conditions was 1452 ms ($SEM = 80$). Skewness and kurtosis of the RT distribution did not deviate from normality (both $ps > 0.05$). A WM load \times Instruction repeated-measures ANOVA showed the predicted main effect for WM such that RT was longer in the high load condition than in the low load condition, $F_{(1, 14)} = 15.53$, $p < 0.001$, partial $\eta^2 = 0.53$. As well, we observed the predicted main effect for Instruction in which RT was longer in the lie condition than in the truth condition, $F_{(1, 14)} = 55.03$, $p < 0.001$, partial $\eta^2 = 0.80$ (**Table 1**). The interaction between WM Load and Instruction was not reliable, $F_{(1, 14)} = 0.45$, $p = 0.51$, partial $\eta^2 = 0.03$.

Mean accuracy across all conditions was 94.4% ($SEM = 0.01$). Skewness and kurtosis of the accuracy distribution did not deviate from normality (both $ps > 0.05$). A WM load \times Instruction repeated-measures ANOVA showed the predicted main effect for WM load: accuracy was lower in the high load condition than in the low load condition, $F_{(1, 14)} = 8.61$, $p < 0.05$, partial $\eta^2 = 0.38$. As well, we observed the predicted main effect for instruction: accuracy was lower in the lie condition than in the truth condition, $F_{(1, 14)} = 7.32$, $p < 0.05$, partial $\eta^2 = 0.34$ (**Table 1**). The interaction between WM Load and Instruction was not reliable, $F_{(1, 14)} = 0.45$, $p = 0.56$, partial $\eta^2 = 0.03$.

fMRI

Using an event-related design, at the first level of analysis (i.e., subject level in SPM8) we specified regressors corresponding to the four conditions, as well as ITI and motor response. Although incorporated into the design, ITI and motor response were modeled out of the analyses by assigning null weights to their regressors. Given the main effect of WM load, we investigated the direct contrast of high vs. low WM load. This demonstrated significant activation in the

middle frontal gyrus, precuneus, intraparietal sulcus, supplementary motor area, caudate, dorsolateral PFC, and cerebellum (**Table 2**). This pattern is consistent with the well-established role of the frontoparietal system in WM, and indeed specifically as observed within the Sternberg paradigm (Zarahn et al., 2006).

Next, given the main effect of Instruction, we investigated the direct contrast of lying–truthful reporting. This demonstrated significant activation in middle and superior frontal gyri, bilateral precuneus, and middle temporal gyrus (**Table 2**). The middle frontal gyrus and precuneus were activated in the lying–truthful reporting contrast in Vartanian et al. (2012) and elsewhere (e.g., Ganis et al., 2003).

We had hypothesized that successful lying would place greater demands on inhibition under high WM load than under low WM load, but that truthful reporting would not place similar demands on inhibition under the same conditions. To test this hypothesis we selected those trials for which an accurate response was collected and compared responses under high and low WM load conditions. We did a direct contrast of the high vs. low WM load condition, but akin to what Vartanian et al. (2012) did, we selected for our Small Volume Correction in SPM8 a spherical region of interest (ROI) in the right inferior lateral prefrontal cortex (PFC) (coordinates of the center of mass $x = 51$, $y = 21$, $z = 12$) with a radius of 10 mm. This specific ROI was selected from Goel and Dolan (2003) in which it was associated with inhibition in reasoning. The same ROI was used by De Neys et al. (2006) as the ROI for inhibition in decision making. As shown in **Figure 2**, the high–low WM load contrast revealed significant activation in the right inferior frontal gyrus under instructions to lie (BA 45) (54, 30, 8, $z = 2.49$, $p = 0.006$). This activation was not present under instructions to tell the truth. Critically, an interaction analysis revealed a significantly greater difference between high and low WM load under instructions to lie than to tell the truth in two areas also located in the inferior frontal gyrus (BA 45) (62, 20, 16,

Table 2 | Activated regions corresponding to reported contrasts.

	BA	L	Z	x	y	z
HI WM LOAD–LOW WM LOAD						
Middle frontal gyrus	6	l	5.28	−50	8	38
	6	r	4.08	34	0	26
Precuneus	7	l	4.95	−4	−74	60
Intraparietal sulcus	40	l	4.39	−30	−58	44
Cerebellum	−	l	4.31	−8	−90	−24
Supplementary motor area	6	l	3.92	−4	−4	60
Caudate	−	r	3.92	12	−8	20
Dorsolateral PFC	46	r	3.74	48	22	30
LYING–TRUTHFUL REPORTING						
Middle frontal gyrus	6	l	5.18	−38	16	46
Superior frontal gyrus	6	l	4.67	−4	32	50
Precuneus	7	l	4.19	−44	−50	46
	7	r	3.84	42	−60	42
Middle temporal gyrus	21	r	3.65	60	−32	−10

Regions are designated using MNI coordinates; BA indicates Brodmann area; L indicates laterality; l and r indicate left and right hemispheres, respectively; Z indicates z -score.

Table 1 | Response time (in ms) and accuracy (expressed as a percentage) with standard errors (SE) in brackets for each condition.

Instructions				
Lie			Truth	
WM load	RT (SE)	Accuracy (SE)	RT (SE)	Accuracy (SE)
High	1719 (99)	91 (0.08)	1332 (84)	94 (0.06)
Low	1555 (91)	94 (0.01)	1203 (71)	98 (0.01)

$z = 3.43$, $p < 0.001$; 54, 26, 10, $z = 3.13$, $p < 0.001$) (**Figure 3**). In other words, high WM load activated the right inferior frontal gyrus more when lying successfully than when telling the truth successfully.

DISCUSSION

Our results are consistent with previous work reporting increased RT in response to increased WM load (e.g., Sternberg, 1966) and the requirement to lie (e.g., Holden and Hibbs, 1995). Our findings and interpretation are also consistent with Vrij et al.'s (2010) hypothesis that one strategy to detect deception is the placement of greater cognitive load on a suspect. More critically, however, our neurological data showed that when WM capacity is depleted, inhibitory workload (as measured by the BOLD signal) is increased specifically for those trials on which participants were required to suppress the truth and respond with a lie.

Several papers in the neuroscience literature have demonstrated that deception activates neural systems underlying WM and executive functions (for reviews see Spence et al., 2004; Sip et al., 2008; Abe, 2009, 2011). The involvement of the PFC has been a recurrent theme, particularly because of its known role in inhibiting behavior, which in the case of lying involves the suppression of truthful responses. Our results contribute to

this literature by demonstrating the role of the right inferior frontal gyrus in successful lying in the high vs. low WM load condition¹. Specifically, we have shown that WM load differentially impacts brain function in right inferior frontal gyrus under instructions to lie, but not under instructions to tell the truth. We postulate that the particular role of the right inferior frontal gyrus during a lie is to suppress the truthful response, and that suppression requires more effort when WM is taxed. This interpretation is also consistent with the RT difference observed between the high and low load conditions when participants were instructed to lie.

One could argue that our participants might have adopted a task-switching strategy under instructions to lie. In other words, the instruction to lie could invoke a switch in the mapping between stimulus and response, and as such reveal little about the participants' intention to lie. Our survey of the recent task-switching literature suggests that such an interpretation is unlikely. Specifically, in several of the papers we reviewed (e.g., Meiran, 2000; Vu and Proctor, 2004; Crone et al., 2006), when participants were cued to indicate the mapping to be applied to the stimulus, the cost in terms of response time for switching tended to be around 100 ms. If our lying manipulation triggered a strategy in which participants simply reassigned the stimulus-response mapping depending on the cue, we should have witnessed a similar cost in response time. By contrast, our comparatively large RT cost (over 300 ms) suggests more effortful processing of the stimulus, and that the activation exhibited in the right inferior frontal gyrus is more likely associated with lying rather than simple task switching.

It has been noted before that an important limitation of studies of lie detection involves the use of experimental designs in which participants were instructed to lie on demand (see Sip et al., 2008). This criticism raises important concerns about the ecological validity of the employed methodologies and by extension, empirical findings. Two recent studies have challenged this criticism by enabling participants to engage in spontaneous lying. Greene and Paxton (2009) instructed their participants to predict the outcomes of computerized coin flips while they were being scanned with fMRI. Correct predictions were rewarded by monetary gain. Importantly, in some trials participants were rewarded based on self-reported accuracy. This allowed them to gain money dishonestly by lying about the accuracy of their predictions. Indeed, given this opportunity many participants behaved dishonestly by lying about their predictions, assessed by improbably high levels of deviation from chance (i.e., 50% for coin flips). Their fMRI results revealed that lying was associated with neural activity in anterior cingulate cortex (ACC), dorsolateral PFC and the inferior frontal gyrus. In addition, activation in these regions was also associated with individual differences in the frequency of lying. This individual-differences result is particularly interesting because it links a tendency to engage in lying

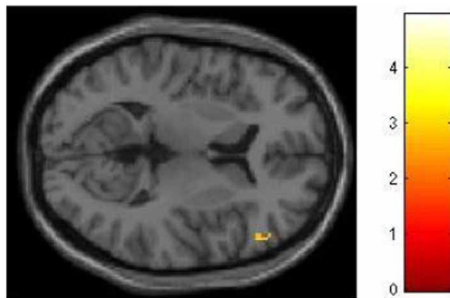


FIGURE 2 | The right inferior frontal gyrus is activated more for successful lying under high WM load than low WM load. SPM rendered into standard stereotactic space and superimposed on to transverse MRI in standard space. The bar graph represents the strength of the activation (T -score).

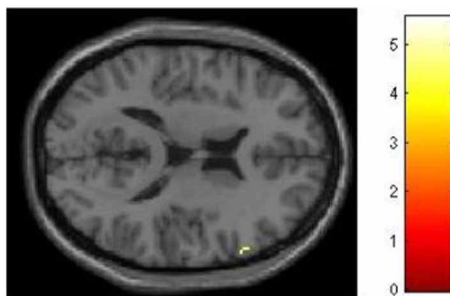


FIGURE 3 | High working memory load activates the right inferior frontal gyrus more when lying successfully than when telling the truth successfully. SPM rendered into standard stereotactic space and superimposed on to transverse MRI in standard space. The bar graph represents the strength of the activation (T -score).

¹Despite the low z -score in relation to the simple main effect of WM load under instructions to lie, the results were consistent with the interaction analysis that also revealed significantly greater difference in activation in right inferior frontal gyrus between high and low WM load under instructions to successfully lie than to tell the truth.

to the same region that Vartanian et al. (2012) found to predict lying skill—the right inferior frontal gyrus.

In a more recent study, Ding et al. (2013) used near-infrared spectroscopy (fNIRS) to study spontaneous deception. NIRS is a non-invasive imaging method that allows *in-vivo* photometric measurement of changes in the concentrations of oxygenated and deoxygenated hemoglobin in the cortex, and can thus be used to characterize physiological blood oxygenation changes in relation to cognitive tasks. The participants' task was to predict, on each trial, the side of the screen in which a coin would appear. Participants put each of their hands in one of two drawers of a desk (so their hand movements would not be directly visible to the experimenter). Participants made their predictions by moving their hand corresponding to the predicted side. Following the presentation of the coin on the screen, a message on the screen asked them whether they had guessed the location of the coin correctly. However, unbeknownst to the participants, the experimenters had installed hidden cameras inside the drawers to record the movement of each participant's hands. This enabled the experimenters to determine, on a trial-by-trial basis, whether the participants had engaged in spontaneous deception. The results demonstrated that lying was correlated with increased activity in left superior frontal gyrus (BA 6)—the area also activated in Vartanian et al. (2012) and the present study in the lying–truthful contrast. Thus, the studies by Greene and Paxton (2009) and Ding et al. (2013) demonstrate that the PFC plays a role in deception—regardless of whether it occurs spontaneously or is triggered on demand. Nevertheless, given that standard fMRI activation patterns are expressed as subtractions, not only is the choice of an appropriate control condition vis-à-vis lying critical for meaningful interpretation of results (Friston et al., 1996), but also vital for a meaningful comparison of the findings reported across laboratories.

Based on recent theoretical and methodological advances in the neuroscience of deception, it would appear that neuroimaging has the potential to eventually develop into a useful part of the forensic toolkit for lie detection (for reviews see Abe, 2009, 2011). However, important questions remain unanswered. For example, because neuroimaging studies are correlational, they cannot definitively determine the necessity of any brain region for deception. Evidence that can determine necessity is provided by loss-of-function studies that investigate permanent inability to lie as a function of neuropsychological impairment, or a transient inability to lie due to “temporary lesions” instantiated using transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS) (for a review see Luber et al., 2009). Unfortunately, evidence from loss-of-function studies regarding the role of PFC in deception has been inconsistent. For example, Luber and colleagues, using a variant of the Guilty

Knowledge Test [adapted from Langleben et al. (2002, 2005)], applied TMS pulses to left DLPFC and parietal cortex to disrupt the neural circuitry shown to be correlated with the formation of deceptive responses. The results demonstrated that TMS pulses applied exclusively to the parietal cortex increased RT by 20%, whereas the same stimulation applied to left DLPFC alone had no effect on RT. These results cast doubt on the necessity of PFC for the formation of lies (see also Verschuere et al., 2012). On the other hand, Priori et al. (2008) found that applying anodal tDCS to bilateral DLPFC did increase RT for denial lies. The inconsistency suggests that continued study is needed to determine precisely the conditions under which PFC and its subregions necessarily contribute to specific aspects of deception.

In addition, there does not appear to be an activation pattern that is unique to lying or deception (Wolpe et al., 2005; Appelbaum, 2007; Sip et al., 2008). Rather, as is the case with other higher-order mental processes such as reasoning and decision making (Goel, 2007; Frank et al., 2009), lying and deception appear to be built on multiple neural systems that are differentially activated as a function of task and contextual demands. In the case of lying and deception those processes include, among others, WM, error monitoring, response selection, and target detection (Hester et al., 2004; Huettel and McCarthy, 2004; Zarah et al., 2006). This makes the use of fMRI for lie detection in forensic and legal settings challenging, given that practitioners in applied settings will be unable to make clear-cut judgments of guilt based on fMRI data alone. However, neuroimaging data could comprise one of many components of a broader arsenal for detecting deception. For example, according to the “information gathering” approach to lie detection, interviewers are instructed to focus on gathering verbal information from suspects that can be subsequently checked for inconsistencies against available evidence (Vrij et al., 2010). The approach is predicated on not focusing on a single cue, but rather collecting and cross-referencing their consistency. By providing neural information, neuroimaging evidence can contribute to the forensic decision-making apparatus in this context. This componential approach was reinforced in a recent report by the National Academy of Sciences (2008). We suggest that a broad set of metrics that combines verbal, non-verbal, and neural data provides the most promising framework for lie detection in the lab and elsewhere.

ACKNOWLEDGMENTS

We thank Fred Tam, Caron Murray, Ruby Endre, Rafal Janik, Sofia Chavez, Sheila Petrongolo, and Lisa Carnduff for their research assistance. This research was supported by Technology Investment Fund project 15dz01 under the direction of the third author.

REFERENCES

- Abe, N. (2009). The neurobiology of deception: evidence from neuroimaging and loss-of-function studies. *Curr. Opin. Neurol.* 22, 594–600. doi: 10.1097/WCO.0b013e328332c3cf
- Abe, N. (2011). How the brain shapes deception: an integrated review of the literature. *Neuroscientist* 17, 560–574. doi: 10.1177/1073858410393359
- Appelbaum, P. S. (2007). Law and psychiatry: the new lie detectors: neuroscience, deception, and the courts. *Psychiatr. Serv.* 58, 460–462. doi: 10.1176/appi.ps.58.4.460
- Aron, A. R., Fletcher, P. C., Bullmore, E. T., Sahakian, B. J., and Robbins, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nat. Neurosci.* 6, 115–116. doi: 10.1038/nn1003
- Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8, 170–177. doi: 10.1016/j.tics.2004.02.010
- Beattie, G. W. (1981). A further investigation of the cognitive interference hypothesis of gaze patterns during conversation. *Br. J. Soc. Psychol.* 20, 243–248. doi: 10.1111/j.2044-8309.1981.tb00493.x

- Crone, E. A., Wendelken, C., Donohue, S. E., and Bunge, S. A. (2006). Neural evidence for dissociable components of task-switching. *Cereb. Cortex* 13, 475–486. doi: 10.1093/cercor/bhi127
- De Neys, W., Vartanian, O., and Goel, V. (2006). Smarter than we think: when our brains detect that we are biased. *Psychol. Sci.* 19, 483–489. doi: 10.1111/j.1467-9280.2008.02113.x
- Ding, X. P., Gao, X., Fu, G., and Lee, K. (2013). Neural correlates of spontaneous deception: a functional near-infrared spectroscopy (fNIRS) study. *Neuropsychologia* 51, 704–712. doi: 10.1016/j.neuropsychologia.2012.12.018
- Frank, M. J., Cohen, M. X., and Sanfey, A. G. (2009). Multiple systems in decision-making: a neurocomputational perspective. *Curr. Dir. Psychol. Sci.* 18, 73–77. doi: 10.1111/j.1467-8721.2009.01612.x
- Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., and Dolan, R. J. (1996). The trouble with cognitive subtraction. *Neuroimage* 4, 97–104. doi: 10.1006/nimg.1996.0033
- Ganis, G., Kosslyn, S. M., Stose, S., Thompson, W. L., and Yurgelun-Todd, D. A. (2003). Neural correlates of different types of deception: an fMRI investigation. *Cereb. Cortex* 13, 830–836. doi: 10.1093/cercor/13.8.830
- Goel, V. (2007). The anatomy of deductive reasoning. *Trends Cogn. Sci.* 11, 435–441. doi: 10.1016/j.tics.2007.09.003
- Goel, V., and Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition* 87, B11–B22. doi: 10.1016/S0010-0277(02)00185-3
- Greene, J. D., and Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12506–12511. doi: 10.1073/pnas.0900152106
- Hester, R., Fassbender, C., and Garavan, H. (2004). Individual differences in error processing: a review and reanalysis of three event-related fMRI studies using the GO/NOGO task. *Cereb. Cortex* 14, 986–994. doi: 10.1093/cercor/bhh059
- Holden, R. R., and Hibbs, N. (1995). Incremental validity of response latencies for detecting fakers on a personality test. *J. Res. Pers.* 29, 362–372. doi: 10.1006/jrpe.1995.1021
- Huetzel, S. A., and McCarthy, G. (2004). What is odd in the oddball task? Prefrontal cortex is activated by dynamic changes in response strategy. *Neuropsychologia* 42, 379–386.
- Langleben, D. D., Loughhead, J. W., Bilker, W. B., Ruparel, K., Childress, A. R., Busch, S. I., et al. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Hum. Brain Mapp.* 26, 262–272. doi: 10.1002/hbm.20191
- Langleben, D. D., Schroeder, L., Maldjian, J. A., Gur, R. C., McDonald, S., Ragland, J. D., et al. (2002). Brain Activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage* 15, 727–732. doi: 10.1006/nimg.2001.1003
- Luber, B., Fisher, C., Appelbaum, P. S., Ploesser, M., and Lisanby, S. H. (2009). Non-invasive brain stimulation in the detection of deception: scientific challenges and ethical consequences. *Behav. Sci. Law* 27, 191–208. doi: 10.1002/bsl.860
- Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychol. Res.* 63, 234–249. doi: 10.1007/s004269900004
- National Academy of Sciences. (2008). *Emerging Cognitive Neuroscience and Related Technologies*. Washington, DC: The National Academies Press.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63. doi: 10.1016/j.tics.2005.12.004
- Priori, A., Mameli, F., Cogiamanian, F., Marceglia, S., Tiriticco, M., Mrakic-Sposta, S., et al. (2008). Lie-specific involvement of dorsolateral prefrontal cortex in deception. *Cereb. Cortex* 18, 451–455. doi: 10.1093/cercor/bhm088
- Quas, J. A., Davis, E. L., Goodman, G. S., and Myers, J. E. B. (2007). Repeated questions, deception, and children's true and false reports of body touch. *Child Maltreat.* 12, 60–67. doi: 10.1177/1077559506296141
- Sip, K. E., Roepstorff, A., McGregor, W., and Frith, C. D. (2008). Detecting deception: the scope and limits. *Trends Cogn. Sci.* 12, 48–53. doi: 10.1016/j.tics.2007.11.008
- Spence, S. A., Hunter, M. D., Farrow, T. F. D., Green, R. D., Leung, D. H., Hughes, C. J., et al. (2004). A cognitive neurobiological account of deception: evidence from functional neuroimaging. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 359, 1755–1762. doi: 10.1098/rstb.2004.1555
- Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652
- Vartanian, O., Kwantes, P., and Mandel, D. R. (2012). Lying in the scanner: localized inhibition predicts lying skill. *Neurosci. Lett.* 529, 18–22. doi: 10.1016/j.neulet.2012.09.019
- Verschuere, B., Schuhmann, T., and Sack, A. T. (2012). Does the inferior frontal sulcus play a functional role in deception? A neuronavigated theta-burst transcranial magnetic stimulation study. *Front. Hum. Neurosci.* 6:284. doi: 10.3389/fnhum.2012.00284
- Vrij, A., Fisher, R., Mann, S., and Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends Cogn. Sci.* 10, 141–142. doi: 10.1016/j.tics.2006.02.003
- Vrij, A., Granhag, P. A., and Porter, S. B. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychol. Sci. Public Interest* 11, 89–121. doi: 10.1177/1529100610390861
- Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., and Bull, R. (2008). Increasing cognitive load to facilitate lie detection: the benefit of recalling an event in reverse order. *Law Hum. Behav.* 32, 253–265. doi: 10.1007/s10979-007-9103-y
- Vu, K.-P. L., and Proctor, R. W. (2004). Mixing compatible and incompatible mappings: elimination, reduction, and enhancement of spatial compatibility effects. *Q. J. Exp. Psychol. A* 57, 539–556. doi: 10.1080/02724980343000387
- Wolpe, P. R., Foster, K. R., and Langleben, D. D. (2005). Emerging neurotechnologies for lie-detection: promises and perils. *Am. J. Bioeth.* 5, 39–49. doi: 10.1080/15265160590923367
- Zarahn, E., Rakitin, B. C., Abela, D., Flynn, J., and Stern, Y. (2006). Distinct spatial patterns of brain activity associated with memory storage and search. *Neuroimage* 33, 794–804. doi: 10.1016/j.neuroimage.2006.07.020

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 April 2013; accepted: 07 September 2013; published online: 03 October 2013.

Citation: Vartanian O, Kwantes PJ, Mandel DR, Bouak F, Nakashima A, Smith I and Lam Q (2013) Right inferior frontal gyrus activation as a neural marker of successful lying. *Front. Hum. Neurosci.* 7:616. doi: 10.3389/fnhum.2013.00616

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Her Majesty the Queen in Right of Canada. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Prospects of functional magnetic resonance imaging as lie detector

Elena Rusconi^{1,2*} and Timothy Mitchener-Nissen¹

¹ Department of Security and Crime Science, University College London, London, UK

² Department of Neurosciences, University of Parma, Parma, Italy

Edited by:

Andrea Szymkowiak, University of
Abertay Dundee, UK

Reviewed by:

Elliot Berkman, University of
Oregon, USA
Luca Sammiceli, University of
Bologna, Italy
Federico G. Pizzetti, Università degli
Studi di Milano - Dipartimento di
Studi Internazionali, Giuridici e
Storico-Politici, Italy

*Correspondence:

Elena Rusconi, Department of
Security and Crime Science,
University College London,
35 Tavistock Square, London,
WC1H 9EZ, UK
e-mail: elena.rusconi@gmail.com

Following the demise of the polygraph, supporters of assisted scientific lie detection tools have enthusiastically appropriated neuroimaging technologies “as the savior of scientifically verifiable lie detection in the courtroom” (Gerard, 2008: 5). These proponents believe the future impact of neuroscience “will be inevitable, dramatic, and will fundamentally alter the way the law does business” (Erickson, 2010: 29); however, such enthusiasm may prove premature. For in nearly every article published by independent researchers in peer reviewed journals, the respective authors acknowledge that fMRI research, processes, and technology are insufficiently developed and understood for gatekeepers to even consider introducing these neuroimaging measures into criminal courts as they stand today for the purpose of determining the veracity of statements made. Regardless of how favorable their analyses of fMRI or its future potential, they all acknowledge the presence of *issues yet to be resolved*. Even assuming a future where these issues are resolved and an appropriate fMRI lie-detection process is developed, its integration into criminal trials is not assured for the very success of such a future system may necessitate its exclusion from courtrooms on the basis of existing legal and ethical prohibitions. In this piece, aimed for a multidisciplinary readership, we seek to highlight and bring together the multitude of hurdles which would need to be successfully overcome before fMRI can (if ever) be a viable applied lie detection system. We argue that the current status of fMRI studies on lie detection meets neither basic legal nor scientific standards. We identify four general classes of hurdles (scientific, legal and ethical, operational, and social) and provide an overview on the stages and operations involved in fMRI studies, as well as the difficulties of translating these laboratory protocols into a practical criminal justice environment. It is our overall conclusion that fMRI is unlikely to constitute a viable lie detector for criminal courts.

Keywords: fMRI, lie detection, evidence, scientific validity, human rights

INTRODUCTION

In recent years researchers in cognitive neuroscience have started to investigate the neural basis of complex mental processes including moral beliefs, intentions, preferences, self-knowledge, social interactions, and consciousness. Influential neuroscientists are introducing the idea that our traditional notions of crime and punishment (and the laws built upon them) should be challenged, and if necessary modified, to make them more *human-friendly*. Recent empirical findings with neuroimaging techniques challenge the central idea of free will around which much of the criminal law has been shaped (see e.g., Gazzaniga, 2008). Additionally, structural MRI evidence is making inroads in courts around the World [see e.g., *Commonwealth of Pennsylvania v. Pirela*, 2007; *Caso Bayout—Corte d'Assise d'Appello di Trieste (n.5/2009) del 18 settembre 2009*; *Tribunale di Como (n.536/2011) del 20 maggio 2011*] and it seems that not before long functional Magnetic Resonance Imaging (fMRI) scans will be routinely requested by the defense when searching for either mitigating factors, such as anatomo-functional abnormalities, and/or the presence of any crucial memories when self-reports can be doubted (e.g., Abbott,

2001, 2007; Hughes, 2010). On similar grounds, and concomitant with attempts to promote fMRI as a mind-reading tool (see Logothetis, 2008, for a specialist overview), fMRI has been proposed as a possible state-of-the-art tool for detecting both malignancy and deception in criminal courts even though it has not yet been considered admissible evidence (e.g., *US v. Semrau*, 2010; <http://blogs.law.stanford.edu/lawandbiosciences/2010/06/01/fmri-lie-detection-fails-its-first-hearing-on-reliability/>), also see Sip et al., 2007, 2008 and Haynes, 2008, for contrasting specialist views on applications in lie detection). In addition to raising questions regarding fMRI's reliability as a lie detecting tool according to scientific standards, such advocacy raises ethical and legal issues that are common to any putative lie detection technology thus engaging the attention of lawyers, ethicists, and philosophers.

Despite all these concerns fMRI is already being advertised as a scientifically proven lie detector by private companies having strong links with academia (see No Lie MRI—<http://noliemri.com/> and CEPHOS—<http://www.cephoscorp.com/>), one that has not (yet) been subjected to the same regulation as the

polygraph and thus is not considered an illegal means of assessment in pre-employment settings. As “trust” is increasingly prioritized in certain business sectors, top-tier corporations may be tempted to assess the trustworthiness of their key current and future employees by requesting they undergo a lie detection test via fMRI. However, in the more conservative criminal justice sector, several hurdles confront any use of fMRI as a viable lie detector. We will attempt herein to provide a realistic and accessible evaluation of such hurdles by discussing those questions raised by fMRI use for lie-detection purposes in criminal courts.

But firstly in the following two sections we take a brief look at the basics of this technique in order to gauge an impression of what types of evidence fMRI currently may and may not be able to provide. For although most neuroscientists would agree that fMRI should not be used as a lie detector, especially within its current form (e.g., Grafton et al., 2006; Tovino, 2007), the debate has recently seen the identification of a possible route toward the use of fMRI for lie detection by separating scientific from legal standards (Schauer, 2010) or basic from translational research (Langleben and Moriarty, 2013).

FUNCTIONAL MAGNETIC RESONANCE IMAGING: BASICS

fMRI is one of the most popular measurement techniques in cognitive neuroscience. It has been in use for about 20 years and is qualified as *correlational* because it records brain states in parallel with ongoing mental activity and/or behavior, thus permitting the establishment of correlational links between them. However, it does not allow researchers to establish a causal connection between brain states and behaviors or supposed mental processes. In most fMRI studies, brain states are the dependent variable measured during manipulation of the stimulus/task condition. Whether any specific local or systemic pattern of brain states is a necessary determinant of its associated behavior it cannot be determined with fMRI only. For this reason, fMRI is routinely used in basic research as a mainstay method to measure brain function and its data are often triangulated with data from complementary techniques (e.g., event-related potentials, transcranial magnetic stimulation), in a quest for converging evidence about mental processes and brain substrates.

As implied by its name, fMRI makes use of strong magnetic fields to create images of biological tissue. Depending on the *pulse sequence*¹ of the electromagnetic fields it generates, an MRI scanner can detect different tissue properties and distinguish between tissue types. Scanners are used to acquire both *brain structural* information (e.g., allowing a fine distinction between white and gray matter, producing images of the static anatomy of the brain), and *functional*² information such as measurements of local changes in blood oxygenation within the brain over time;

the most common form of fMRI study. Because blood oxygenation levels change rapidly (i.e., after 1–2 s) following the activity of neurons in a brain region, fMRI allows researchers to localize brain activity on a second-by-second basis and within millimeters of its origin (Logothetis and Pfeuffer, 2004). These changes in blood oxygenation occur naturally and internally as part of normal brain physiology and, because the pulse sequence does not alter neuronal firing or blood flow, fMRI is considered a non-invasive technique (Huettel et al., 2009).

Central to cognitive fMRI studies are the concepts of differences and similarities between maps of *blood oxygenation level-dependent (BOLD) signal*³ that are recorded in concomitance with different experimental conditions. In classical fMRI designs and in most of the available lie detection studies BOLD responses are evaluated in relative terms as the result of a contrast between two or more conditions. For example, maps of the BOLD signal that are recorded while a participant is lying can be contrasted with either maps recorded when the participant is at rest or when is telling the truth. Inferences about the neural correlates of lying are typically drawn from an analysis of the pattern of differences and/or similarities between BOLD signal maps across *lying* and *not-lying* conditions⁴. In principle, any design difference (e.g., the use of a different stimulus or the requirement of additional mental operations given the same stimulus) between the lying condition and any other condition with which it is compared might lead to the recruitment of different brain regions to perform the task. Therefore, the more accurately the *not-lying* and *lying* conditions are matched, the more precise the conclusions that can be drawn about the neural correlates uniquely associated with lying. While this type of analysis is not the only possible or optimal way to draw informative inferences from fMRI data (e.g., Sartori and Umiltà, 2002) such contrast between conditions is a basic standard in fMRI research. We would like to draw the attention here on the fact that the possibility to interpret as specific correlates of lying any findings ultimately resides in the original choice and design of experimental and control conditions. More recent approaches to the discrimination between *lying* vs. *not-lying* correlates include data-driven pattern classification algorithms (e.g., Davatzikos et al., 2005; Kozel et al., 2005),

¹ A pulse sequence is the series of changing magnetic field gradients and oscillating electromagnetic fields defined by the user that allows the MRI scanner to tune on and create images sensitive to a target physical property. Different pulse sequences, for example, are used when collecting structural data and functional brain activations.

² The term “functional” refers to changes in brain function and regional levels of activation over time.

³ fMRI is based on the difference in magnetic resonance signals from oxyhemoglobin and deoxyhemoglobin and builds on the fact that active brain regions tend to use more oxygen than relatively inactive regions. Soon after a brain region has been activated by a cognitive event or task, the local microvasculature responds to increased oxygen consumption by increasing the flow of new arterial blood (i.e., blood rich in oxyhemoglobin) to the region. As a consequence, the relative concentration of deoxyhemoglobin decreases, thus causing localized changes in the magnetic resonance signal. These changes are known as blood oxygenation level-dependent (BOLD) signal (Purves et al., 2013).

⁴ For a broad overview on other relevant testing paradigms—that can also be used in fMRI studies—we direct the reader toward Gamer (2011). For an insightful discussion on the complexity of deception and the unlikelihood of encompassing it by using simple tasks that require participants to lie in response to certain stimuli and tell the truth in response to others, we direct the reader toward Sip et al. (2007). Our discussion will provide prototypical examples in order to highlight general principles; it is by no means aimed to provide a comprehensive and systematic description of the vast literature on the topic.

which are bound to the possibility of an independent and objective classification of lie vs. truth (Sip et al., 2008). Common to all approaches within brain data analyses however is that that which is identified as the “correlates of lying” even at the individual level would be expected to emerge *across several lying trials* thereby capturing similarities across different instances of “lying,” rather than simply representing singularities associated with an *individual instance* of lying.

When evaluating fMRI evidence, with an eye on applying it to a real-world problem, it is important not only to be aware of basic experimental design principles but also of the peculiar requirements of the technique and its limitations (Spence, 2008). In this regard both scanner reliability and staff technical skills are fundamental to the internal validity⁵ of fMRI-testing protocols. This is often an issue since control conditions need to be carefully matched to experimental conditions in order to unequivocally isolate the *construct of interest*⁶. However, even with an elegant design the reliability and localization of BOLD signals depends on the extent to which participants perform their tasks accurately, consistently, and in compliance with all instructions (for example, not moving their head as movement will degrade the image). To double-check participant's compliance, behavior should be monitored and recorded during the scanning session whenever possible (e.g., by recording reaction times and accuracy in a task), and in experiments involving arousal or emotional stimuli, skin conductance, heart-rate or salivary hormones could be also monitored to provide converging information.

The outcome of data analysis is a function of a series of consensus-based decisions, including options and parameters chosen for realignment, normalization, and smoothing, statistical models for analyses, and the associated correction criteria that can be more or less conservative. Finally, strategic decisions may guide choice of the evidence that will find its way in the final report on a peer-reviewed journal. Although raw data may be requested by anyone for further analyses, most readers will exclusively rely on the information provided in a polished report. Furthermore, in very competitive scientific environments there is no incentive for investigators to try and replicate their own findings as journals typically promote the publication of novel designs rather than replications (e.g., Giner-Sorolla, 2012), and in real practice it is very unusual to see a brain imaging experiment precisely repeated within and between laboratories. This may prove especially problematic when trying to identify a well-known and reliable protocol for potential applications in the real world. Finally, numerous safety exclusion criteria apply which limits the generalizability of fMRI results and prevents its universal use (see a typical participant screening checklist at <http://airto.hosted.ats.ucla.edu/BMCweb/Consent/SafetyScreen.html>).

⁵Internal Validity refers to the appropriateness of construct operationalization and of experimental design in order to test the hypothesis of interest. It guarantees that any obtained effects may be univocally attributed to the experimental manipulation. Clearly, the use of expensive and fancy techniques does not guarantee by itself that experimental results are meaningful and interpretable.

⁶In this context, the “construct of interest” is the brain fingerprint of lying.

In summary, protocol design determines how fMRI evidence can be interpreted, full compliance is required from participants, and the final evidence reflects choices, assumptions and data transformations based on current scientific standards and consensus criteria but also on publication strategies. Finally, not everybody can undergo fMRI. So the question remains, can its potential contributions as a lie detector outweigh its intrinsic limitations?

THE LYING BRAIN

Many people believe they are very good at detecting deceit and that certain signs give away when somebody is lying: liars would talk too much and tell stories far more elaborate and detailed than required by the context; they would never gaze interlocutors straight in the eyes or would stare at them too intensely; or they would cross their arms or their legs; or a combination of behaviors (e.g., Houston et al., 2012). Yet studies show that the vast majority of onlookers correctly distinguish *truth* from *lies* when told by a stranger only about 54% of the time (i.e., they are only slightly better than chance). Notably, this same level of (in)accuracy holds true even for professional categories such as lawyers, policemen, magistrates, and psychiatrists (Bond and DePaulo, 2006).

Conversely, the ability to lie develops spontaneously (it is typically absent in children with neuro-developmental impairments, like autism). Lying is fundamental to healthy behavior, as shown by the disastrous social interactions of patients with orbito-frontal lesions. Indeed some of these patients become notoriously tactless—which in a final analysis can be achieved by always being completely frank and honest. The literature on orbito-frontal patients suggests in turn that the ability to lie depends on the integrity of localized neural circuits (e.g., Damasio, 1994).

Recent attempts have been made with fMRI to specify the neural correlates of lying or deception (see Christ et al., 2009; Abe, 2011 and Gamer, 2011 for recent overviews and meta-analyses; see Sip et al., 2007 for a discussion on deception and lying from a cognitive neuroscience perspective). In one of the typical experiments, researchers ask participants to answer truthfully to some questions/stimuli and lie in response to others. The BOLD contrast⁷ between the two conditions (i.e., the pattern of BOLD signals detected when the participant is lying minus the pattern of BOLD signals detected when the participant is being truthful; also indicated in the specialist literature as *Lie > Truth*) is expected to enable the identification of brain regions whose activation is significantly correlated with lying. Accordingly, several studies identified a network of parieto-frontal⁸ areas that are significantly more engaged when the individual is lying. As the opposite contrast (i.e., *Truth > Lie*) does not usually detect any regions that are significantly more engaged, most neuroscientists infer that lying requires extra-effort compared to responding truthfully. Such extra-effort is possibly aimed at inhibiting the

⁷The difference in signal on fMRI images from different experimental conditions as a function of the amount of deoxygenated hemoglobin.

⁸The term “parieto-frontal” areas denotes brain regions in the parietal and frontal lobes. The parietal lobe is located on the posterior and dorsal surfaces of the cerebrum. The frontal lobe is the most anterior lobe of the cerebrum.

truth and/or producing an alternative response that sounds realistic enough. In studies employing ecologically plausible stimuli, activation of regions in the limbic system (a deep brain structure traditionally associated with emotional responses) has also been associated with lying (e.g., Hakun et al., 2009). Note however that this does not imply in any mechanistic way that a person is lying when the same region of the limbic system or network of parietofrontal areas activates during a task (e.g., Poldrack, 2006, 2010; see also following *Scientific Hurdles* section, point 1). Finally, a great part of research on the neural correlates of deception has been focused on group-level results (i.e., results that are obtained by averaging data from several participants), whereas any real-world application would require a differential approach (i.e., it should provide evidence that is informative and predictive at the individual level).

Within a basic cognitive neuroscience perspective, fMRI research on deception can indeed aspire to provide correlation maps that possibly reflect the difference between deceitful and truthful responses. In order to obtain knowledge about the anatomo-functional substrates that are causally related to lying, and disambiguate potentially spurious activations, evidence would need to be collected with complementary techniques (e.g., with neurological lesion or non-invasive brain stimulation studies). fMRI is thus useful inasmuch as it hands over to techniques with complementary inferential power a map for (1) identifying cortical networks that play a necessary role in deception, and (2) testing their role by directly manipulating an individual's ability to deceive. This information could then feed back into fMRI maps and enable the identification of the most relevant correlates of lying for applicative purposes. The ability to establish causal links between brain substrates and behavior resides in the fact that the functionality of the brain tissue underlying stimulators can be temporarily modulated (e.g., see ; Nitsche et al., 2008; Sandrini et al., 2011). For example, by modulating the activity of frontal lobe areas with non-invasive brain stimulation, Priori et al. (2008) were able to interfere with intentional deception by slowing down the production of untruthful responses (see also Mameli et al., 2010). Karim et al. (2010) could enhance the ability to lie by modulating activity in a contiguous part of the frontal lobe, the anterior prefrontal cortex. It thus seems possible to manipulate efficiency in lie production by targeting specific brain regions (see Luber et al., 2009, for a discussion of related ethical implications), although careful task analysis, replication and clarification of the underlying mechanisms of action of non-invasive brain stimulation techniques need to be carried out before endorsing any mass applications. This should suggest how in basic neuroscience (1) fMRI can contribute to our models of the brain substrates of lying, however for completeness its evidence is best integrated with evidence from complementary techniques, (2) fMRI evidence alone does not provide compelling evidence as to whether certain neural substrates are strictly necessary to the process of lying. Other techniques may help restrict the focus to a subset of potential substrates.

As a final point it is worth remembering that in basic research a participant's compliance with instructions is almost taken for granted as there is no rational reason why a participant might benefit from not following them. Quite the opposite situation

may arise in a criminal forensic setting however, whereby it is not difficult to imagine that either intentional (e.g., adopting countermeasures) or non-intentional (e.g., due to alterations in one's emotional state) factors may lead to inconclusive results. In this respect, a recent study by Ganis et al. (2011) has eloquently shown how easy it is to "fool" an fMRI test for participants who have been trained in the use of task-tailored countermeasures.

fMRI AS LIE DETECTOR IN CRIMINAL COURTS

THE SCIENTIFIC HURDLES

Legal systems are not new to influences from the cognitive neurosciences. For example, admissible MRI evidence showing the absence of frontal lobe maturation in the brains of teenagers contributed to the elimination of the death penalty for minors in some US states (frontal lobes are causally implicated in decision-making and the control of impulsive reactions; e.g., Damasio, 1994; Coricelli and Rusconi, 2011). Additionally structural brain scans are widely admissible at sentencing and are now almost invariably present in capital cases. However, when it comes to lie detection not all procedures have proven acceptable with polygraphs failing to attain general admissibility in criminal courts⁹ with the exception of New Mexico.

Despite this final fact, in 2006 two private bodies *No Lie MRI* and *Cephos Corporation* were launched with the goal of bringing fMRI lie detection to the public for use in legal proceedings, employment screening, and national security investigations. Detection accuracy was claimed to be as high as 90% (compared to a purported 70% for polygraphs). Attempts are being made to admit fMRI evidence in criminal courts; for example at the end of 2009 tests performed by No Lie fMRI were presented as evidence by the defense in a child protection hearing to prove innocence claims of a parent accused of committing sexual abuse. Had they been admitted that would have been the first time fMRI was used in an American court (Simpson, 2008). They were not but it might only be a matter of time before judges form the opinion that fMRI may provide relevant scientific evidence (Aharoni et al., 2012) opening the door to their wider admissibility.

Within this and the following sections we summarize and bring together the multitude of hurdles which need to be overcome before fMRI can ever be successfully integrated into criminal trials. Our discussions are primarily restricted to the English common law system of adversarial justice as applied throughout the United Kingdom, the United States, and Australia amongst others, as opposed to the continental European mixed adversarial-inquisitorial civil law systems. This decision is based on the particular nature of adversarial trials, with its competing prosecution and defense counsels who in turn can engage the services of competing expert neuroimaging witnesses, which may exacerbate some of the issues surrounding fMRI evidence discussed herein.

Legally, for scientific evidence to be admissible in criminal trials it must meet the legal standards as set down in the relevant jurisdiction, be these common law requirements such as

⁹In some states polygraph evidence is permitted when both the prosecution and defense agree to its admissibility, while in others such evidence cannot be admitted even when both parties would otherwise agree.

either the test under *Frye v United States* (1923: 293 F.1013) or the succeeding requirements under *Daubert v Merrell Dow Pharmaceuticals Inc.* (1993: 509, U.S.579) as applied in *Kumho Tire Co. v Carmichael* (1999: 526, U.S.137), the presence of statutory requirements, international conventions, Federal Rules of Evidence, or any permutation of these. Drawing from these various requirements there are general principles that scientific evidence must be both *relevant* and thereby possessing probative value, as well as being *reliable*. It is primarily this second concept of reliability that is our focus here.

Within the specific constraints of the criminal law we can comprehend scientific evidence as being reliable if, amongst other things: the methods and results are both consistent and consistently applied; the accuracy of results meets an acceptably high standard while both false positives and false negatives are minimized; what practitioners *believe* is being measured is *actually* being measured; the processes being measured are both understood by scientists and are agreed upon by scientists working in the field or who choose to examine the processes; and the scientific processes being relied upon apply equally to all individuals regardless of any internal or external traits or influences, or if there is variation this has been addressed in relation to the individual at hand. While these requirements may appear somewhat ill-defined to the objective scientist, they reflect the style of judge-made legalistic tests whereby relatively broad requirements may be set down. Within the field of law this flexibility is not seen as a vice as it both allows a future court to judge a case on its merits and does not undermine the role of the jury as the final arbiter of truth.

From the published fMRI literature it unavoidably emerges that fMRI technology has not reached this *reliability* threshold. Issues which require addressing by cognitive neuroscientists are set out below:

- (1) Assumptions and inferences underlying fMRI processes and technologies need to be confirmed (or dispelled) so as to give credence to the scientific claims being made. Cognitive neuroscience, for example, assumes that complex thoughts have a physical counterpart that is both accessible and interpretable with technologies such as fMRI (Erickson, 2010). Many fMRI researchers operate on the basic assumption that lying involves additional efforts than telling the truth, which in turn can be signaled by heightened blood flow in specific brain regions (Gerard, 2008). However, several fMRI studies have been employing “reverse inference” as a central feature, whereby the activation of certain brain regions (X) is taken as evidence of a particular cognitive function (Y). As thoroughly discussed by Poldrack (2006), such inferences are only deductively valid if brain state X only occurs when cognitive function Y is engaged (i.e., if a selective association between X and Y is established), yet this one-to-one matching is not the case. Rather many-to-many matching of brain states to mental states are observed, and thus valid reverse inferences cannot be made here. What is required first of all is the creation of a robust “cognitive ontology” specifying the component brain operations that comprise specific mental functions, even before trying
- to establish univocal associations with functional anatomy (Poldrack, 2010). Furthermore, data-driven pattern analyses approaches (e.g., Haynes, 2008), although more current in terms of the methodology and less constrained by theoretical assumptions, still rely on the objective identification of what a lie is (e.g., Sip et al., 2008). However, this is not always possible and is especially unlikely in forensic contexts where lie detectors would be employed when neither facts nor subjective intentions can be directly verified. The validity of underlying assumptions must be addressed and a wide consensus reached within the scientific community before possible applications of the technology can achieve broad credibility.
- (2) To achieve internal validity, it needs to be conclusively determined that what is being measured is actually evidence of deception and not unrelated cognitive processes, and this needs to be determinable for each and every response given by every future individual undergoing fMRI questioning when operational. Contrary to public expectations lie detectors like fMRI are not *mind readers*, do not actually detect deception, and will never provide details of what has actually happened in complicated cases. Rather they merely detect and measure manifestations of thoughts through changes in oxygenated blood which proponents consider denotes lying (Andrewartha, 2008; Holley, 2009). What fMRI lie detection actually depends upon is an ability to detect the suppression of competing responses, and yet remains hamstrung by the inability to determine what these competing responses are and what the suppression of these implies. As pointed out by Grafton and colleagues, “[m]any defendants [while testifying] do inhibit their natural tendency to blurt out everything they know. They are circumspect about what they say. Many of them also suppress expressions of anger and outrage at accusation. Suppressing natural tendencies is not a reliable indicator of lying, in the context of a trial” (Grafton et al., 2006, pp. 36–37). Critics have also gone on to argue that any exhibited increase in blood flow detected by fMRI may result from alternate neurological process such as anxiety, fear, or other heightened emotional states which are unrelated to the question of deception. In other words, just because the prefrontal cortex is activated during deception it does not follow that every time the prefrontal cortex activates the individual is lying (Fox, 2009; Moreno, 2009). Furthermore, assuming increased oxygenated blood flow in specific brain regions denotes deception, scientists have not agreed with a degree of precision as to what these specific regions are (Gerard, 2008). Should cognitive neuroscientists successfully address these questions, the machines and processes they develop will need to be constantly identifying and correcting for extraneous mental activity during the questioning process of criminal suspect and witnesses, otherwise any results will be an open target for legal challenge by opposing expert counsel.
- (3) The question of individual differences affecting fMRI results needs to be answered. The importance of this issue for a technology seeking both legitimacy and broad application cannot be underestimated. Neuroimaging devices need to be able to identify and correct for variances in individuals’ brains and

not operate on shared but unproven assumptions that all or most brain's process lies similarly (Ellenberg, 2009; Holley, 2009). This includes correcting for variations in brain processes based on age; particularly juveniles. They need to be able to cope with the types of individuals usually encountered by law enforcement officers, including substance addicts, those with high incentives to lie, and those with mental disorders. Doubts already exist as to whether fMRI would be usable for those presenting with conditions such as delusions and amnesic disorders with confabulation (Langleben et al., 2006). Finally there is the issue of possible differences in outputs resulting from the social diversity of those tested, given that what is considered a lie is a matter of social convention which may vary on a cultural basis (Holley, 2009).

- (4) The question of whether subjects or questioners can manipulate the fMRI baseline or response data needs to be addressed. To measure neurophysiological changes in the brain, neuroimaging devices must be able to create a reliable baseline against which comparisons can be formed. In this regard fMRI depends upon the cooperation of the subject and is highly vulnerable to countermeasures (e.g., Ganis et al., 2011). Trained participants can alter test results by engaging in some taxing activity like mental calculations during control sequences which will enormously reduce the power of the contrast between truthful statements and lies. Furthermore, it is not yet clear whether extensive rehearsing of a story which subsequently requires virtually no mental effort to retell will diminish detection rates (see Ganis et al., 2003 for a first attempt to distinguish spontaneous from rehearsed lies; Gerard, 2008). A separate issue of manipulation relates to both experimenter and questioner expectancy effects. Consider a scenario whereby fMRI evidence is considered admissible and a criminal suspect is being questioned while undergoing an fMRI. Undoubtedly, the suspect will be influenced by the questioner's expectancy of guilty (for example via the questioner's tone of voice, their reactions, mannerisms, etc.). Accordingly, these experimenter effects (or questioner expectancy effects) are likely to influence the suspect's neurological processes. They would need to be both recognized and treated as noise, making it more difficult to determine what the observed brain activity actually means.
- (5) The subjectivity inherent in fMRI analysis algorithms needs to be acknowledged and these algorithms opened up for scrutiny. It has been claimed that fMRI output is superior to previous assisted lie detection methods partly because of the automated interpretation of results with computer algorithms which reduce the risk of human error by minimizing experimenter bias and subjectivity (Bilz, 2008–2009; Gerard, 2008). However, algorithms are not purely objective artifacts; they encapsulate and reproduce all the subjectivity, bias, and assumptions of the programmers, and some of these may differ each and every time they are applied. This can introduce errors that bias results in an unevenly distributed way across individuals. Furthermore, as Simpson (2008) notes, separate research groups have devised their own independent statistical methods for identifying brain activity they each believe consistent with lying. This development of diverse approaches increases the variability of results and the general reliability of the technique (see Bennett and Miller, 2010, for a thorough discussion of the issue); the consequences of which should not be underestimated and will be amplified in an applied setting. For if competing (but accepted) algorithms produce conflicting results when interpreting fMRI questioning data then opposing prosecution and defense experts will each exploit the one which best serves their purposes in court leading to a stalemate of probative value and a negation of fMRI evidence.
- (6) We need to determine the percentage of the population who for various reasons are unable to undertake an fMRI, as well as the nature of those reasons. fMRI is highly sensitive to movement, requiring subjects remain virtually motionless for long periods of time during questioning for the slightest head movement can wreck the resulting image. According to a review by Alexander (2007) of fMRI trials published between 2001 and 2006, ~20% of subjects (38 of 192) were rejected because of head motion artifacts or insufficient data. Add to this the physical construction of MRI machines which involves the use of powerful magnets and the number of groups who will not be able to undergo fMRI questioning grows. This includes those with medical conditions such as Parkinson disease which prevents then remaining still without medication, those suffering claustrophobia, people with medical implants, metal pins, piercings, and shrapnel; all of which may preclude fMRI questioning (Gerard, 2008; Holley, 2009; see Appendix A).
- (7) Questions over the methodological validity of past and future fMRI studies must be answered. Many of the fMRI trials to date have compared group differences rather than individuals, and the few accuracy levels reported range between 78 and 90%. This discrepancy within detection rates counts against fMRI gaining admissibility within criminal trials and will not be remedied until more studies are published for peer review (Gerard, 2008). Furthermore, according to Moreno (2009) a recent meta-analysis of existing neuroscience data taken from published studies revealed that more than half of these employed defective research methods producing dubious results as a result of distorted data and biased correlation analysis; specifically *non-independence error* (see Vul et al., 2009).
- (8) To attain external validity, experiments need to be applicable beyond highly controlled laboratory settings to confrontational, emotional “high-stakes” criminal justice situations. Criminal investigations and trials are confrontational and may represent high personal stakes for those involved. This will affect the individuals' mental state and underlying neurological processes (see e.g., Sip et al., 2010, 2012; for evidence of modulation on both deception behavior and its brain correlates in social contexts). As Andrewartha (2008, p. 93) states, “[t]his perhaps explains why considerable judicial criticism has been made of the purported reliability of utilizing lie detector machines in litigation. The propriety of equating simulated scientific testing with real life scenarios for the purpose of evidence is highly questionable.” Unless researchers can show fMRI testing is suitably robust for the criminal

law setting then this technology will remain unacceptable. Randomized controlled trials with currently available testing protocols may not be a straightforward solution to this. In addition to result interpretation problems and the current lack of a comprehensive model of deception, “translational validation” requires access to real-world situations with minimal interference and the possibility to derive an objective index of performance for deception detection. The outcome of court proceedings, for example, could not be taken as an objective parameter for the discrimination between lie and truth (whatever lie detection task is being translationally validated). It is instead already possible to predict that the introduction of fMRI evidence will significantly influence juror decision-making, if unchallenged (see e.g., McCabe et al., 2011).

According to skeptics the enthusiasm for brain imaging and related “mind reading” applications largely overestimate its current ability to identify unique neural correlates of complex mental functions such as lying (but see Haynes and Rees, 2006; Haynes, 2008). Brain activations look extremely persuasive but they result from a long series of manipulations, assumptions, and interpretations. A precise and robust model of the mental processes involved in lying should guide hypotheses about brain activations, however such a generally accepted model remains absent. In addition, lies can be of different types (i.e., denying an event that has occurred vs. making up a slightly different story vs. telling a truth which will be interpreted as a lie; for example, think of a betrayed partner asking “with whom have you had dinner last night?” and the cheater sarcastically replying: “with my lover, obviously!” thereby telling the truth with the intent to make it sound like a lie). The context of basic fMRI experiments is artificial and one often has to sacrifice external for internal validity, and any attempts to make them more similar to real world scenarios will almost inevitably undermine internal validity. Finally the available literature cannot be generalized to all populations for lie detection protocols have not been tested on juveniles, the elderly, or individuals with problems of substance abuse, antisocial personality, mental retardation, head injury, dementia.

In summary, while fMRI may be a useful research tool in combination with other techniques to clarify the mechanisms involved in lying, and its degree of sensitivity and specificity in lie detection may be higher than that of the polygraph, most scientists currently agree that fMRI research evidence is still weak and lacks both external and construct validity (Spence, 2008). We also must conclude that the current state of the science does not at this time meet the legal standards for admissibility in court proceedings (see Simpson, 2008 and Merikangas, 2008, for exhaustive discussions).

LEGAL AND ETHICAL HURDLES

Since the 1920s proponents of assisted lie detection technologies have been predicting their inevitable acceptance by courts; first for polygraphs and now for neuroimaging technologies (Gerard, 2008), however to date fMRI evidence has never successfully been admitted in court for determining the veracity of statements by witnesses or defendants. This reflects the deep skepticism held

by the judiciary as to the reliability of assisted lie detection techniques. This skepticism is partly borne out of the failure of the polygraph and now threatens to taint this new generation of neuroimaging technologies. The perverse irony for the cognitive neuroscientists who have been developing these new technologies in a conscious effort to address the legal short-comings of polygraphs, is that while techniques like fMRI might well-tick the boxes of reliability and objectivity when perfected, the solution of bypassing physiological responses in favor of the direct recording of neural activity may itself constitute grounds for the judiciary to reject neuroimaging technologies. Not because such solutions will necessarily lack reliability or objectivity, but because they potentially infringe other human/constitutional rights and legal principles. The developers of neuroimaging technologies need to acknowledge and engage with these legal issues *before* they seek to impose their new techniques into criminal courts if they are to maximize their chances of winning over the already skeptical judicial gate-keepers. For should they fail to find a way to square their new technologies with the existing legal principles set out below, then without legislative intervention their technologies will remain excluded from criminal courts.

- (1) Possible constitutional and human rights violations (illegal search, right to silence, freedom of thought, right to privacy, human dignity, right to integrity of the person, and protection of personal data):

Looking across various common law legal systems, a number of constitutional principles, and human rights conventions¹⁰ will be engaged to differing degrees within different jurisdictions by the neuroimaging processes of fMRI. Ultimately without legislative intervention it will be the respective national courts who will be forced to rule on each of these issues, either when parties first seek to introduce fMRI evidence of statement veracity into criminal trials, or upon appeals to the first convictions/acquittals where this technology played a material part in arriving at a verdict. It is not our intention to examine each of these in depth here, rather to discuss broadly the various legal hurdles which must be addressed if fMRI is to find its place within criminal trials for determining the veracity of statements made.

The first set of issues is whether fMRI questioning constitutes a *search* of the subject, and when such a search will be considered lawful or unlawful. Discussions in this area tend to center on the US Constitutional Fourth Amendment protecting against unreasonable or unlawful searches (see Pardo, 2006, and Holley, 2009, amongst others for in-depth discussions on this point). A view exists that neuroimaging techniques will constitute a legitimate search under established legal doctrine should neural activity be

¹⁰The sources examined here are: the US Constitution (limited in application to US citizens within US territories), the European Convention on Human Rights (produced by, and applying to, the 47 member states of the Council of Europe, and overseen by the European Court of Human Rights), and the Charter of Fundamental Rights of the European Union (applicable to all citizens and residents of the 28 member states of the EU, this Charter enshrines a range of personal, civil, and social rights and existing conventions and treaties (including the European Convention on Human Rights) into EU law thus ensuring their legal certainty).

equated to other forms of physical evidence gathered from the human body, such as blood or DNA sampling, fingerprints, voice sampling, etc., providing probable cause exists justifying such sampling (Pardo, 2006). However, it is easy to conceptualize neural activity as distinct from other forms of physiological evidence. For example, while we can manipulate neural activity by conducting mathematical problems in our head, we cannot change our DNA profile through thought processes. What legal weight such a distinction would carry is moot until tested in court. The more challenging question is whether or not authorities should be allowed to record our neural activity without our consent, or even our knowledge. When confronted with this problem courts will be forced to either shoehorn this new technology into existing legal frameworks governing conceptually similar subject matter (i.e., DNA, blood, fingerprints, etc.) or produce new bespoke legal frameworks for their governance. In the latter case, the form of any new framework cannot be predicted. A final difficult question is whether police could require a person to undergo an fMRI test without a warrant, with no clear consensus existing between commentators on this point (see Pardo, 2006 and Holley, 2009).

Another set of issues is whether fMRI questioning undermines the right to silence and the right not to self-incriminate. Neuroimaging technology has the potential to undermine these rights if it can operate without the individual needing to speak. Within the United States, the Supreme Court has previously speculated that “the involuntary transmission of incriminating lie-detection evidence would violate a suspect’s right to silence” (Simpson, 2008: 767). Under the European Convention on Human Rights (ECHR) whilst there is no explicit protection against self-incrimination, in the case of *Funke v France* (A/256-A, 1993; 1 C.M.L.R. 897, ECHR) the European Court of Human Rights (ECtHR) was explicit that the right not to self-incriminate is an implicit component of one’s *Right to a fair trial* under Article 6 ECHR (Jackson, 2009) though it is not an absolute right (Berger, 2006). The ECtHR in *Saunders v United Kingdom* (1997, 23 EHRR 313) drew a distinction between material which respects the will of the suspect to remain silent and materials which exists independent of the suspect’s will such as DNA, blood, urine, and breath. Unfortunately what they left for a future court to decide is whether or not an individual’s brain activity exists independent of their will to remain silent?

It must also be asked whether questioning in fMRI without consent engages Article 8 Right to respect for private and family life and Article 9 Freedom of thought, conscience and religion of the ECHR. Article 8(1) has been broadly interpreted in the past, and it is readily conceivable that processes which seek to determine the veracity of our statements by measuring neural activity will engage this right. The question will turn on whether or not police will be able to conduct questioning in fMRI under the Article 8(2) qualifications of national security, public safety, and crime prevention, and what protections will be needed to ensure proportionality. The answer to this might well be tied with Article 9, for to allow the state to access thoughts without consent and knowledge may have a chilling effect on both individuals and society as they seek to exercise their freedom of thought. Courts may well seek to impose stringent safeguards on neuroimaging

technology to prevent both the overuse and misuse of them if they feel these rights are threatened.

Additionally a number of rights within the Charter of Fundamental Rights of the European Union (CFREU) may also prove challenging for fMRI use in court proceedings. Article 1 *Human dignity* states that human dignity is inviolable and must be respected and protected. It is possible to argue that fMRI questioning without consent undermines an individual’s dignity. Article 3 *Right to the integrity of the person* potentially poses the greatest challenge; especially Article 3(1) that everyone has the right to respect for their *physical* and *mental integrity*. This right may also be engaged by enforced or non-consensual fMRI questioning, especially given the express recognition of both physical and mental components. It is worth noting here that within France, Art.45 of LOI n°2011-814 of 7 July 2011 (i.e., post ratification of the CFREU) which created Art.16-14 of the French Civil Code¹¹ specifically limits the use of brain imaging techniques to medical purposes, scientific research, and in the context of judicial enquiries carried out by experts. But most importantly that the express and informed consent of the individual *must be obtained in writing* prior to *any imaging*, and that this consent is revocable at any time. Given how few national legislators have specifically acknowledged the use of neuroimaging in judicial proceedings, let alone the issue of consent, this early approach by the French government takes on considerable significance.

Finally Article 8 CFREU *Protection of personal data* poses a number of interesting questions. Firstly would fMRI data constitute personal data and thus fall under the protection of this article? Given the uniquely nature of such imaging in relation to an individual this must surely be the case. Assuming this is correct, under Article 8(2) everyone has the right to access their personal data and to have it rectified. As a result given the current fallibility of fMRI evidence, one could always argue that an interpretation of such data is incorrect and must be rectified. It would be interesting to see the effects on the admissibility of fMRI evidence in courts where one party seeks to challenge the ostensibly incorrect interpretation of their fMRI results by the other.

(2) Compelled questioning and covert surveillance:

The issues of compelled questioning and (as the technology develops) covert neuroimaging surveillance are ones which courts will be forced to face given the potentially profound impact covert surveillance of this nature will have on society as a whole. One of the concerns raised is the potential for authorities to use these technologies for *fishing trips* whereby police would question an individual to determine whether they have committed criminal acts in the past without any pre-existing evidence or reasonable suspicion. For police to search before they suspect is to undermine the presumption of innocence upon which our common law legal systems are built. Though as Pardo (2006) notes, it is not certain that such actions will be prevented under current regimes.

¹¹Taken from the French Civil Code, Book I: People, Title I: The Civil Rights, Chapter IV: the use of brain imaging techniques.

(3) Probative value, unfair prejudice, and undermining the Province of the Jury:

For evidence (including scientific evidence) to be admissible in criminal courts it must be *relevant*, thus possessing *probative value*. The probative value of evidence can be defined as “the extent to which [this evidence] increases or decreases the probability of a fact in issue” (Dennis, 2007, p.108). Thus, for fMRI evidence the probative value is the extent to which it increases or decreases the *subjective veracity probability* of a declarant’s statement; i.e., how it affects the factual probability that a person does or does not believe what they are saying.

However, probative value also refers to the *degree of relevance* evidence possesses, which is the extent to which evidence influences the probability of a fact in issue in the mind of a rational juror. Within England and Wales if the judge considers the probative value of evidence will have a prejudicial effect on this juror “disproportionate to the rational strength of the evidence as a means of proof, [then] the exclusionary discretion is available [to the judge] to prevent an accused suffering prejudice” (Dennis, 2007, p.108); thus the judge can exclude such disproportionate evidence. An example of evidence excluded may include full-color graphic photographs of injuries when a party seeks to admit these in addition to clear and factual medical reports. Similarly, within the US Federal Rules of Evidence (Rule 403, Federal Rules of Evidence) where the probative value of otherwise admissible evidence is substantially outweighed by the danger of unfair prejudice then this evidence too can be excluded. The question therefore becomes, will the courts reject *prima facie* admissible fMRI evidence on the basis that, because of its nature (or its presentation) it risks unfairly prejudicing the accused? Fears have been raised that: the graphic nature of fMRI evidence will result in unfair prejudice; that scientific lie detection evidence will unduly influence and taint jury deliberations; that jurors will not use their intuition and independent reasoning to critically challenge neuroimaging evidence; and that through function-creep such evidence will trespass into the *Province of the Jury* by effectively usurping their role as arbiter of fact (Gerard, 2008; see Weisberg et al., 2008 for evidence that fMRI images are seen as more compelling than other types and formats of data).

Supporters of neuroscience technologies consider concerns over the undermining of judges and juries as unfounded; rather neuroimaging evidence will simply make the predictions of veracity by jurors and judges more reliable (Pardo, 2006). Pardo makes the argument that:

Because even a highly reliable neuroscience test would not establish knowledge or lies directly, jurors would still need to play their traditional role in assessing it. In making these assessments, the jury would, for example, consider whether other evidence regarding credibility should override the test results, rendering the test conclusion unlikely. (2006: 318)

While this statement seeks to defend and support fMRI in criminal courts, it unwittingly demonstrates the danger that this technology will import unfair prejudice into criminal trials. To

explain; by rightly accepting that even a highly reliable neuroimaging test does not directly establish knowledge or lies one must ask “what is the point of introducing evidence to a jury from a technology that cannot provide direct evidence as to the veracity of statements made but is still marketed and promoted as a scientifically accurate lie detector¹²?” The obvious danger here is that the nuance between *lie detection* and *statement veracity* will not be clearly explained at the start of a case and/or not maintained and reinforced as the case progresses, leading juries to overestimate the capabilities of this technology. This is highlighted by the remainder of the above quote where the neuroimaging test results are already being presented by the author as the *de facto* position of truth, one which can only be overridden should *other evidence regarding credibility override the test results*; i.e., the tests shall be the truth unless you can prove otherwise. This statement, while seeking to defend neuroimaging technologies, actually serves to highlight the potential disproportional probative effect of neuroscience lie detectors. Cognitive neuroscientists must be careful not to overplay what these technologies can offer criminal courts nor their vision of the potential future role of neuroscience within criminal courts, lest they overplay themselves out of the courtroom altogether.

(4) Right to a fair trial:

Depending on how questioning in fMRI is conducted for criminal trials it can be argued that the fairness of trials will be placed at risk unless *all* parties to the trial are subjected to pre-trial questioning in fMRI. Presenting fMRI evidence from only one party to the case may result in an artificial disparity of evidence; i.e., the neuroimaging evidence plus testimony vs. testimony without neuroimaging evidence. Justice may now depend on whether or not a jury will question a technology promoted as a highly accurate lie detector, so to ensure parity of arms and a fair trial all parties should be subjected to pre-trial questioning in fMRI if they are ever introduced. Of course such a scenario depends on all the parties being capable of undergoing fMRI testing which is not the case when the victim is dead or comatose. In these circumstances fMRI evidence may need to be prohibited.

Nevertheless, it is conceivable that fMRI evidence may become admissible solely as a defense instrument given that structural brain scans have already found acceptance and are widely admissible as mitigating evidence during sentencing proceedings. Indeed such use may help ensure a trial is ultimately fair. However, any attempt to extrapolate from this niche application such that fMRI evidence can be used throughout the entirety of

¹²Both of the two commercial companies offering fMRI detection service specifically and deliberately promote their technologies as scientific lie detection tools and not as veracity probability enhancement tools; No Lie MRI claims their technology ‘represents the first and only direct measure of truth verification and lie detection in human history’ (see <http://noliemri.com/>), while CEPHOS Corp claims to have developed ‘the latest, most scientifically advanced, brain imaging techniques for scientifically accurate lie detection’ (see <http://www.cephoscorp.com/>).

the trial *but only by the defense* represents an arguably unacceptable asymmetry of measures; one which potentially undermines the overall fairness of the trial (both actual and perceived) and the rights of victims.

CONCLUDING POINTS

Our discussion throughout has focused on the scientific, legal, and ethical hurdles facing those seeking to introduce fMRI evidence into trials as a means of assisting judges and juries in determining the veracity of statements made. Schauer (2010) suggests that as the goals of the law differ from those of science, what is not good enough for science may yet be good enough for the law and *vice versa*. However, following our assessments of the science underpinning fMRI as a lie detector and how this relates to the law, we must conclude that the current state of this technology, and potentially the technology *per se*, fails to meet either acceptable scientific or legal standards.

The evaluation of fMRI accuracy in lie detection—in some cases claimed to be as high as 0.90—is indeed based on laboratory experiments conducted with compliant participants, which is unlikely to be true of most legal settings where non-compliance and the use of countermeasures would make its accuracy figure drop dramatically (e.g., Ganis et al., 2011). In the cognitive neurosciences fMRI is not sufficient by itself to unveil which brain areas are epiphenomenal, which are strictly necessary to lying. It may thus pick up some noise together with the real signal. Even if it was possible to produce a correlational map where a constant pattern could be detected indicating a lie, issues of replicability and generalizability across conditions and participants could be raised. And even more so in those cases where facts are unknown to the tester and there is no objective reality against which to establish whether a person is lying or not. From the legal perspective, until the science behind fMRI testing improves it will not meet the relevance and reliability thresholds required for any scientific evidence to be admissible in criminal trials. The assumptions, inferences, and questions of internal validity which so pervade current fMRI testing and analysis need to be addressed. As does the challenge of successfully applying this technology to criminal justice scenarios characterized by their confrontational emotional nature and the personal high-stakes involved for the participants.

Neuroimaging in courts also raises the specter of potential constitutional and human rights violations. Questions arise as to whether or not such testing constitutes an illegal search as well as how it respects rights to privacy, silence, thought, and a fair trial are all engaged by this technology yet left unanswered. Unless the admissibility decision is taken out of the hands of the judiciary by politicians, which is itself a likely scenario, ultimately it will be for the courts to decide the fate of fMRI evidence in criminal trials. Given the range and depth of the legal and ethical issues identified in the earlier sections, the likely outcome will probably fall on a spectrum somewhere between outright rejection through to some form of restricted and regulated usage, as opposed to the highly unlikely scenario of *carte blanche* acceptance.

What we have not discussed within this paper are both the *operational* and *social* barriers to the widespread use of fMRI testing in criminal trials. These barriers are potentially just as daunting as their scientific and legal counterparts.

From the *practical operational perspective*, issues which require future examination include; the cost of purchasing, staffing, and maintaining sufficient fMRI machines to cater for a national justice system; the additional time and monetary costs fMRI testing will add to criminal cases; how fMRI testing can be made to work within adversarial systems of questioning and cross-examination based on earlier responses; and the lack of a courtroom-friendly portable fMRI system. Additionally there are questions specific to the assessment algorithms used when interpreting fMRI response data: will only a single universal *official* algorithm be allowed?; will commercial patented algorithms be admissible if they are not completely open for inspection and independent verification?; and finally what happens when new algorithms and new fMRI scanners are inevitably developed as the science is continually refined which prove to be more reliable and sensitive than previous algorithms/machines? It is conceivable that those who maintain their innocence and are appealing their conviction under the previous technology will seek to be re-tested with the new machines and the new algorithm in an effort to prove their innocence placing a further burden on the criminal justice system. All of these points possess the potential to impact upon the fairness of future trials.

A final hurdle to the widespread introduction of fMRI testing is *societal acceptability*, without which technologies such as neuroimaging techniques for determining the veracity of statements within criminal trials will lack both public confidence and legitimacy. Future research needs to gauge the levels of public support for such technologies, for even if neuroimaging proves superior to humans as arbiters of statement veracity in criminal courts, this fact in of itself may not be enough for the public to accept their introduction if they are apprehensive or hostile to what such technologies represent for their future. We cannot escape from asking the question, *will people accept mind reading machines?* This is obviously not what the current generation of neuroimaging technologies is, but they are a small step down this long path.

Our societies have developed to both accept and respect an individual's right to keep secrets, and in so doing they do not seek to override human beings' evolved capacity to keep secrets, for a society where individuals are denied secrets is not a human society as we know it. The developers and proponents of fMRI testing must respect this fact and engage society in their research as it progresses. Otherwise they may find they successfully negotiate the frying-pan of scientific and technical challenges in perfecting fMRI testing only to be consumed by a fire of legal, ethical, social, and political opposition.

ACKNOWLEDGMENTS

The authors would like to thank Roberto Cubelli for his comments on a previous version of the manuscript, the Editor and three Reviewers for their very constructive comments. This study was funded by EPSRC (grant number EP/G037264/1).

REFERENCES

- Abbott, A. (2001). Into the mind of a killer. *Nature* 410, 296–298. doi: 10.1038/35066717
- Abbott, A. (2007). Abnormal neuroscience: scanning psychopaths. *Nature* 450, 942–944. doi: 10.1038/450942a
- Abe, N. (2011). How the brain shapes deception: an integrated review of the literature. *Neuroscientist* 17, 560–574. doi: 10.1177/1073858410393359
- Aharoni, E., Sinnott-Armstrong, W., and Kiehl, K. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *J. Abnorm. Psychol.* 121, 484–497. doi: 10.1037/a0024796
- Alexander, A. (2007). Functional magnetic resonance imaging lie detection: is a “Brainstorm” heading toward the “Gatekeeper”? *Houston J. Health Law Policy* 7, 1–56.
- Andrewartha, D. (2008). Lie detection in litigation: science or prejudice? *Psychiatry Psychol. Law* 15, 88–104. doi: 10.1080/13218710701873940
- Bennett, C., and Miller, M. (2010). How reliable are the results from functional magnetic resonance imaging? *Ann. N.Y. Acad. Sci.* 1191, 133–155. doi: 10.1111/j.1749-6632.2010.05446.x
- Berger, M. (2006). Compelled self-reporting and the principle against compelled self incrimination: some comparative perspectives. *Eur. Hum. Rights Law Rev.* 1, 25–38.
- Bilz, K. (2008–2009). Self-Incrimination doctrine is dead; long live self-incrimination doctrine: confessions, scientific evidence, and the anxieties of the liberal state. *Cardozo Law Rev.* 30, 807–869.
- Bond, C., and DePaulo, B. (2006). Accuracy of deception judgments. *Pers. Soc. Psychol. Rev.* 10, 214–234. doi: 10.1207/s15327957pspr1003_2
- Christ, S., Van Essen, D., Watson, J., Brubaker, L., and McDermott, K. (2009). The Contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cereb. Cortex* 19, 1557–1566. doi: 10.1093/cercor/bhn189
- Coricelli, G., and Rusconi, E. (2011). “Probing the decisional brain with TMS and tDCS,” in *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*, eds M. Schulte-Mecklenbeck, A. Kühberger, and R. Ranyard (New York, NY: Taylor and Francis), 205–222.
- Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York, NY: Putnam Publishing.
- Davatzikos, C., Ruparel, K., Fan, Y., Shen, D., Acharyya, M., Loughhead, J., et al. (2005). Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage* 28, 663–668. doi: 10.1016/j.neuroimage.2005.08.009
- Dennis, I. (2007). *The Law of Evidence*. London: Sweet and Maxwell Ltd.
- Ellenberg, C. (2009). Lie detection: a changing of the guard in the quest for truth in court? *Law Psychol. Rev.* 33, 139–149.
- Erickson, S. (2010). Blaming the brain. *Minn. J. Law Sci. Technol.* 11, 22–77.
- Fox, D. (2009). The right to silence as protecting mental control. *Akron Law Rev.* 42, 763–801. doi: 10.1093/acprof:oso/9780199599844.003.0017
- Gamer, M. (2011). “Detecting of deception and concealed information using neuroimaging techniques,” in *Memory Detection: Theory and Application of the Concealed Information Test*, eds B. Verschuere, G. Ben-Shakhar, and E. Meijer (New York, NY: Cambridge University Press), 90–113. doi: 10.1017/CBO9780511975196.006
- Ganis, G., Kosslyn, S., Stose, S., Thompson, W., and Yurgelun-Todd, D. (2003). Neural correlates of different types of deception: an fMRI investigation. *Cereb. Cortex* 13, 830–836. doi: 10.1093/cercor/13.8.830
- Ganis, G., Rosenfeld, J., Meixner, J., Kievit, R., and Schendan, H. (2011). Lying in the scanner: covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *Neuroimage* 55, 312–319. doi: 10.1016/j.neuroimage.2010.11.025
- Gazzaniga, M. (2008). The law and neuroscience. *Neuron* 60, 412–415. doi: 10.1016/j.neuron.2008.10.022
- Gerard, E. (2008). Waiting in the Wings? the admissibility of neuroimaging for lie detection. *Dev. Ment. Health Law* 27, 1–32.
- Giner-Sorolla, R. (2012). Science or art? How aesthetic standards grease the way through the publication bottleneck but undermine science. *Perspect. Psychol. Sci.* 7, 562–571. doi: 10.1177/1745691612457576
- Grafton, S., Sinnott-Armstrong, W., Gazzaniga, S., and Gazzaniga, M. (2006). Brain scans go legal. *Sci. Am.* 17, 30–37. doi: 10.1038/scientificamericanmind1206-30
- Hakun, J., Ruparel, K., Seelig, D., Busch, E., Loughhead, J., Gur, R., et al. (2009). Towards clinical trials of lie detection with fMRI. *Soc. Neurosci.* 4, 518–527. doi: 10.1080/17470910802188370
- Haynes, J. (2008). Detecting deception from neuroimaging signals – a data-driven perspective. *Trends Cogn. Sci.* 12, 126–127. doi: 10.1016/j.tics.2008.01.003
- Haynes, J., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534. doi: 10.1038/nrn1931
- Holley, B. (2009). It's All in your head: neurotechnological lie detection and the fourth and fifth amendments. *Dev. Ment. Health Law* 28, 1–24.
- Houston, P., Floyd, M., and Carnicero, S. (2012). *Spy the Lie*. New York, NY: St Martin's Press.
- Huettel, S., Song, A., and McCarthy, G. (2009). *Functional Magnetic Resonance Imaging, 2nd Edn*. Sunderland: Sinauer Associates Inc.
- Hughes, V. (2010). Science in court: head case. *Nature* 464, 340–342. doi: 10.1038/464340a
- Jackson, J. (2009). Re-conceptualizing the right of silence as an effective fair trial standard. *Int. Comp. Law Q.* 58, 835–861. doi: 10.1017/S0020589309001407
- Karim, A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., et al. (2010). The truth about lying: inhibition of the anterior prefrontal cortex improves deceptive behavior. *Cereb. Cortex* 20, 205–213. doi: 10.1093/cercor/bhp090
- Kozel, F. A., Johnson, K., Mu, Q., Grenesko, E., Laken, S., and George, M. (2005). Detecting deception using functional magnetic resonance imaging. *Biol. Psychiatry* 58, 605–613. doi: 10.1016/j.biopsych.2005.07.040
- Langleben, D. D., and Moriarty, J. C. (2013). Using brain imaging for lie-detection: where science, law and research policy collide. *Psychol. Public Policy Law* 19, 222–234. doi: 10.1037/a0028841
- Langleben, D., Dattilo, F., and Guthei, T. (2006). True lies: delusions and lie-detection technology. *J. Psychiatry Law* 34, 351–370.
- Logothetis, N. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878. doi: 10.1038/nature06976
- Logothetis, N., and Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magn. Reson. Imaging* 22, 1517–1531. doi: 10.1016/j.mri.2004.10.018
- Luber, B., Fisher, C., Appelbaum, P., Ploesser, M., and Lisanby, S. (2009). Non-invasive brain stimulation in the detection of deception: scientific challenges and ethical consequences. *Behav. Sci. Law* 27, 191–208. doi: 10.1002/bsl.860
- Mameli, F., Mrakic-Sposta, S., Vergari, M., Fumagalli, M., Macis, M., Ferrucci, R., et al. (2010). Dorsolateral prefrontal cortex specifically processes general – but not personal – knowledge deception: multiple brain networks for lying. *Behav. Brain Res.* 211, 164–168. doi: 10.1016/j.bbr.2010.03.024
- McCabe, D. P., Castel, A. D., and Rhodes, M. G. (2011). The influence of fMRI lie detection evidence on juror decision-making. *Behav. Sci. Law* 29, 566–577. doi: 10.1002/bsl.993
- Merikangas, J. (2008). Commentary: functional MRI lie detection. *J. Am. Acad. Psychiatry Law* 36, 491–498.
- Moreno, J. (2009). The future of neuroimaged lie detection and the law. *Akron Law Rev.* 42, 717–737.
- Nitsche, M., Cohen, L., Wassermann, E., Priori, A., Lang, N., Antal, A., et al. (2008). Transcranial direct current stimulation: state of the art 2008. *Brain Stimul.* 1, 206–223. doi: 10.1016/j.brs.2008.06.004
- Pardo, M. (2006). Neuroscience evidence, legal culture, and criminal procedure. *Am. J. Crim. Law* 33, 301–337.
- Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63. doi: 10.1016/j.tics.2005.12.004
- Poldrack, R. (2010). Mapping mental function to brain structure: how can cognitive neuroimaging succeed? *Perspect. Psychol. Sci.* 5, 753–761. doi: 10.1177/1745691610388777
- Priori, A., Mameli, F., Cogiamanian, F., Marceglia, S., Tiriticco, M., Mrakic-Sposta, S., et al. (2008). Lie-specific involvement of dorsolateral prefrontal cortex in deception. *Cereb. Cortex* 18, 451–455. doi: 10.1093/cercor/bhm088
- Purves, D., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., and Woldorff, M. G. (2013). *Principles of Cognitive Neuroscience*. Sunderland: Sinauer Associates Inc.
- Sandrini, M., Umiltà, C., and Rusconi, E. (2011). The use of transcranial magnetic stimulation in cognitive neuroscience: a new synthesis of methodological issues. *Neurosci.*

- Biobehav. Rev.* 35, 516–536. doi: 10.1016/j.neubiorev.2010.06.005
- Sartori, G., and Umiltà, C. (2002). How to avoid the fallacies of cognitive subtraction in brain imaging. *Brain Lang.* 74, 191–212. doi: 10.1006/brln.2000.2334
- Schauer, F. (2010). Neuroscience, lie-detection, and the law. *Trends Cogn. Sci.* 14, 101–103. doi: 10.1016/j.tics.2009.12.004
- Simpson, J. (2008). Functional MRI lie detection: too good to be true? *J. Am. Acad. Psychiatry Law* 36, 491–498.
- Sip, K., Lynge, M., Wallentin, M., McGregor, W., Frith, C., and Roepstorff, A. (2010). The production and detection of deception in an interactive game. *Neuropsychologia* 48, 3619–3626. doi: 10.1016/j.neuropsychologia.2010.08.013
- Sip, K., Roepstorff, A., McGregor, W., and Frith, C. (2007). Detecting deception: the scope and limits. *Trends Cogn. Sci.* 12, 48–53. doi: 10.1016/j.tics.2007.11.008
- Sip, K., Roepstorff, A., McGregor, W., and Frith, C. (2008). Response to Haynes: there's more to deception than brain activity. *Trends Cogn. Sci.* 12, 127–128. doi: 10.1016/j.tics.2008.01.004
- Sip, K., Skewes, J., Marchant, J., McGregor, W., Roepstorff, A., and Frith, C. (2012). What if I get busted? deception, choice, and decision-making in social interaction. *Front. Neurosci.* 6:58. doi: 10.3389/fnins.2012.00058
- Spence, S. (2008). Playing devil's advocate: the case against fMRI lie detection. *Legal Criminol. Psychol.* 13, 11–25. doi: 10.1348/135532507X251597
- Tovino, S. (2007). Functional neuroimaging and the law: trends and directions for future scholarship. *Am. J. Bioeth.* 7, 44–56. doi: 10.1080/15265160701518714
- Vul, E., Harris, C., Winkielman, P., and Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect. Psychol. Sci.* 4, 274–290. doi: 10.1111/j.1745-6924.2009.01125.x
- Weisberg, D., Keil, F., Goodstein, J., Rawson, E., and Gray, J. (2008). The seductive allure of neuroscience explanations. *J. Cogn. Neurosci.* 20, 470–477. doi: 10.1162/jocn.2008.20040
- Received: 31 July 2013; paper pending published: 27 August 2013; accepted: 03 September 2013; published online: 24 September 2013.
- Citation: Rusconi E and Mitchener-Nissen T (2013) Prospects of functional magnetic resonance imaging as lie detector. *Front. Hum. Neurosci.* 7:594. doi: 10.3389/fnhum.2013.00594
- This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Rusconi and Mitchener-Nissen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.



Addressing social resistance in emerging security technologies

Timothy Mitchener-Nissen *

Department of Security and Crime Science, University College London, London, UK

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

Aaron Winter, University of Abertay, UK

Nicola Lettieri, ISFOL, Italy

***Correspondence:**

Timothy Mitchener-Nissen,
Department of Security and Crime
Science, University College London,
35 Tavistock Square, London WC1H
9EZ, UK
e-mail: t.nissen@ucl.ac.uk

In their efforts to enhance the safety and security of citizens, governments and law enforcement agencies look to scientists and engineers to produce modern methods for preventing, detecting, and prosecuting criminal activities. Whole body scanners, lie detection technologies, biometrics, etc., are all being developed for incorporation into the criminal justice apparatus.¹ Yet despite their purported security benefits these technologies often evoke social resistance. Concerns over privacy, ethics, and function-creep appear repeatedly in analyses of these technologies. It is argued here that scientists and engineers continue to pay insufficient attention to this resistance; acknowledging the presence of these social concerns yet failing to meaningfully address them. In so doing they place at risk the very technologies and techniques they are seeking to develop, for socially controversial security technologies face restrictions and in some cases outright banning. By identifying sources of potential social resistance early in the research and design process, scientists can both engage with the public in meaningful debate and modify their security technologies before deployment so as to minimize social resistance and enhance uptake.

Keywords: technology, social, resistance, ethics, security

INTRODUCTION

Social constructionism is a sociological theory of knowledge which holds that our knowledge of the world is not derived from observing nature, rather that it is constructed through the social interactions and processes of people (Burr, 2003). By adopting a social constructionist perspective one can comprehend the phenomena of criminality and criminal behavior as existing from the moment individuals and societies began socially constructing and adopting laws which proscribed certain acts or omissions as constituting criminal activities (Newburn, 2007). Under this formulation that which is considered *criminal* can differ spatially (different countries, states, districts, towns have different laws), by ascribed categories (i.e., different laws for different religious groups, genders, sexual orientations, professions and/or social classes) and temporally (laws are not “set-in-stone,” rather are subject to change). Yet while laws can change to reflect both prevailing social views and the organization of activities within a society, the slow pace of this change often results

in the law struggling to catch up. The advent of the digital age, the pace of technological development, and the widespread adoption of technologies in many societies all pose challenges for the application of existing laws and the timely creation of new ones.

This paper begins by examining the phenomenon whereby states embrace technologies as solutions or *fixes* for the problem of crime. The negative consequences of this policy in the form of social resistance are then discussed. Finally the question is asked as to why the design and implementation of emerging security technologies continues to repeat mistakes observed in previous technologies? Four answers are provided here, including; (i) the paucity of social education within science, technology, engineering, and mathematics (STEM) courses, (ii) the lack of priority afforded social and ethical issues within the research and design of security technologies, (iii) a general failure by STEM practitioners in comprehending the importance of social acceptability to the technologies they create, and (iv) restricted public engagement.

SECURITY TECHNOLOGIES AS TECHNOLOGICAL FIXES

The development and interpretation of new technological advancements have been adopted with considerable enthusiasm by governments, law enforcements agencies, universities and private companies as potential methods for preventing, detecting, and prosecuting criminal activities. In this regard they represent *technological fixes* for the social problem of crime; a technological fix is broadly defined as a technological solution for *solving* social problems (Weinberg, 1967) reflecting the views of technological optimists. Technology

¹ I am using this umbrella term to cover all the organisations involved in every stage of the prevention, detection, and prosecution of criminal activities. This includes: (i) the work of the security services with their roles of collecting intelligence (both domestic and international) to protect the national security and economic well-being of a nation as well as supporting the prevention and detection of serious crimes; (ii) domestic law enforcements organisations such as police and border agencies with their various roles in preventing, detecting and deterring criminal activities, as well as gathering evidence to assist in the prosecution of those accused of committing crimes; and (iii) the criminal court system, including the prosecution and defence who make use of scientific evidence and experts when furthering the case of their clients.

is presented as a panacea for social problems by being cheaper and more effective than alternative human-centric approaches for dealing with issues which negatively impact society.

The current range of technological fixes designed specifically for addressing crime (hereafter referred to as *security technologies*²) continues to increase as scientists and engineers seek to apply the knowledge and approaches of their specific fields to this particular goal. Whole body scanners at airports utilize X-ray backscattering or millimeter wave technology so as to identify metallic and non-metallic objects, plastic and liquid explosives, flora, fauna, drugs, and cash, concealed within or beneath the clothing of passengers (European Commission, 2010; Mitchener-Nissen et al., 2012). Data mining, being the application of database technology and techniques (such as modeling and statistical analysis) to data to identify valid, novel, implicit and potentially useful information and patterns within that data, is employed with the aim of analysing intelligence and detecting terrorist activities, fraud, and other criminal patterns (Tien, 2004; Steinbock, 2005; Schermer, 2011). The use of biometrics enables crime-scene technologies that can assist in the identification and prosecution of offenders (such as DNA databases and fingerprinting technologies), tackling identity fraud, and counteracting illegal immigration (Grijpink, 2006; Goldstein et al., 2008). And to assist in the investigation and prosecution of criminal acts, lie detection technologies designed to directly access brain function (including fMRI and EEG) are trying to be developed by researchers and private companies (Wolpe et al., 2010). This selection represents a tiny snapshot of the cornucopia of security technologies both under development and already implemented.

RESULTING SOCIETAL RESISTANCE

Without further examination it would be tempting to conclude that security technologies do indeed constitute justifications for Weinberg's vision of technological fixes as the solution to social problems. However, the notion of the technological fix has been subject to robust criticism. It has been described as "a quick cheap fix using inappropriate technology that creates more problems than it solves" (Rosner, 2004). The truth of this statement is evident within the social controversies (or in the case of the lie detection technologies, the possible future social controversies) produced by each of the security technology examples provided above. Whole body scanners have been accused of conducting digital strip-searches (Klitou, 2008), and the backscatter variation is to be removed from US airports because of the images produced. Data mining has been associated with both a fear of totalitarian-style state observation, as well as the targeting of individuals by governments (Steinbock, 2005). Different biometric technologies can discriminate against various groups within society and are plagued by the problem of false positives (Hunter, 2005; Whitley and Hosein, 2010).

²By *security technologies* I am referring to the product of an engineering endeavour which seeks to deter, prevent, detect or prosecute crimes, and/or enhance the security of individuals, their property, or the state (including its infrastructure).

Additionally the UK's DNA database (the largest in the world) has created controversy by holding the details of innocent people and a disproportionate number of samples from ethnic minorities. And the new generation of potential lie-detection technologies have faced criticism over the potential ethical, social, and legal implications of their operation to existing social and legal institutions should they ever be made to definitively and consistently "work." This social resistance to a security technology begins individually, as solitary citizens question the rationale and/or operation of a particular measure. These may be individuals who actively critique government security policy, those who prioritize privacy and liberty, or as is often the case these are individuals who find themselves adversely impacted upon by a security technology without just cause. For example; individuals who are incorrectly prevented from flying because either they have the same name as another person on a no-fly list, or their details have been added in error to such a list without them being previously notified or provided a way to rectify this error. Recognition of an individual's issues with a security technology can now begin to coalesce into social resistance once knowledge of their plight becomes known to others. The media, lawyers, NGO's, social activists, political figures, and independent commissioners amongst others can all assist in raising awareness here, which in turn can influence other citizens thereby snowballing the effect and reducing support for the security technology in question.

The manifestation of social resistance present in the technologies discussed above represents only a snapshot of the controversies produced by security technologies which have in the past undermined their social acceptability and widespread uptake. In an on-going examination of security technologies which have evoked social resistance, I have identified numerous recurring controversies which continue to arise within new security technologies with depressing regularity. These can be organized into eight high-level categories; the causing of physical and mental harm, questions of legality, financial costs, liberties and human rights issues, broader public responses, issues of functionality, security and safety issues, and abuse/misuse issues. A selection of commonly recurring controversies includes; privacy concerns, function creep, false positive/negative rates, lack of public trust, the failure of a technology to achieve what its designers claim it can do, and the potential for the technology to be abused by the state.

WHY NEW SECURITY TECHNOLOGIES REPEAT THE MISTAKES OF THE PAST

The question which needs addressing here is why have lessons not been learnt such that new security technologies consistently evoke such ethical and social controversy? I suggest there are four complementary elements underpinning the answer to this question. The first is the paucity of social and ethical education within university STEM courses. Within university engineering courses in the UK it is highly likely that a student can (and will) complete their education without ever undertaking a single lecture on the importance of identifying and incorporating social and ethics factors into

their work. This is despite the creation of the field of engineering ethics which arose in the early 1980s following a number of technological developments, designs and failures which negatively impacted human wellbeing (Johnson and Wetmore, 2008). The situation is repeated within the hard sciences with the possible exception of medical ethics. For those who counter with the claim that ethics and ethical research is ensured by the presence of university ethics boards; while a particular research or design project may meet all official conduct requirements such that it is considered ethical, this does not mean that what is being undertaken or created will be accepted by the public. The diverse groups which comprise a society ultimately determine what is considered socially or ethically acceptable, and yet university engineering and hard science courses regularly fail new researchers and designers by not equipping them with an understanding of this fact nor the tools to adequately interact with the public.

The second element in the lack of priority afforded social and ethical issues within research and design projects. Interviews with engineers and scientists engaged in the process of designing and developing new security technologies have highlighted a clear hierarchical structure to the design process. For commercial projects it begins with cost; if it is determined that there is not a viable market for a product then it will not be produced. If this test is passed and the project is considered feasible than design specifications are produced in accordance with the client's requirements and the product is created. Similarly with university research projects, the presence of funding and/or the potential for future commercial exploitation dictates the research undertaken. When this is directed toward addressing perceived security deficiencies the focus is on attaining a specific security goal. These processes leave little space for the consideration and incorporation of social and ethical issues—the focus is on “can we achieve what we have set out to achieve,” and not “is this a socially acceptable way of achieving the desired goals” or “are these goals socially acceptable *per se*.”

The third element is a general failure by scientist and engineers to comprehend just how important social acceptability is to the life cycles of their technologies. In the majority, scientists and engineers do not develop an appreciation of the importance of identifying and addressing social concerns until they are confronted by social resistance; a point often reached *after* a product has been released to market.

The fourth element is the challenge of, and the resistance to, achieving effective public engagement in relation to the design of security technologies. The arguments in favor public engagement hold that just as democracy derives its legitimacy through participation, so too will increasing participation within the development of new or controversial technologies help to infuse the finished products with similar legitimacy and reduce societal resistance. The primary argument against is that lay people are handicapped by a lack the technological literacy, or access to and understanding of, security-sensitive intelligence, which together constrain their ability to provide relevant input or

make informed decisions. But as Kleinman (2005) highlights, the flawed nature of such views is driven home by the fact that experts³ are never value-neutral, unbiased, all-seeing individuals; rather are bounded by the nature of their expert knowledge and will necessarily view a phenomenon from a partial perspective. In other words, experts are handicapped to view the world through blinkers and in this respect have similarities with the very lay public whose input they would seek to exclude.

By introducing socially unacceptable technologies in the first place, trust in both the developers and the end-users (i.e., governments and agencies of the state) is threatened, research and design capacity is diverted from acceptable technologies, and money is wasted that could otherwise have been used for legitimate programmes. The challenge becomes identifying what is acceptable and unacceptable *before* a technology is developed and deployed. By accepting that judgments over acceptability of a technology differ between social groups and that rejection of a technology can lead to its permanent inferiority through neglect (MacKenzie and Wajcman, 1999), the consideration of wider social and ethical issues upstream in the design process to anticipate and mitigate negative social reactions becomes both a valid and logical response.

CONCLUSION

The list of technologies developed which have been banned or their use restricted in various societies, (not necessarily because of deficiencies in the underlying science) but because the developers did not seek to anticipate and mitigate social resistance through upstream design modifications is long and growing. It includes backscatter body scanners, instances of data mining, less lethal weapons, polygraph lie detectors, CCTV, national ID card, etc.

To avoid the ignominy of this situation for emerging security technologies developers must take meaningful steps to identify sources of potential social resistance early in the research and design process. This requires truly reflexive engagement with the public to identify concerns which then can be translated into upstream design requirements; thereby heading off social resistance before it coalesces and becomes synonymous with the technology being developed. The enormity of this challenge cannot be overestimated for if a proposed technology cannot be created in such a fashion which respects and reflects the values held within a society, then those developing the technology are wasting valuable time, money and resources on research which will ultimately be rejected.

ACKNOWLEDGMENTS

This research was funded by the Engineering and Physical Sciences Research Council of the United Kingdom through their Centers for Doctoral Training programme, specifically the Security Science Doctoral Research Training Centre (UCL SECRiT) based at University College London.

³In this case STEM practitioners, state officials, and law-enforcement/intelligence officers.

REFERENCES

- Burr, V. (2003). *Social Constructionism*, 2nd Edn. Hove: Routledge.
- European Commission. (2010). *Communication from the Commission to the European Parliament and the Council on the Use of Security Scanners at EU airports* COM(2010) 311, Brussels: European Commission.
- Goldstein, J., Angeletti, R., Holzbach, M., Konrad, D., and Snijder, M. (2008). *Large-scale Biometrics Deployment in Europe: Identifying Challenges and Threats*. Luxembourg: European Commission Joint Research Centre.
- Grijpink, J. (2006). An assessment model for the use of biometrics. *Comput. Law Secur. Rep.* 22, 316–319. doi: 10.1016/j.clsr.2006.03.007
- Hunter, P. (2005). London terrorist attacks heat up identity card debate and highlight uncertainties over their efficacy. *Comput. Fraud Secur.* 7, 4–5. doi: 10.1016/S1361-3723(05)70230-X
- Johnson, D., and Wetmore, J. M. (2008). “STS and ethics: implications for engineering ethics,” in *The Handbook of Science and Technology Studies*, 3rd Edn., eds E. J. Hackett, O. Amsterdamska, M. Lynch, and J. Wajcman (Cambridge, MA: The MIT Press), 567–581.
- Kleinman, D. (2005). *Science and Technology in Society: From Biotechnology to the Internet*. Malden, MA: Blackwell Publishing.
- Klitou, D. (2008). Backscatter body scanners – a strip search by other means. *Comput. Law Secur. Rep.* 24, 316–325. doi: 10.1016/j.clsr.2008.05.005
- MacKenzie, D., and Wajcman, J. (1999). *The social shaping of technology*, 2nd Edn. Buckingham: Open University Press.
- Mitchener-Nissen, T., Bowers, K., and Chetty, K. (2012). Public attitudes to airport security: the case of whole body scanners. *Secur. J.* 25, 229–243. doi: 10.1057/sj.2011.20
- Newburn, T. (2007). *Criminology*. Cullompton: Willan Publishing.
- Rosner, L. (2004). *The Technological Fix: How People Use Technology to Create and Solve Problems*. New York, NY: Routledge.
- Schermer, B. (2011). The limits of privacy in automated profiling and data mining. *Comput. Law Secur. Rev.* 27, 45–52. doi: 10.1016/j.clsr.2010.11.009
- Steinbock, D. (2005). Data matching, data mining, and due process. *Georgia Law Rev.* 40, 1–84.
- Tien, L. (2004). Privacy, technology and data mining. *Ohio North. Univ. Law Rev.* 30, 389–415.
- Weinberg, A. (1967). *Reflections on Big Science*. Cambridge, MA: MIT Press.
- Whitley, E., and Hosein, G. (2010). *Global Challenges for Identity Policies*. Basingstoke: Palgrave Macmillan.
- Wolpe, P., Foster, K., and Langleben, D. (2010). Emerging neurotechnologies for lie-detection: promises and perils. *Am. J. Bioeth.* 10, 40–48. doi: 10.1080/15265161.2010.519238

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 April 2013; accepted: 31 July 2013; published online: 20 August 2013.

Citation: Mitchener-Nissen T (2013) Addressing social resistance in emerging security technologies. *Front. Hum. Neurosci.* 7:483. doi: 10.3389/fnhum.2013.00483

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Mitchener-Nissen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Can laptops be left inside passenger bags if motion imaging is used in X-ray security screening?

Marcia Mendes^{1,2}, Adrian Schwaninger^{1,2*} and Stefan Michel^{1,2}

¹ School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland (FHNW), Olten, Switzerland

² Center for Adaptive Security Research and Applications (CASRA), Zürich, Switzerland

Edited by:

Andrea Szymkowiak, University of
Aberystwyth, UK

Reviewed by:

Elena Rusconi, University College
London, UK

Lynne M. Coventry, Northumbria
University, UK

*Correspondence:

Adrian Schwaninger, School of
Applied Psychology, Institute
Humans in Complex Systems,
University of Applied Sciences and
Arts Northwestern Switzerland,
Riggenbachstrasse 16, 4600 Olten,
Switzerland
e-mail: adrian.schwaninger@fhnw.ch

This paper describes a study where a new X-ray machine for security screening featuring motion imaging (i.e., 5 views of a bag are shown as an image sequence) was evaluated and compared to single view imaging available on conventional X-ray screening systems. More specifically, it was investigated whether with this new technology X-ray screening of passenger bags could be enhanced to such an extent that laptops could be left inside passenger bags, without causing a significant impairment in threat detection performance. An X-ray image interpretation test was created in four different versions, manipulating the factors packing condition (laptop and bag separate vs. laptop in bag) and display condition (single vs. motion imaging). There was a highly significant and large main effect of packing condition. When laptops and bags were screened separately, threat item detection was substantially higher. For display condition, a medium effect was observed. Detection could be slightly enhanced through the application of motion imaging. There was no interaction between display and packing condition, implying that the high negative effect of leaving laptops in passenger bags could not be fully compensated by motion imaging. Additional analyses were carried out to examine effects depending on different threat categories (guns, improvised explosive devices, knives, others), the placement of the threat items (in bag vs. in laptop) and viewpoint (easy vs. difficult view). In summary, although motion imaging provides an enhancement, it is not strong enough to allow leaving laptops in bags for security screening.

Keywords: aviation security, X-ray screening, threat detection, human factors, motion imaging, multiple views, laptop screening

INTRODUCTION

A secure air transportation system is vital for society and economy. Aviation security measures have been increased substantially in response to several successful and attempted terrorist attacks since September 11, 2001. One major aspect in this field is the mandatory process of baggage screening using X-ray machines. Before entering the secure area of an airport, all passengers, as well as members of airline and airport staff have to pass the security checkpoints to have themselves and all their belongings screened. The security checkpoint is a socio-technical system consisting of human and technical elements, working together. The goal is that no threat items are brought past security checkpoints and onto an airplane. Strong efforts are being made in order to improve and further develop X-ray screening equipment. Yet, the final decision whether threat items are contained in the baggage still relies on human operators (screening officers) who visually inspect the X-ray images provided by the machine. As a consequence, man-machine system performance depends on human factors and display technology (e.g., Bolfig et al., 2008; Koller et al., 2008; von Bastian et al., 2008, 2010; Michel and Schwaninger, 2009; Graves et al., 2011). When evaluating new technological developments with regard to their added value for security screening purposes, this should be taken into account appropriately (see also Yoo and Choi, 2006; Yoo, 2009).

In X-ray screening, three image-based factors have been identified as relevant for human operators to detect threat items in X-ray images (Schwaninger, 2003b; Hardmeier et al., 2005, 2006; Schwaninger et al., 2005a). The first one is the view difficulty of an object, resulting from the position of a threat item in a bag (effect of viewpoint). The second factor is the superposition of an item by other objects contained in the bag (effect of superposition). The third factor refers to the complexity of a bag, which depends on the number and type of objects in the bag (effect of bag complexity). The intensity with which X-rays can penetrate through materials in a bag depends on the specific material density of a substance (e.g., Brown et al., 1995). Therefore, the material density of the items contained in a bag will also affect the factors superposition and bag complexity and thus will influence the difficulty to detect threat items. Schwaninger et al. (2005b) have developed algorithms to automatically estimate X-ray image difficulty based on viewpoint, superposition, and bag complexity. Their algorithms were highly correlated with human perception of the above mentioned image-based factors and could well predict human threat detection performance (see also Schwaninger et al., 2007; Bolfig et al., 2008).

State-of-the-art X-ray screening equipment is able to provide high quality images with good image resolution. Yet, the detection of threat items in X-ray images remains a challenging task

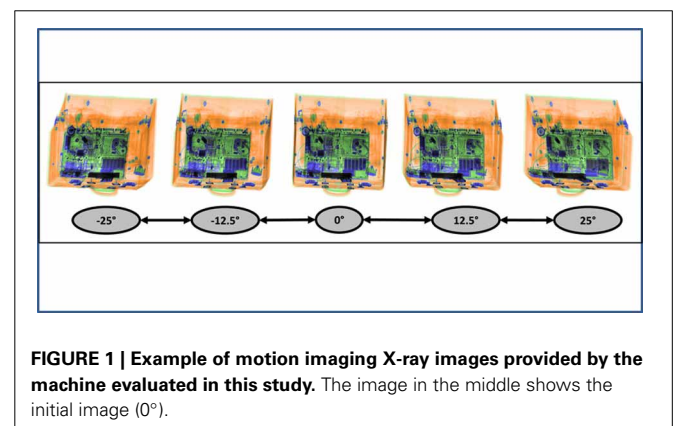
for screening officers and becomes even more difficult when dense objects, such as large electronic devices, are contained in the baggage. Due to their compact construction, electronic devices (e.g., laptops) are hard to penetrate. Hence, they can conceal other parts of luggage or could be used to intentionally hide threat items (e.g., an improvised explosive device, IED). Especially when single view X-ray systems are used or even multi-view systems, if the additional views do not provide enough meaningful information, the inspection becomes difficult. Threat items which are behind, in front of, or hidden inside a laptop case become very challenging or even impossible for human operators to recognize (see also von Bastian et al., 2008). In a previous paper, Mendes et al. (2012) documented how threat detection can be substantially impaired when laptops are not taken out of passenger bags and a threat item (e.g., an IED) is placed either behind, in front of, or within a laptop. The present paper extends these results by investigating how a new technology which allows presenting bags in multiple views as an image sequence (i.e., motion imaging) could possibly reduce such an impairment.

Considering the large number of views which can be produced by a single object, the question arises how objects can be recognized when presented in unusual views. In the object recognition literature, two types of theories can be distinguished (see Peissig and Tarr, 2007; Kravitz et al., 2008): viewpoint-invariant theories (e.g., Marr, 1982; Biederman, 1987) and viewpoint-dependent theories (e.g., Poggio and Edelman, 1990; Bülthoff and Edelman, 1992; Tarr, 1995). Most viewpoint-invariant theories assume that objects are stored in visual memory by their component parts and their spatial relationship (see Marr and Nishihara, 1978; Biederman, 1987). Once a particular object has been stored, recognition of that object should be unaffected by the viewpoint (including novel viewpoints), given that the necessary features can be recovered from this view (Burgund and Marsolek, 2000). The viewpoint-dependent theories propose that objects are not stored in memory as rotation invariant structural descriptions, but in a viewer centered format. Thus, if an object has never been seen from a certain viewpoint and is therefore not stored in visual memory, recognition is impaired if view-invariant features are not available (Kosslyn, 1994; Bülthoff and Bülthoff, 2006; Schwaninger, 2005). Several studies on viewpoint-dependent theories could show that viewpoint can strongly affect recognition performance (e.g., Bülthoff and Edelman, 1992; Edelman and Bülthoff, 1992; Humphrey and Khan, 1992; Graf et al., 2002). Even though our visual perception can be considered highly robust with respect to changes of viewpoint, we are more facile with certain views relative to others, such as often encountered views and views that make larger numbers of surfaces available (Palmer et al., 1981; Blanz et al., 1999). Such views have also been referred to as “canonical” views. Research in aviation security X-ray screening has shown that threat items are easier to identify when depicted in frontal (canonical) views than when horizontally or vertically rotated (e.g., Michel et al., 2007; Bolfiging et al., 2008; Koller et al., 2008). Consequently, having machines featuring multiple X-ray images of the same bag from different viewpoints could ease recognition of threat items in passenger bags for screening officers.

At present, most of the machines deployed at airports provide single view images, which do not allow screening officers to analyze an image from different viewpoints. A human operator will only be able to identify a threat item and make a correct decision if the threat can be recognized in the provided single view image (Schwaninger, 2003b; Schwaninger et al., 2005a; Graves et al., 2011). Considering the above mentioned image based factors (viewpoint, superposition and bag complexity) and the density of electronic devices, it becomes evident why most international and national regulations specify that portable computers and other large electronic devices shall be removed from passenger bags and screened separately at security checkpoints (e.g., the current regulation of the European Commission, 2010). Based on the model by Schwaninger et al. (2005b) one would predict that leaving laptops in passenger bags results in decreases of threat detection performance due to increases of superposition and bag complexity. Threat items placed behind, in front of, or inside a laptop could become very challenging for human operators to detect. Moreover, recognition would become additionally challenging if in the provided X-ray image the threat item would be depicted from a difficult viewpoint (e.g., vertically or horizontally rotated).

This study was conducted to examine the above mentioned effects by comparing conventional single view display technology to a new technology. More specifically, a new X-ray screening machine featuring “motion imaging” was tested. “Motion imaging” means that five images are available, which are rotated around the vertical axis. These can be either displayed in a short video sequence or can each be statically viewed. In relation to the initial image (0°), the angles of the five images are -25° , -12.5° , 0° , 12.5° , 25° (see Figure 1).

One could hypothesize that through the application of motion imaging and the availability of multiple views, recognition of certain objects could become easier. There are several possible advantages dynamic displays may confer over static ones (Vuong and Tarr, 2004). For example, object motion may enhance the recovery of information about shape (e.g., Ullmann, 1979). Furthermore, it may provide observers with additional views of objects (Pike et al., 1997), or it may allow observers to anticipate views of objects (Mitsumatsu and Yokosawa, 2003). Moreover, when objects rotate in depth, certain features can become visible while others become obscured (Vuong and Tarr, 2004). Thus,



objects could become less superimposed and could possibly be displayed from an easier viewpoint (i.e., from a more canonical perspective).

The first goal of our study was to determine whether motion imaging improves detection of threat items in passenger bags. The second goal was to investigate whether leaving laptops in passenger bags results in a decrease of detection performance (effect of superposition and bag complexity), while the third goal was to evaluate whether such an effect can be compensated when motion imaging is available. Additional analyses were carried out to examine effects depending on different threat categories (guns, IEDs, knives, others), the placement of the threat items (in bag vs. in laptop) and the viewpoint effect (easy vs. difficult view).

METHODS AND MATERIALS

An image interpretation test containing bags and laptops was created in four versions to examine the factors display condition (single vs. motion imaging) and packing condition (laptops inside vs. laptops outside). Each test version differed with regard to these two factors (see section Experimental Design). Four experimental groups with certified screening officers were formed. Each group conducted one of the test versions. Detection performance scores and reaction times (RTs) of all groups were compared to evaluate the effects of the above mentioned factors.

PARTICIPANTS

The study was conducted with 80 airport security screening officers employed at an international European airport. All participants were certified screeners, meaning they were all qualified, trained and certified according to the standards set by the national appropriate authority (civil aviation administration) and consistent with the European Regulation (European Commission, 2010). The screening officers were randomly distributed into four different experimental groups (A, B, C, and D, 20 per group). **Figure 2** illustrates the experimental design. In order to verify that all experimental groups were comparable with regard to the screeners' X-ray image interpretation competency, all participants conducted the X-Ray Competency Assessment Test (X-Ray CAT) before the main experiment was carried out. The X-Ray

CAT for cabin baggage screening is a standardized instrument to measure X-ray image interpretation competency of airport security screening officers and has been applied in several previous scientific studies (Koller and Schwaninger, 2006; Michel et al., 2007; Koller et al., 2008). It is currently used for screener certification at several European airports. The test consists of 256 trials and is based on 128 different color X-ray images of passenger bags, which are each used twice: once without (non-threat image) and once containing a threat object (threat image). For more information on the X-Ray CAT see Koller and Schwaninger (2006). Average detection performance scores (A')¹ of all four groups were compared using *post-hoc* pairwise comparisons with Bonferroni correction. No significant differences between the groups could be found (all p values >0.05), implying that they were comparable regarding their image interpretation competency. The average age of the participants was $M = 40.69$ years ($SD = 10.78$), with a range between 22 and 58 years. 53% of the participants were female. The average amount of job experience was $M = 4.95$ years ($SD = 4.49$, range: 0.5–23 years). Between-participants analyses of variance showed no differences between the experimental groups with regard to age [$F_{(3, 76)} = 1.57$, $p = 0.204$, $\eta^2 = 0.058$] or job experience [$F_{(3, 74)} = 0.66$, $p = 0.579$, $\eta^2 = 0.026$].

EXPERIMENTAL DESIGN

All experimental groups conducted a computer-based X-ray image interpretation test. During the test, color X-ray images of passenger bags and laptops were displayed, sometimes containing threats (threat images) and sometimes without any threat items (non-threat images). Images were displayed in random order. All participants were exposed to every image and had to decide whether the bags and laptops could be regarded as harmless (OK) or whether they contained a threat item (NOT OK). Each test condition differed with regard to the factors display condition (single view vs. motion imaging) and packing condition (laptops inside vs. outside of passenger bags). **Figure 2** displays the experimental design of the study. The following four

¹For details on the calculation of A' see section Results and Discussion.

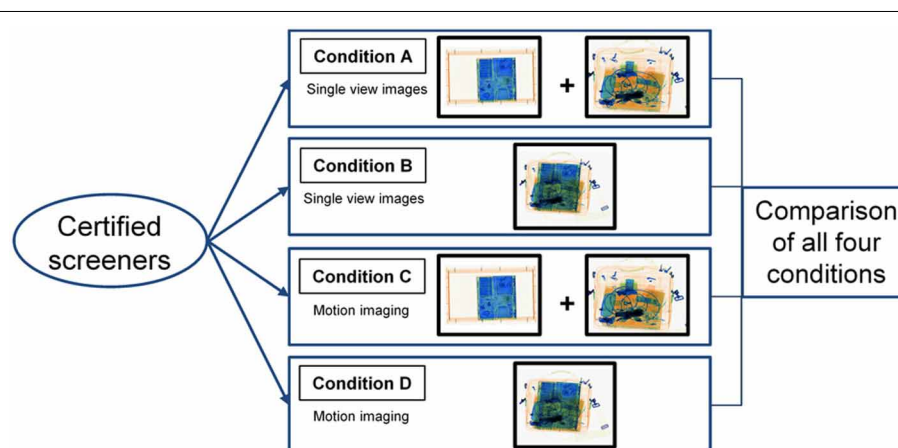


FIGURE 2 | Experimental design (between-participants) for the comparison of the four different conditions.

different experimental conditions were conducted and compared to examine the effects and interactions of the above mentioned two factors using a between-participants design:

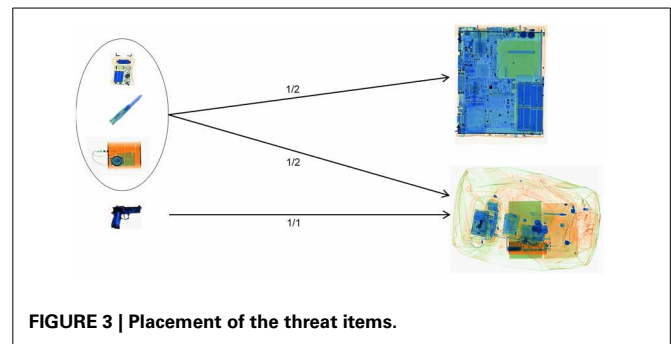
- (A) Single view images, laptops and passenger bags screened separately.
- (B) Single view images, laptops are left inside the passenger bags.
- (C) Motion imaging, laptops and passenger bags screened separately.
- (D) Motion imaging, laptops are left inside the passenger bags.

In all test conditions the same bags were presented to the screening officers. Originally, every bag contained a laptop. In conditions A and C the laptops were taken out of the bag and screened separately, whereas in conditions B and D the laptops were left inside the passenger bags. This allowed examining the effects of superposition and bag complexity caused by laptops. **Figure 2** illustrates the two different packing conditions (laptop inside vs. laptop outside).

In conditions A and B, images of the baggage and laptop could only be seen from one single viewpoint. Conditions C and D allowed examining the images from different viewpoints through motion imaging. As explained in the introduction, one important objective of this study was to test whether motion imaging could enhance the inspection of passenger bags to such an extent that laptops could be left inside passenger bags without affecting detection performance negatively. Would detection performance scores still be significantly higher in condition A compared to D, one could conclude that the detection of threat items is significantly impaired when laptops are left inside passenger bags, even when motion imaging is available.

IMAGE INTERPRETATION TEST

The image interpretation test was based on a representative set² of 96 passenger bags (defined by screening experts from a specialized police organization), all of which originally contained laptops. All test images were recorded with the machine evaluated in this study. The test images were created and recorded in collaboration with aviation security experts from a specialized police organization and former airport security screening officers now employed by CASRA. As explained above, in conditions A and C the laptops were taken out of the bags and recorded separately, whereas in conditions B and D the laptops were left inside the bags. Each bag/laptop-combination was used twice, once containing a threat item in either the bag or the laptop, and once without any threat item. The test contained a representative sample of threat items selected and developed (the IEDs in laptops) by experts from an airport police department. These could be divided into four different threat categories: guns, IEDs, knives and other threat items (e.g., electric shock devices, etc.). For all categories except guns, in half of the cases the threat items were placed in the bag, while in the other half of the cases the threat items were placed within the laptop (see **Figure 3**). Due to their size, it would not have been



realistic to place guns inside a laptop. Moreover, the factor viewpoint was included in the test design. For those threat items placed inside the bags, half were positioned in easy views and half in difficult views. Easy view means that threat items were depicted from a frontal/canonical view in the X-ray image, while for difficult view the threat items were horizontally or vertically rotated. All the threat items placed inside the laptop cases were positioned in easy views. As laptops are comparably flat, it would have been difficult to place threat items in vertically or horizontally rotated positions. The IEDs which were placed inside the laptops were specifically built into the cases. It must be considered that since an IED consists of several component parts, it becomes more difficult to determine what the canonical view and thus an easy view would be. Each threat category contained 24 items. Therefore, the number of test images for the conditions were the following:

- **Tests A and C (laptops and bags screened separately):**
 4×24 threat images (60 bags and 36 laptops)
 +96 non threat BAG images
 +96 non threat LAPTOP images
 = 288 test images
- **Tests B and D (laptops inside bags):**
 4×24 combined threat images (60 threats in bags and 36 threats in laptops)
 +96 combined non threat images
 = 192 test images

PROCEDURE

All participants were invited to the experimenters' facilities to conduct the test. Four computer workstations with the corresponding consoles of the tested machine and 19" TFT monitors were set up in a normally lit room. X-ray images covered about 2/3 of the computer screen. The distance to the monitor was ~60 cm. Four participants at a time were tested. Before the test started, all participants received a short introduction by the test supervisor, explaining the test procedure and introducing the new technology of motion imaging. All participants were able to try out the console and view test images for ~20 min, in order to become familiar with the images, the technology and the handling of the console. Pre-testing had shown that this amount of time was enough to get well acquainted with the console and it was also recommended by the manufacturer. After a break of 10 min the actual test started. Tests were conducted quietly and individually, and under supervision. The test images remained on the screen

²The set was based on a two-month data collection, assessing the contents and types of bags that were passing through the security checkpoints at an international European airport.

until the participant either pressed the “OK” or “NOT OK” and the “move belt forward” button. RTs were measured in milliseconds and correspond to the amount of time it took for a screening officer to come to a decision and press the “OK” or “NOT OK” button after the first image pixel of the bag/laptop appeared on the screen. There was no time limit set for viewing an image. However, participants were instructed to inspect the images as quickly and accurately as possible. Breaks of 10 min were taken in 30 minute-cycles, to avoid eyestrain and fatigue, and to make sure that especially those participants conducting tests A and C (288 images instead of 192 images, see section Image Interpretation Test) would not become too tired toward the end. All participants completed the test in less than 2 h, including breaks.

RESULTS AND DISCUSSION

According to signal detection theory (Green and Swets, 1966), there are four possible outcomes to a screener’s response when judging an X-ray image as either OK or NOT OK: hit, false-alarm, correct rejection and miss (Schwaninger, 2003a; Hofer and Schwaninger, 2004). In this study, A' was applied as a measure for detection performance (Pollack and Norman, 1964). A' is a measure of sensitivity which is commonly used for a variety of tasks including screener certification and competency assessments (Hofer and Schwaninger, 2004; Koller and Schwaninger, 2006; Michel et al., 2010). It considers the hit rate as well as the false-alarm rate and can be calculated using the following formula (Grier, 1971):

$$0.5 + [(H - F)(1 + H - F)]/[4H(1 - F)] \quad (1)$$

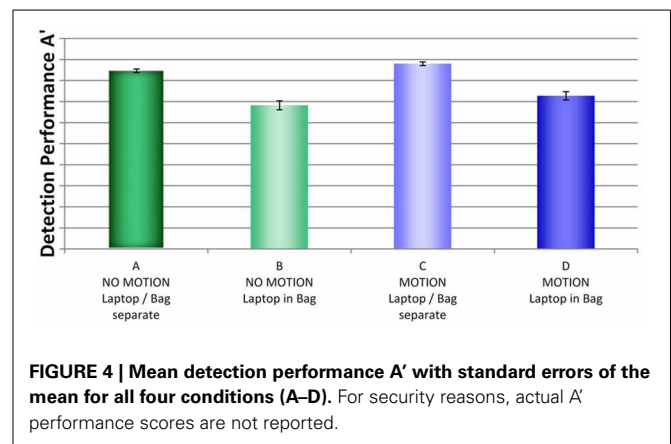
$$0.5 + [(F - H)(1 + F - H)]/[4F(1 - H)] \quad (2)$$

H is the hit rate and F the false alarm rate. If performance is below chance, i.e., when $H < F$, equation (2) must be used (Aaronson and Watts, 1987).

Due to the security confidential nature of performance values, these are not displayed in this paper. In order to provide meaningful results, relative differences and effect sizes are reported. All reported effect sizes are interpreted based on Cohen (1988). For t -tests, d between 0.20 and 0.49 represents a small effect size; d between 0.50 and 0.79 represents a medium effect size; $d \geq 0.80$ represent a large effect size. For analysis of variance (ANOVA) statistics, η^2 between 0.01 and 0.05 represents a small effect size; η^2 between 0.06 and 0.13 represents a medium effect size; $\eta^2 \geq 0.14$ represents a large effect size.

COMPARISON OF DETECTION PERFORMANCE BY CONDITION

Figure 4 shows a comparison of detection performance scores by condition (A, B, C, and D)³. Most remarkable seems to be



the effect of packing condition. Performance was much better in conditions A and C, where laptops and bags were screened separately, compared to conditions B and D, where laptops were left inside the passenger bags. The graph also suggests that performance was slightly better when motion imaging was available (condition C compared to A and condition D compared to C, respectively). The ANOVA with the between-participants factors display condition (no motion vs. motion) and packing condition (laptop separate vs. laptop in bag) revealed a large main effect for packing condition, $F_{(1, 76)} = 105.22$, $p < 0.001$, $\eta^2 = 0.581$, and a medium main effect for display condition, $F_{(1, 76)} = 5.05$, $p < 0.05$, $\eta^2 = 0.062$. There was no interaction between display and packing condition, $F_{(1, 76)} = 0.361$, $p = 0.55$, $\eta^2 = 0.005$. Thus, although motion imaging enhanced detection performance slightly, it could not compensate the negative effects on detection performance resulting from leaving laptops inside bags. Further, the direct comparison of condition D (motion imaging available, laptops in bags) and condition A (no motion imaging available, laptops and bags screened separately) revealed a highly significant effect, $t_{(26)} = 5.89$, $p < 0.001$, $d = 1.86$. This also shows that although motion imaging did improve detection performance (as shown by the main effect in the ANOVA), the large negative effect of packing condition could not be compensated.

In sum, the results imply that the packing condition had a high impact on detection performance. Motion imaging resulted in better detection but could not fully compensate the effect of packing condition (i.e., impaired detection when leaving laptops in passenger bags). The large main effect for packing condition is consistent with the assumption that the well-documented effects of superposition and bag complexity (Schwaninger et al., 2005a,b, 2007; Hardmeier et al., 2005, 2006; Bolting et al., 2008; von Bastian et al., 2008) increase when laptops are left in passenger bags, resulting in impairments of threat detection performance.

A more detailed analysis was conducted by looking at each threat category separately. As can be seen in Figure 5, large differences between conditions, but also between threat categories were

³Due to the packing condition, the proportion of target present and target absent trials differed for conditions A/C and B/D (in conditions A and C the ratio is 1:2; in conditions B and D the ratio is 1:1). According to signal detection theory (Green and Swets, 1966), different ratios of target present and target absent trials can result in a criterion shift (i.e., changes in hit and false alarm rates). Measures of detection performance in terms of sensitivity such as d' and A' are thought to be relatively independent of criterion shifts, which could also be shown in studies on target prevalence (e.g., Gur et al., 2003;

Wolfe et al., 2007; Wolfe and Van Wert, 2010). Therefore, it can be assumed that the different proportions on target present trials in conditions A/C and B/D did not affect detection performance (A') results in this study.

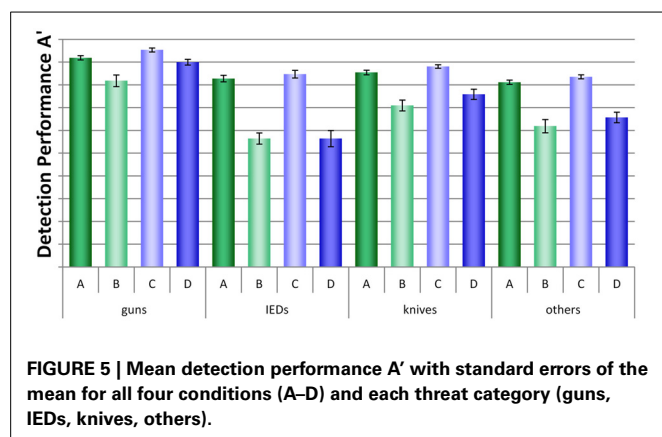


FIGURE 5 | Mean detection performance A' with standard errors of the mean for all four conditions (A–D) and each threat category (guns, IEDs, knives, others).

Table 1 | Results of the ANOVAs conducted with detection performance (A') as dependent variable⁴.

Detection performance	df	F	η^2	p-value
(A)				
Threat category (T)	2.04, 155	186.96	0.711	<0.001
Display condition (D)	1, 76	4.49	0.056	<0.05
Packing condition (P)	1, 76	108.12	0.587	<0.001
$T \times D$	2.04, 155	2.43	0.031	0.09
$T \times P$	2.04, 155	42.21	0.357	<0.001
$D \times P$	1, 76	0.25	0.003	0.62
$T \times D \times P$	2.04, 155	1.15	0.015	0.320
(B)				
View difficulty (V)	1, 76	14.23	0.158	<0.001
Threat category (T)	2.23, 170	122.89	0.618	<0.001
Condition (C)	3, 76	28.78	0.532	<0.001
$V \times C$	3, 76	0.33	0.013	0.805
$T \times C$	6.70, 170	5.07	0.167	<0.001
$V \times T$	2.38, 181	52.03	0.406	<0.001
$V \times T \times C$	7.12, 181	2.66	0.095	<0.05

found. A mixed-design ANOVA with the within-participants factor threat category (guns, IEDs, knives, others) and the between-participants factors display condition (no motion vs. motion) and packing condition (laptop separate vs. laptop in bag) revealed large significant main effects for the factors threat category and packing condition and a medium effect for display condition (for details, see **Table 1A**). The interaction between threat category and packing condition was also highly significant, implying that leaving laptops in passenger bags affected performance differently, depending on threat category. None of the other interactions reached statistical significance. As **Figure 5** indicates, IEDs and other threats were most difficult to detect, especially in conditions B and D. In general, a slight advantage of motion imaging could be observed (compare condition C to A, and D to B), which according to **Figure 5** was most evident for guns.

⁴In all analyses of variance in this study were Mauchly's test indicated that the assumption of sphericity had been violated, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity.

Table 2 | Results of the two-tailed independent samples t -tests comparing detection performance A' between conditions A and D for each threat category (guns, knives, IEDs, others)⁵.

	$t_{(35)}$	$t_{(25)}$	$t_{(27)}$	p	d
A–D (guns)	1.22			0.23	0.39
A–D (IEDs)		6.90		<0.001	2.37
A–D (knives)			3.85	<0.01	1.30
A–D (others)		6.13		<0.001	2.10

Additionally, we conducted direct comparisons between conditions D (motion imaging available, laptops in bags) and A (no motion imaging available, laptops and bags screened separately) for each threat category, to further examine whether for certain threat types the negative effect on detection performance of leaving laptops in bags could be fully compensated by motion imaging. For all threat categories except guns, large significant differences were revealed (see **Table 2**). This further indicates that even though motion imaging did improve detection performance (as shown by the main effect in the ANOVA, see above), it could not compensate the large negative effect of packing condition. Only for the detection of guns, motion imaging seemed to have helped to compensate the negative effect of leaving laptops in bags (which could explain the marginally significant interaction ($p = 0.09$) between threat category and display condition in **Table 1A**).

As described earlier, half of the threat items were placed inside laptops and half were placed inside the bags (except for guns, which could not be placed inside laptops). **Figure 6** displays how detection performance differed for each condition with regard to threat category and the placement of threat items. Again, threat items were detected better when the bags and laptops were screened separately (conditions A and C). Planned comparisons were conducted for each condition and threat category (except for guns, as all guns were placed inside the bags), to compare the differences between detection performance with regard to the placement of threats for each condition (see **Table 3**). Biggest differences were found for IEDs. For each condition, detection performance was worse when the IEDs were built into the laptops, compared to when they were placed inside the bags. However, while for conditions A and C detection performance was still relatively high, the scores achieved in conditions B and D were much lower for the IEDs within the laptops. For the threat categories knives and others, in most conditions detection performance was higher when these were placed inside the laptops. This could be explained by the fact that all threat items placed within the laptops were positioned in easy views (see Method section), and thus were easier to recognize. For IEDs, this effect was not observed. As the IEDs were specifically built into the laptops and since an IED consists of several component parts, it becomes more difficult to determine what actually the canonical/frontal view and thus an easy view would be (see Method section).

⁵In all t -tests of this study where Levene's test indicated unequal variances, degrees of freedom were adjusted using the default procedure in SPSS.

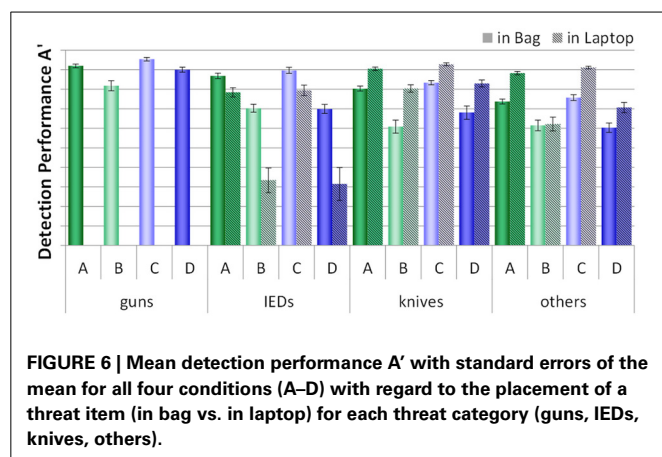


FIGURE 6 | Mean detection performance A' with standard errors of the mean for all four conditions (A–D) with regard to the placement of a threat item (in bag vs. in laptop) for each threat category (guns, IEDs, knives, others).

Table 3 | Results of two-tailed paired samples *t*-tests comparing the detection performance A' with regard to the placement of each threat item (in laptop vs. in bag) for each threat category (guns, IEDs, knives, others) and each condition (A–D).

	<i>t</i> ₍₁₉₎	<i>p</i>	<i>d</i>
IEDs			
A	3.30	<0.01	0.74
B	5.98	<0.001	1.34
C	3.76	<0.01	0.84
D	4.83	<0.001	1.08
KNIVES			
A	–12.68	<0.001	–2.84
B	–7.76	<0.001	–1.73
C	–9.07	<0.001	–2.03
D	–5.32	<0.001	–1.19
OTHERS			
A	–14.24	<0.001	–3.18
B	–0.31	0.76	–0.07
C	–11.23	<0.001	–2.51
D	–5.45	<0.001	–1.22

In order to examine the viewpoint effect and whether this effect was influenced by condition, detection performance scores of all conditions were compared, broken up by easy vs. difficult view (see Figure 7). Since all threat items placed inside the laptop cases were positioned in easy views, this analysis was only conducted for the threat items placed inside the bags. A mixed-design ANOVA with the within-participants factors view difficulty (easy vs. difficult view) and threat category (guns, IEDs, knives, others) and the between-participants factor condition (A, B, C, D) revealed large significant main effects for all three factors (see Table 1B). There was no significant interaction between view difficulty and condition, while all other interactions were significant. Therefore, a viewpoint effect could clearly be observed, which differed with regard to threat category. However, view difficulty was not significantly affected by condition. Interestingly, as Figure 7 indicates, throughout all conditions guns, IEDs and knives were detected better when depicted in easy views, while for the category others this was the other way around. The category others

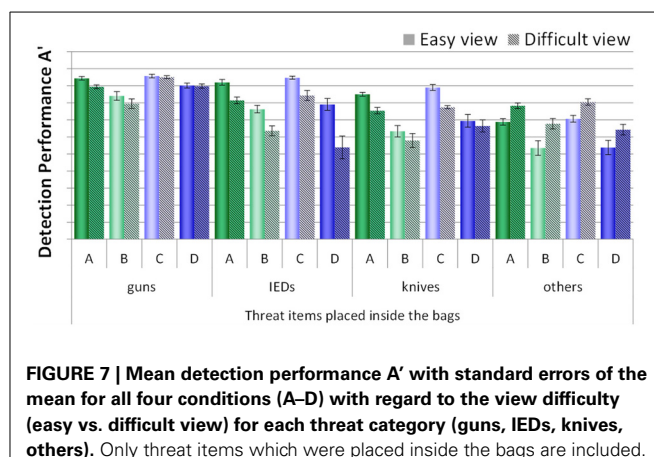


FIGURE 7 | Mean detection performance A' with standard errors of the mean for all four conditions (A–D) with regard to the view difficulty (easy vs. difficult view) for each threat category (guns, IEDs, knives, others). Only threat items which were placed inside the bags are included.

contained a very heterogeneous group of threat items (e.g., pepper spray, taser, throwing star, etc.). Hence, it could have been that the screening officers were more familiar with those threat items positioned in difficult views, and therefore recognized these more easily. As Figure 7 further indicates, for the category guns, motion imaging seemed to have been of help to reduce the view-point effect (see conditions C and D). This is consistent with the results reported above (see Table 2) and makes sense if one takes into account that guns change their shape more drastically than other objects when rotated. Thus, motion imaging can be more effective for supporting the recognition of guns.

COMPARISON OF REACTION TIMES BY CONDITION

Figure 8 shows the average reaction times (RTs, converted into seconds) for all conditions and threat categories. For all categories, a similar pattern can be observed: More time was needed in conditions B and D where laptops were left inside the bags. Most time was needed in condition D, where motion imaging was available. As Figure 8 implies, remarkable differences can be observed between the threat categories and conditions. A repeated-measures ANOVA (see Table 4) revealed large significant main effects for the factors threat category (guns, IEDs, knives, others) and condition (A, B, C, D). The interaction between both factors was also significant, implying that the size of the differences in RTs between the conditions varied with regard to threat category. As displayed in Figure 8, all conditions achieved fastest RTs for the category guns, while longest RTs are clearly observed for the category IEDs.

To determine which condition actually took the longest time to complete the test all RTs for each security screener in each condition were summed and averaged across screening officers. Figure 9 displays these results. As described in the Method section, for test conditions A and C where laptops and bags were displayed separately, 288 images were displayed. In test conditions B and D, 192 images were shown. Even though fewer images were viewed in condition D, compared to conditions A and C, altogether, more time was needed to inspect these test images. While conditions A, B, and C did not differ from each other significantly, large differences were observed between each of these three conditions with condition D (see Table 5). These results indicate that

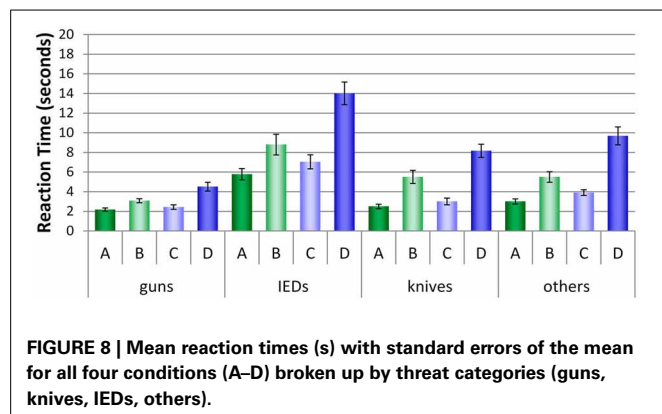


FIGURE 8 | Mean reaction times (s) with standard errors of the mean for all four conditions (A–D) broken up by threat categories (guns, knives, IEDs, others).

Table 4 | Results of the repeated-measures ANOVA conducted with reaction time (RT).

Reaction time	df	F	η^2	p-value
Threat category (T)	1.89, 144	155.04	0.671	<0.001
Condition (C)	3, 76	25.26	0.499	<0.001
T × C	5.67, 144	8.11	0.242	<0.001

even though fewer images had to be viewed when laptops were kept in passenger bags, altogether more time was needed to apply motion imaging and investigate these images thoroughly. Thus, while motion imaging provides a security advantage, it comes with a certain cost of efficiency.

SUMMARY AND CONCLUSIONS

The benefits of an X-ray machine featuring a new technology offering multiple views of X-ray images and motion imaging were evaluated and compared to single view imaging. In specific, it was investigated whether leaving laptops inside passenger bags resulted in a decrease of detection performance and whether such an effect could be compensated by motion imaging. The results revealed that threat detection performance was much better when laptops and bags were screened separately (see also Mendes et al., 2012). Leaving laptops inside passenger bags resulted in a clear decrease of threat detection performance, supporting the view that increases in superposition and bag complexity affect detection performance negatively (Schwaninger et al., 2005b, 2007; Bolting et al., 2008). Motion imaging technology could slightly improve threat detection performance. Yet, it could not compensate the negative effect of leaving laptops inside bags. Highest detection performance was achieved when motion imaging was available and laptops and bags were screened separately.

More detailed analyses indicate that performance differed remarkably with regard to the different threat categories [guns, improvised explosive devices (IEDs), knives, others]. IEDs and the others threat category were most difficult to detect, especially when laptops were not removed from passenger bags. Only a small advantage of motion imaging was observed. Merely for the detection of guns, motion imaging seemed to be of substantial benefit. Further analyses regarding the placement of threat items (in bag vs. in laptop) indicated that IEDs were particularly

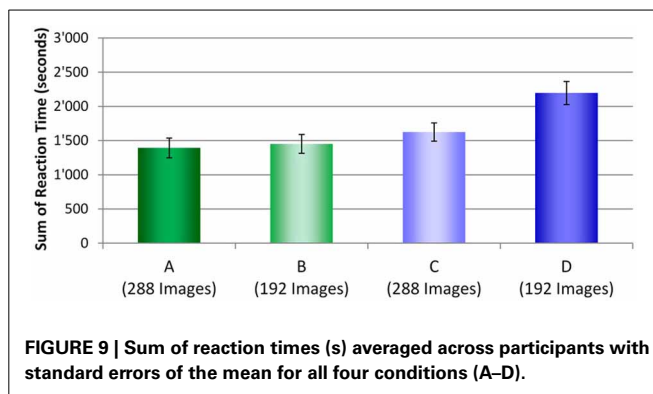


FIGURE 9 | Sum of reaction times (s) averaged across participants with standard errors of the mean for all four conditions (A–D).

Table 5 | Results of pairwise comparisons with Bonferroni correction for the sums of reaction times (in seconds) of all four conditions (SPSS Bonferroni adjusted *p*-values are quoted).

Comparison	Mean difference RT	SE	<i>p</i>	<i>d</i>
A–B	–58	207.5	1.000	–0.09
A–C	–231	207.5	1.000	–0.32
A–D	–802	207.5	<0.01	–1.14
B–C	–173	207.5	1.000	–0.29
B–D	–744	207.5	<0.01	–1.08
C–D	–570	207.5	<0.05	–0.84

difficult to detect when these were built into the laptop cases. Specifically when laptops were left inside the bags, threat detection performance was quite low compared to when the laptops were displayed separately. Thus, when no automatic explosives detection is available and laptops are not removed from passenger bags, the detection of explosives and bombs, in particular, is impaired. For the categories knives and others, detection performance was higher when these were placed inside the laptops. This could be due to the fact that—for practical reasons—all threat items placed inside the laptops had to be positioned in easy views (canonical views). In general, threat items depicted in more difficult views were harder to detect. These findings are consistent with previous research on viewpoint effects, which showed that recognition of items depicted in frontal/canonical view is easier (e.g., Michel et al., 2007; Bolting et al., 2008; Koller et al., 2008). Only for the category others, this effect was the other way round. As the category others contained a very heterogeneous group of threat items, possibly screening officers were more familiar with the items positioned in difficult views and thus detected these better. Results also showed that in general more time was needed to inspect the images when laptops were left inside the bags. Longest RTs were found when laptops were not removed from bags and motion imaging was applied. Thus, providing additional views is paid for by increasing RT (see also von Bastian et al., 2008). Even though fewer images were viewed when laptops were left inside the passenger bags, altogether more time was needed to apply motion imaging and inspect these images properly. Keeping factors such as throughput and efficiency at security checkpoints in mind, screening time is an important point to consider.

Technology for security screening will constantly be developed further. Yet, the final decision on whether threat items are contained in luggage still rely on human operators, who inspect the luggage based on an image provided by a machine. The presented study underlines the importance of thoroughly evaluating any new technological features with regard to their added value provided to the screening officers, prior to implementing these in the airport environment. In this study, only a slight benefit of motion imaging technology was revealed. No real advantage could be observed for the detection of IEDs, while the results do suggest that for certain objects such as guns, the rotation and availability of different viewpoints through motion imaging could improve identification. As previous research has shown (e.g., Michel et al., 2007) guns change their shape more drastically than other objects when rotated. Thus, one could assume that motion imaging would possibly be more helpful also for the detection of other threat types if larger rotations and more views are available (or even fully rotatable 3D images, see below).

All in all, the detection of threat items in cabin baggage screening currently still seems more reliable when laptops are taken out of passenger bags. Therefore, the outcomes of this study underline the appropriateness and importance of current regulations

specifying that portable computers should be removed from passenger bags for X-ray screening. However, this might be reconsidered if effective and efficient automatic threat detection is available, which is particularly important for IEDs (see e.g., Singh and Singh, 2003; Eilbert, 2009; Mery et al., 2013). Furthermore, if more rotation in depth would be available, higher benefits could possibly be expected, which is of particular importance regarding new technological developments such as computer tomography offering 3D views. Effects of superposition and viewpoint could be reduced further and RTs could be decreased if screening officers can directly navigate to their preferred view of a bag image. In combination with automated threat detection this could possibly result in substantially higher human-machine system performance (see e.g., Flitton et al., 2010; Megherbi et al., 2010). However, this would have to be examined in further studies.

ACKNOWLEDGMENTS

We are thankful to aviation security experts from the German Federal Police Technology Center for providing their valuable expertise and support for the creation and recording of x-ray images. We thank Zurich State Police, Airport Division, for providing screeners and supporting this study.

REFERENCES

- Aaronson, D., and Watts, B. (1987). Extension of Grier's computational formulas for A' and B' to below-chance performance. *Psychol. Bull.* 102, 439–442. doi: 10.1037/0033-2909.102.3.439
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147. doi: 10.1037/0033-295X.94.2.115
- Blanz, V., Tarr, M. J., Bülthoff, H. H., and Vetter, T. (1999). What object attributes determine canonical views? *Perception* 28, 575–600. doi: 10.1068/p2897
- Bolfing, A., Halbherr, T., and Schwaninger, A. (2008). "How image based factors and human factors contribute to threat detection performance in X-ray aviation security screening," in *HCI and Usability for Education and Work Lecture Notes in Computer Science*, Vol. 5298, ed A. Holzinger (Berlin; Heidelberg: Springer-Verlag), 419–438. doi: 10.1007/978-3-540-89350-9_30
- Brown, W., Emery, T., Gergory, M., Hackett, R., Yates, C., and Ling, J. (1995). *Advanced Physics: Practical Guide*. Harlow: Longman.
- Bülthoff, H. H., and Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc. Natl. Acad. Sci. U.S.A.* 89, 60–64. doi: 10.1073/pnas.89.1.60
- Bülthoff, I., and Bülthoff, H. H. (2006). "Objektwahrnehmung," in *Handbuch der Allgemeinen Psychologie – Kognition*, eds J. Funke and P. A. Frensch (Göttingen: Hogrefe Verlag), 165–172.
- Burgund, E. D., and Marsolek, C. J. (2000). Viewpoint-invariant and viewpoint-dependent object recognition in dissociable neural subsystems. *Psychon. Bull. Rev.* 7, 480–489. doi: 10.3758/BF03214360
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Erlbaum, Hillsdale.
- Edelman, S., and Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Res.* 32, 2385–2400. doi: 10.1016/0042-6989(92)90102-O
- Eilbert, R. F. (2009). "Chapter 6 – X-ray technologies," in *Aspects of Explosives Detection*, eds M. M. Marshall and J. C. Oxley (Oxford: Elsevier), 89–130. doi: 10.1016/B978-0-12-374533-0.00006-4
- European Commission. (2010). Commission regulation (EU) No 185/2010 of March 2010 laying down detailed measures for the implementation of the common basic standards on aviation security. *Official J. Eur. Union* 53, 1–55.
- Flitton, G. T., Breckon, T. P., and Bouallagu, N. M. (2010). "Object recognition using 3D SIFT in complex CT volumes," in *Proceedings of the British Machine Vision Conference*, eds F. Labrosse, R. Zwigelaar, Y. Liu, and B. Tiddeman (BMVA Press), 1–12. doi: 10.5244/C.24.11
- Graf, M., Schwaninger, A., Wallraven, C., and Bülthoff, H. H. (2002). *Psychophysical Results from Experiments on Recognition and Categorisation. Information Society Technologies (IST) Programme, Cognitive Vision Systems CogVis (IST-2000-29375)*, Tübingen: Max Planck Institute for Biological Cybernetics.
- Graves, I., Butavicius, M., MacLeod, V., Heyer, R., Parsons, K., Kuester, N., et al. (2011). "The role of the human operator in image-based airport security technologies," in *Innovations in Defence Support Systems – Socio-Technical Systems, Studies in Computational Intelligence*, Vol. 338, eds L. C. Jain, E. V. Aidman, and C. Abeynake (Berlin; Heidelberg: Springer Verlag), 147–181. doi: 10.1007/978-3-642-17764-4_5
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: computing formulas. *Psychol. Bull.* 75, 424–429. doi: 10.1037/h0031246
- Gur, D., Rockette, H. E., Armfield, D. R., Blachar, A., Bogan, J. K., Brancatelli, G., et al. (2003). Prevalence effect in a laboratory environment. *1. Radiology* 228, 10–14. doi: 10.1148/radiol.2281020709
- Hardmeier, D., Hofer, F., and Schwaninger, A. (2005). The x-ray object recognition test (x-ray ort) – a reliable and valid instrument for measuring visual abilities needed in x-ray screening. *IEEE ICCST Proc.* 39, 189–192.
- Hardmeier, D., Hofer, F., and Schwaninger, A. (2006). "Increased detection performance in airport security screening using the X-Ray ORT as pre-employment assessment tool" in *Proceedings of the 2nd International Conference on Research in Air Transportation, ICRAT 2006* (Belgrade), 393–397.
- Hofer, F., and Schwaninger, A. (2004). Reliable and valid measures of threat detection performance in X-ray screening. *IEEE ICCST Proc.* 38, 303–308.
- Humphrey, G. K., and Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Can. J. Psychol.* 46, 170–190. doi: 10.1037/h0084320
- Koller, S., and Schwaninger, A. (2006). "Assessing X-ray image interpretation competency of airport security screeners," *Proceedings of the 2nd International Conference on Research in Air Transportation, ICRAT 2006* (Belgrade), 399–402.
- Koller, S. M., Hardmeier, D., Michel, S., and Schwaninger, A. (2008). Investigating training, transfer, and

- viewpoint effects resulting from recurrent CBT of X-ray image interpretation. *J. Transp. Secur.* 1, 81–106. doi: 10.1007/s12198-007-0006-4
- Kosslyn, S. M. (1994). *Image and Brain. The Resolution of the Imagery Debate*. Cambridge, MA: MIT Press.
- Kravitz, D. J., Vinson, L. D., and Baker, C. I. (2008). How position dependent is visual object recognition? *Trends Cogn. Sci.* 12, 114–122. doi: 10.1016/j.tics.2007.12.006
- Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.
- Marr, D., and Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. B* 200, 269–294. doi: 10.1098/rspb.1978.0020
- Megherbi, N., Flitton, G. T., and Breckon, T. P. (2010). “A classifier based approach for the detection of potential threats in CT based baggage screening,” in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP)*, (Hong Kong), 1833–1836. doi: 10.1109/ICIP.2010.5653676
- Mendes, M., Schwaninger, A., Strebel, N., and Michel, S. (2012). “Why laptops should be screened separately when conventional x-ray screening is used,” in *Proceedings of the 46th IEEE International Carnahan Conference on Security Technology* (Boston, MA). doi: 10.1109/CCST.2012.6393571
- Mery, D., Mondragon, G., Rizzo, V., and Zuccar, I. (2013). Detection of regular objects in baggage using multiple X-ray views. *Insight* 55, 16–20. doi: 10.1784/insi.2012.55.1.16
- Michel, S., de Ruiter, J. C., Hogervorst, M., Koller, S. M., Moerland, R., and Schwaninger, A. (2007). “Computer-based training increases efficiency in X-ray image interpretation by aviation security screeners,” in *Proceedings of the 41st Carnahan Conference on Security Technology* (Ottawa).
- Michel, S., Mendes, M., and Schwaninger, A. (2010). “Can the difficulty level reached in computer-based training predict results in x-ray image interpretation tests?,” in *Proceedings of the 44th Carnahan Conference on Security Technology* (San Jose, CA).
- Michel, S., and Schwaninger, A. (2009). “Human machine interaction in X-ray screening,” in *Proceedings of the 43rd IEEE International Carnahan Conference on Security Technology* (Zurich). doi: 10.1109/CCST.2009.5335572
- Mitsumatsu, H., and Yokosawa, K. (2003). Efficient extrapolation of the view with a dynamic and predictive stimulus. *Perception* 32, 969–983. doi: 10.1068/p5068
- Palmer, S., Rosch, E., and Chase, P. (1981). “Canonical perspective and the perception of objects,” in *Attention and Performance IX*, eds J. Long and A. Baddeley (Hillsdale, NJ: Lawrence Erlbaum), 135–151.
- Peissig, J. J., and Tarr, M. J. (2007). Visual object recognition: do we know more than we did 20 years ago? *Annu. Rev. Psychol.* 58, 75–96. doi: 10.1146/annurev.psych.58.102904.190114
- Pike, G. E., Kemp, R. I., Towell, N. A., and Phillips, K. C. (1997). Recognizing moving faces: the relative contribution of motion and perspective view information. *Vis. Cogn.* 4, 409–437. doi: 10.1080/713756769
- Poggio, T., and Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature* 343, 263–266. doi: 10.1038/343263a0
- Pollack, I., and Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychon. Sci.* (75), 125–126.
- Schwaninger, A. (2005). “Objekterkennung und Signaldetektion,” in *Praxisfelder der Wahrnehmungspsychologie*, ed B. Kersten (Bern: Huber), 108–132.
- Schwaninger, A. (2003a). Reliable measurement of threat detection. *Airport* 1, 22–23.
- Schwaninger, A. (2003b). Evaluation and selection of airport security screeners. *Airport* 2, 14–15.
- Schwaninger, A., Hardmeier, D., and Hofer, F. (2005a). Aviation security screeners visual abilities and visual knowledge measurement. *IEEE Aerosp. Electron. Syst.* 20, 29–35.
- Schwaninger, A., Michel, S., and Bolting, A. (2005b). Towards a model for estimating image difficulty in x-ray screening. *IEEE ICCST Proc.* 39, 185–188. doi: 10.1109/CCST.2005.1594875
- Schwaninger, A., Michel, S., and Bolting, A. (2007). “A statistical approach for image difficulty estimation in X-ray screening using image measurements,” in *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization* (New York, NY: ACM Press), 123–130. doi: 10.1145/1272582.1272606
- Singh, S., and Singh, M. (2003). Explosives detection systems (EDS) for aviation security. *Signal Process.* 83, 31–55. doi: 10.1016/S0165-1684(02)00391-2
- Tarr, M. J. (1995). Rotating objects to recognize them: a case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychon. Bull. Rev.* 2, 55–82. doi: 10.3758/BF03214412
- Ullmann, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA: MIT Press.
- von Bastian, C. C., Schwaninger, A., and Michel, S. (2008). Do multi-view X-ray systems improve X-ray image interpretation in airport security screening? *Zeitschrift für Arbeitswissenschaft* 3, 166–173.
- von Bastian, C. C., Schwaninger, A., and Michel, S. (2010). Color impact on security screening. *Aerospace Electron. Syst. Mag.* 25, 33–38. doi: 10.1109/MAES.2010.5631724
- Vuong, Q. C., and Tarr, J. T. (2004). Rotation direction affects object recognition. *Vision Res.* 44, 1717–1730. doi: 10.1016/j.visres.2004.02.002
- Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., and Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *J. Exp. Psychol. Gen.* 136, 623–638. doi: 10.1037/0096-3445.136.4.623
- Wolfe, J. M., and Van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Curr. Biol.* 20, 121–124. doi: 10.1016/j.cub.2009.11.066
- Yoo, K. E. (2009). “Screening on the ground, security in the sky,” in *Protecting Airline Passengers in the Age of Terrorism*, eds P. Seidenstat and F. X. Splane (Santa Barbara, CA: Praeger Security International), 63–85.
- Yoo, K. E., and Choi, Y. C. (2006). Analytic hierarchy process approach for identifying relative importance of factors to improve passenger security checks at airports. *J. Air Transp. Manag.* 12, 135–142. doi: 10.1016/j.jairtraman.2005.11.006

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 May 2013; paper pending published: 22 July 2013; accepted: 19 September 2013; published online: 18 October 2013.

Citation: Mendes M, Schwaninger A and Michel S (2013) Can laptops be left inside passenger bags if motion imaging is used in X-ray security screening? *Front. Hum. Neurosci.* 7:654. doi: 10.3389/fnhum.2013.00654

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Mendes, Schwaninger and Michel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Suspiciousness perception in dynamic scenes: a comparison of CCTV operators and novices

Christina J. Howard^{1*}, Tom Troscianko², Iain D. Gilchrist², Ardhendu Behera³ and David C. Hogg³

¹ Division of Psychology, Nottingham Trent University, Nottingham, UK

² School of Experimental Psychology, University of Bristol, Bristol, UK

³ School of Computing, University of Leeds, Leeds, UK

Edited by:

Kenneth C. Scott-Brown, University of Abertay Dundee, UK

Reviewed by:

Anne Hillstrom, University of Portsmouth, UK

Damien Litchfield, Edge Hill University, UK

*Correspondence:

Christina J. Howard, Nottingham Trent University, Burton Street, Nottingham, NG1 4BU, UK

e-mail: Christina.Howard@ntu.ac.uk

Perception of scenes has typically been investigated by using static or simplified visual displays. How attention is used to perceive and evaluate dynamic, realistic scenes is more poorly understood, in part due to the problem of comparing eye fixations to moving stimuli across observers. When the task and stimulus is common across observers, consistent fixation location can indicate that that region has high goal-based relevance. Here we investigated these issues when an observer has a specific, and naturalistic, task: closed-circuit television (CCTV) monitoring. We concurrently recorded eye movements and ratings of perceived suspiciousness as different observers watched the same set of clips from real CCTV footage. Trained CCTV operators showed greater consistency in fixation location and greater consistency in suspiciousness judgements than untrained observers. Training appears to increase between-operators consistency by learning “knowing what to look for” in these scenes. We used a novel “Dynamic Area of Focus (DAF)” analysis to show that in CCTV monitoring there is a temporal relationship between eye movements and subsequent manual responses, as we have previously found for a sports video watching task. For trained CCTV operators and for untrained observers, manual responses were most highly related to between-observer eye position spread when a temporal lag was introduced between the fixation and response data. Several hundred milliseconds after between-observer eye positions became most similar, observers tended to push the joystick to indicate perceived suspiciousness. Conversely, several hundred milliseconds after between-observer eye positions became dissimilar, observers tended to rate suspiciousness as low. These data provide further support for this DAF method as an important tool for examining goal-directed fixation behavior when the stimulus is a real moving image.

Keywords: eye movements, scene perception, expertise, security and human factors, visual search

INTRODUCTION

Studies of naturalistic task performance have used eye movements as a measure of attentional deployment (e.g., Land, 1999; Findlay and Gilchrist, 2003; Underwood et al., 2003). Here we measure eye movements to investigate such attentional deployment in the context of closed-circuit television (CCTV) monitoring. CCTV monitoring is both a good model task in which to study the deployment of goal directed attention more generally and an important task more specifically because of its increased deployment in security and policing.

Recent research has examined human performance in some aspects of CCTV monitoring. For example, Troscianko et al. (2004) showed that people were able to anticipate antisocial behavior in the near future from CCTV footage. Others have examined the limitations of the use of CCTV footage in identifying unfamiliar individuals, although face recognition appears to be surprisingly resistant to viewpoint changes or poor image quality when individuals are familiar to observers (Bruce et al., 2001).

However, much less is known about the dynamic allocation of attention during CCTV monitoring. Stainer et al. (2011) showed that eye movements in multiscreen displays tend to fall near the centres of individual video screens in a multiscreen display, and suggested that the lack of scene continuity and spatial contiguity between individual screens causes each one to be treated as an independent stimulus. However, there appears to be some direct competition between screens: Howard et al. (2011) showed that eye movements are driven to a great extent by the relative suspiciousness of different concurrent video screens in the display. CCTV is clearly a very rich visual stimulus and results are now beginning to emerge on several of the many aspects of human interaction with these stimuli. However, we are not aware of any work that seeks to examine exactly how attention is used within a single screen during on-line monitoring and decision making about video events, and this is addressed by the current study.

Examining how people perceive CCTV footage is one example of the more general task of perception of moving scenes.

Much research has been conducted into the question of how we perceive static scenes and in particular, how long it takes to extract different types of visual information from scenes. Strikingly, the general “gist” of a scene can be processed from extremely brief (less than a tenth of a second) displays (e.g., Rousselet et al., 2005) or from a single glance (Biederman et al., 1974; Fei-Fei et al., 2007). When making global property classifications of a scene (e.g., naturalness, openness) and basic level categorisations (e.g., ocean, mountain), observers can reach asymptote levels of performance in 100 ms (Greene and Oliva, 2009). Little is known, however, about the time course of the perception of dynamic scenes. Of course outside of the laboratory, visual stimuli are rarely static and so it is important to investigate the extent to which the work with static images generalises to moving scenes. We recently examined this issue (Howard et al., 2010) by asking observers to make a continuous semantic judgement about a video of a semi-constrained real-world scenario: a football match. We found that responses continuously lagged behind eye movement behavior by over a second, suggesting that evaluation of moving scenes proceeds relatively slowly.

As well as being a task involving perception of real moving scenes, the task of monitoring CCTV images typically requires observers to search for and assess locations in the scene of maximum perceived suspiciousness. In this sense, whilst this task is very different from traditional visual search, some comparisons can be made. The task here could be considered as a visual search task for an extremely high level semantic target which is visually unspecified and could therefore take many different visual forms. From the traditional visual search literature, although target templates with high specificity are optimal for guiding attention, search can be driven by imprecise target information such as target categories (Malcolm and Henderson, 2009; Schmidt and Zelinsky, 2009; Yang and Zelinsky, 2009). Consistent with this, observers can use flexible target templates for search that are tolerant to some changes in target appearance e.g., changes in scale and orientation (Bravo and Farid, 2009, 2012). However, much less is known about the extent to which attention can be guided by very high-level semantic interpretation of scenes.

Another aspect of performance we addressed in the experiments presented here is the effect of expertise since our observers were both trained CCTV operators and untrained undergraduate observers. Howard et al. (2010) found that expertise affected the pattern of eye movements and the relationship between eye movements and responses. Specifically, individuals with more experience watching football matches made eye movements to goal relevant areas of the scene earlier than non-experts, and were thus able to spend longer evaluating the scene before making their responses. This would suggest that expertise may affect eye movement behavior in this CCTV monitoring task in a similar way i.e., that CCTV operators will be more able to direct their eye movements towards goal relevant areas of the scene than untrained observers. Indeed in a meta-analysis of several hundred effect sizes of expertise on eye movement behavior, Gegenfurtner et al. (2011) recently reported robust effects of expertise acting to increase frequency of fixations on goal relevant information and to reduce latencies for first fixations on the these areas. Some have claimed that this attention to goal relevant information

(and consequently, reduced attention to irrelevant information) underlies the effect of expertise in visual tasks (Haider and Frensch, 1996). This “knowing what to look for” is likely in the CCTV operators since they are familiar with environments shown in the CCTV footage and the likely types of suspicious behaviors that may occur. For this reason, we hypothesised that CCTV operators would be able to process the scenes more efficiently than untrained observers. Scene “gist” or general layout can be extracted very rapidly but more detailed processing of scene content can take many seconds (e.g., Tatler et al., 2003). Given that CCTV monitoring requires complex semantic evaluation of scenes, we reasoned that this task would be likely to incur slower processing times and therefore may be sensitive to the effects of expertise.

There is evidence from a range of visual tasks that expertise affects processing efficiency. For example, in visual search tasks, domain-relevant expertise appears to enable observers to process a wider portion of their visual field at a time during visual search tasks (Hershler and Hochstein, 2009) and expert chess players can process visual information from across the visual field rapidly and with few fixations (Reingold et al., 2001). Similar processing advantages are seen in more applied tasks: expertise in driving facilitates wider visual scanning of road scenes (Underwood et al., 2002) and expertise in musical sight reading fosters greater storage of visual information from fixations on musical text (Furneaux and Land, 1999). We therefore hypothesised that CCTV operators would be able to process the unfolding CCTV scenes more efficiently than untrained observers. We expected CCTV operators to display greater between-observer consistency of gaze locations since their attention should be more consistently drawn to “suspiciousness” rather than other aspects of the scene or events within them.

We present here a task in which the attentional deployment of both trained operators and untrained, naïve observers is measured through eye tracking, whilst monitoring a single scene for potentially suspicious events. In this task, manual responses take the form of pushing a joystick to reflect the current degree to which events in the scene are perceived to be suspicious. The CCTV monitoring task requires continuous appraisal of the semantic content of scenes, and the evaluation of the current intentions and behaviors of people displayed in the scene. We will show a surprising degree of between-observer consistency in eye gaze locations in the scene, particularly between the gaze locations of trained CCTV operators. We will also show that periods of particularly high between-observer consistency in gaze positions are correlated with ratings of perceived suspiciousness in the scene. We will show that durations of between-observer eye position convergence are related to judgments of higher suspiciousness in the scenes, and that CCTV operators show longer periods of eye position convergence than untrained observers.

The current experiment demonstrates that this Dynamic Area of Focus (DAF) method works for a CCTV task as it did for the football watching task (Howard et al., 2010) with a different video stimulus and a very different semantic evaluation task. The DAF method is again shown to be a powerful tool for examining continuous perceptions of dynamic scenes without the need to

analyse the content of the videos nor to measure low level physical characteristics or salience of the stimuli.

EXPERIMENT 1: UNTRAINED OBSERVERS

A computer programme was written in C++ to display a series of one-minute video clips of real CCTV footage obtained from Manchester City Council. Observers viewed a total of 40 one-minute clips comprising four clips from each of 10 different CCTV cameras. Observers viewed the videos in four blocks of ten minutes with breaks in-between, and the order of the 40 clips was quasi-randomised. The 10 cameras were chosen to represent as wide a range as possible in terms of the visual characteristics of the scenes. The ten camera views were as follows: night-time view of a carpark, pedestrian crossing, shopping street underpass, cash point at junction, busy retail street, landscaped open area, pedestrianised street, entrance to nightclub at night, bus stops and city centre street at night.

Observers made a constant judgement about the current perceived level of suspicious events in the scene by moving a joystick. Joystick ratings were sampled at 100 Hz resulting in a series of data points for each one-minute video stimulus. Note that this is a continuous response to a continuous stimulus, and the response takes the form of a rating about the video stimulus. Observers' eye positions were recorded at 25 Hz throughout the task using the ASL Mobile Eye head-mounted eye-tracker and Eye Vision software.

Video clips measuring 27 degrees by 22 degrees of visual angle were projected in a dimly lit room against a white background using a Canon SX6 projector onto a screen at a distance of 1.6 m. Black chequerboard markers subtending 4×4 degrees were placed at each corner of the video display such that a computer algorithm could be used after data collection to stabilise eye position recording for changes in head position.

OBSERVERS

Observers were thirty three undergraduate and postgraduate students at Bristol University, all naïve as to the purpose of the experiment, four of whom were male and twenty-nine female. The mean age of observers was 20 years, ranging from 18 to 34 years. All had normal or corrected-to-normal vision.

PROCEDURE

Observers were given written instructions as follows. They were asked to watch several videos of "urban scenes" and to monitor them for any suspicious events which, if seen in real life, might cause them to alert relevant authorities. They were asked to move a joystick according to what they perceived as being the current level of suspicious behavior in the video. They were told that at all times, the joystick should reflect what they perceived as being the current level of suspicious behavior. For instance, if they thought that the video was showing something very suspicious, they were told to move the joystick fully forwards for the duration of the suspicious events. If they perceived that there was currently absolutely no suspicious behavior, they were told not to move the joystick at all. They were informed that they could push the joystick to any level in-between these two extremes and that we would record the position of the joystick throughout the experiment. The joystick was in part chosen as a method of collecting

responses as in the real CCTV control room where our operators were employed, operators may use a joystick to control the level of zoom on a particular camera, pushing the joystick further to zoom further in and vice versa. Hence this manual response was compatible with behaviors that occur in real CCTV monitoring contexts. Observers were asked to keep their hand on the joystick at all times to minimise the impact of manual reaction times.

EXPERIMENT 2: TRAINED CCTV OPERATORS

The method for Experiment 2 was similar to that used in Experiment 1, apart from the following differences. Video clips measuring 22 degrees by 18 degrees were projected against a white background. Observers viewed a total of 80 one-minute clips comprising eight clips from each of 10 different CCTV cameras. The 10 cameras were the same as those used in Experiment 1, but using twice as many clips from each: an extra four clips were used from each camera in addition to those used in Experiment 1. Observers viewed the videos in eight blocks of ten minutes with breaks in-between, and the order of the 80 clips was quasi-randomised. Observers completed the experiment in two sessions over the week long testing period.

OBSERVERS

Observers were eleven trained CCTV operators working in the Manchester City Council CCTV control room of whom two were female and nine were male. All were naïve as to the purpose of the experiment but were aware that we were investigating the way that operators carry out their job, and the things that they look for whilst monitoring CCTV. All had normal or corrected-to-normal vision. The mean age of observers was 37 years, ranging from 23 to 60 years.

The trained CCTV operators differed from the untrained observers in terms of expertise since they had received training in the task of CCTV monitoring and all were currently employed as CCTV operators in the Manchester control room at the time of testing. Operators ranged in their level of experience from around six months to many years' experience in the job. Different individuals undoubtedly had achieved differing levels of expertise in the task but on average these observers would certainly be more familiar with the task and the types of CCTV images used than the untrained observers.

RESULTS

RATINGS OF PERCEIVED SUSPICIOUS EVENTS

The mean suspiciousness rating across videos for the untrained observers was 0.397 and for the operators was 0.387 (0.417 for operators watching the 40 videos seen by both groups). There was no difference in the overall suspiciousness ratings given by the two groups ($t(39) = 0.574$, $p = 0.57$). The range of suspiciousness ratings across videos for the untrained observers was 2.769 and for the operators was 2.258 (2.313 for the operators when viewing those 40 videos also seen by the untrained observers). The untrained observers showed a greater range of ratings at each given time (between-observer variability in ratings) than the trained observers ($t(39) = 3.16$, $p < 0.01$). In other words, trained observers' ratings were more consistent with one another at any given time.

CONSISTENCY OF EYE GAZE POSITION

For each frame in each video stimulus, we calculated a measure of spread of eye positions. In an example frame, there will be one recorded eye position for each observer, each with a horizontal and vertical position. As a measure of spread in eye positions across observers at a particular time, we took the mean of the interquartile ranges of the horizontal and vertical eye positions. We used the interquartile range as a measure of variability to minimise the influence of position outliers. The subsequent “spread value” is a measure of the extent to which all observers were looking at the same part of the screen at the same time.

The mean spread measure, expressed as a fraction of the size of the display was 19.0% (SD = 3.0%) for the operators (18.8%, SD = 3.0%, for operators watching those 40 videos seen by both groups) and 22.0% (SD = 3.2%) for the untrained observers. Trained observers showed less eye position spread than untrained observers ($t(39) = 6.07, p < 0.01$) indicating that they were more likely to be looking at a similar point in the videos as one another at any particular time.

DYNAMIC AREA OF FOCUS ANALYSIS: RELATIONSHIP BETWEEN RATINGS AND EYE GAZE POSITION

The DAF analysis captures the relationship between moment-by-moment eye movement behavior and judgements of a group of observers viewing the same dynamic stimulus. To perform this analysis, we calculated estimates of the temporal relationship between eye movement behavior and responses. For these and all subsequent analyses, we calculated normalised suspiciousness ratings as follows: we first calculated the total overall mean and standard deviation of suspiciousness ratings for each observer and used these to normalise each observer’s data set. For each video stimulus, we then calculated for each frame, the median of the normalised ratings. We chose the median to minimise the effects of outliers in the data.

To calculate an estimate of the time lag between eye movements and responses, we performed correlations between eye position spread and these normalised manual responses. At each point in time for a particular video, there will be one value of between-observer eye position spread and one value of suspiciousness ratings across observers as defined above. Note that across each whole video, the manual responses and the eye movements are both time series data and hence do not represent a single point in time but rather a continuous stream of events that relate to the continuous video stimulus.

To test for a non-zero lag, we performed each correlation after artificially shifting the eye spread data forwards and backwards in time. For instance, to test for a 100 ms lag, we shifted the eye spread data 100 ms backwards in time relative to the response data and recalculated the correlation value. At the best estimate of the lag, this correlation should be maximally negative. The lag estimate is the estimated time delay between changes in eye movements spread and the manual responses associated with them. For example, a reduction in eye position spread might be associated with an increase in suspiciousness ratings a short while later, whilst an increase in eye position spread is likely to be associated with a decrease in suspiciousness ratings soon afterwards.

Missing data created by these artificial time shifts were replaced with the mean value of spread for that stimulus. We tested each

lag moving in steps of 10 ms through the range of up to 10 seconds both forwards and backwards in time. The results of these lag analyses are shown in **Figures 1, 2** below. Error bars were obtained by bootstrapping: we sampled observers (with replacement) to create bootstrapped “new” data sets and obtained the lag for each of these data sets. This bootstrapping cycle was repeated 50,000 times and the standard error of this set of lags was then calculated.

We reasoned that the spread values and ratings would be related to one another but not necessarily in a straightforward way. Two factors are likely to drive eye movements, namely goal relevance (in this case, suspiciousness) and also low-level image salience differences such as differences in brightness, colour and motion in the scene. In addition, there may be multiple areas of a scene for which either or both of these drivers attracts attention

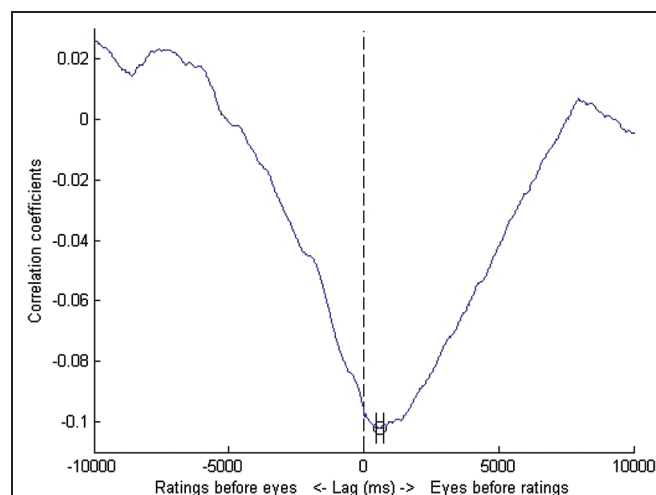


FIGURE 1 | DAF analysis for untrained observers. The maximally negative correlation was obtained at a lag of 580 ms i.e., the eyes led manual responses by a lag of just over half a second. Error bars represent standard errors obtained by bootstrapping.

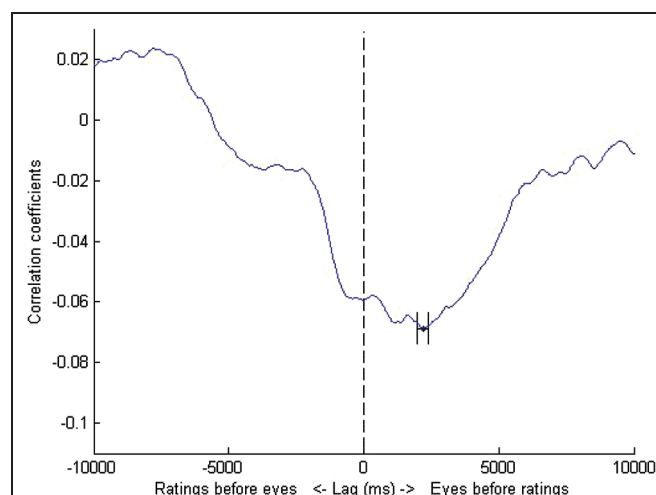


FIGURE 2 | DAF analysis for trained observers. The maximally negative correlation was obtained at a lag of 2180 ms i.e., the eyes led manual responses by a lag of just over two seconds. Error bars represent standard errors obtained by bootstrapping.

at any time. The extent to which observers will tend to fixate the same areas of the screen as one another (producing low spread values) was considered an empirical question. However, changes in eye position spread that occur close in time to changes in suspiciousness ratings may reflect scene events that are goal relevant. Of course, this is not to say that goal relevant events may not also be accompanied by changes in low-level scene salience, but by looking for the antecedents of high and low suspiciousness ratings, we will identify the extent to which eye position spread changes are related to goal relevance. We reasoned that if goal relevance is reliably related to eye position spread, then there will be a negative (and lagged) relationship between eye position spread and suspiciousness ratings. Overall, those events judged to be suspicious will tend to be preceded by different observers looking in similar places, and conversely, that events judged not to be suspicious will tend to be preceded by different observers looking in dissimilar places to one another. Since we are using a suspiciousness judgement along a continuum of joystick positions and not a discrete suspicious/not suspicious judgement, we must also consider that intermediate suspiciousness ratings will tend to be preceded by intermediate spread values, to an extent determined by how suspicious the scenes are judged to be.

Changes in eye position spread driven only by salience (for example, everyone's eyes being drawn to a street light being turned on) will not be accompanied by a change in suspiciousness rating, and hence can only serve to decrease the strength of the correlation. Similarly, if there are two or multiple events in different parts of a scene that appear suspicious at any given time, this would produce high eye position spread measures and high suspiciousness ratings, thus decreasing the strength of the negative correlation between eye spread and ratings.

For both groups of observers at the obtained lags, eye position spread was negatively correlated with response (E1: $r = -0.10$, $p < 0.05$, E2: $r = -0.07$, $p < 0.05$) and these correlations coefficients between videos were significantly more negative than zero (E1: $t(39) = -2.84$, $p < 0.01$, E2: $t(79) = -3.19$, $p < 0.01$). The magnitude of the lag was much greater for the trained than the untrained observers. For those 40 videos seen by both sets of observers, the trained observers also showed a negative correlation between eye position spread and responses ($r = -0.07$, $p < 0.01$). This correlation was maximally negative at a lag of 1130 ms and was significantly more negative than zero ($t(39) = -2.56$, $p = 0.01$). Although this lag value is shorter than that seen when the data is analysed for the whole set of 80 videos seen by trained participants, it is still substantially longer than the lag found for untrained participants.

RELATIONSHIP BETWEEN RATINGS AND EYE GAZE CONVERGENCE

For each frame in each video, we classed the spread measure as either “low” or “not low” using a threshold of one standard deviation below the overall mean spread value. We then calculated the durations of periods of time in which this thresholded spread value remained consistently “low” on subsequent frames. Overall for the untrained observers, the mean duration of these low spread (or equivalently, “convergence”) periods was 128 ms. For the untrained observers, there was a significant correlation (see **Figure 3**) between the mean convergence duration for each one minute video stimulus and the mean suspiciousness rating given

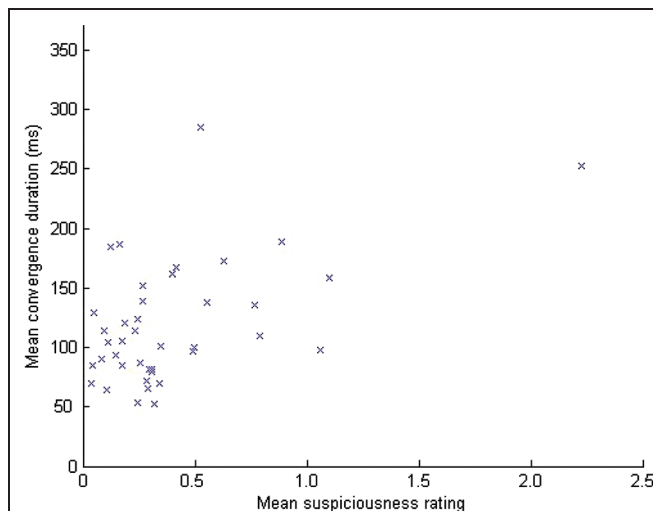


FIGURE 3 | The relationship between convergence period duration and ratings of perceived suspiciousness for untrained observers.

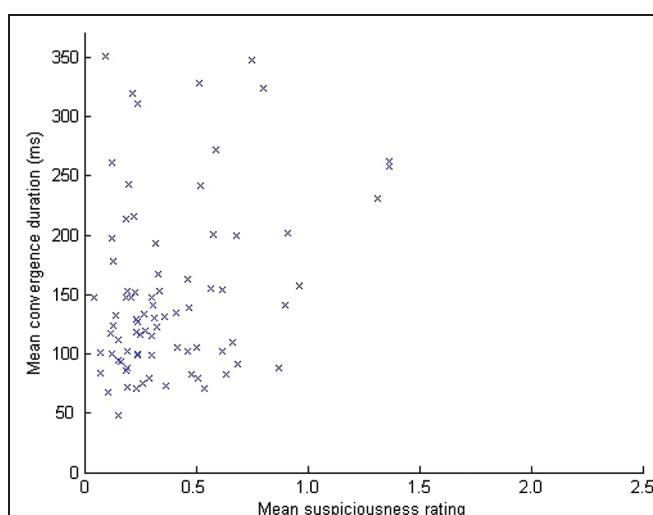


FIGURE 4 | The relationship between convergence period duration and ratings of perceived suspiciousness for trained CCTV operators.

to that video ($r(39) = 0.509$, $p < 0.01$). The data for one of the video stimuli was more than two standard deviations above the mean on both variables of mean convergence duration and mean suspiciousness rating, but the correlation remained significant even after excluding this data point ($r(38) = 0.324$, $p = 0.044$).

Overall for the trained CCTV operators, the mean duration of these low spread “convergence” periods was 151 ms. As shown in **Figure 4**, there was a significant correlation between the mean convergence duration and the mean suspiciousness rating for each one minute video stimulus ($r(79) = 0.296$, $p < 0.01$). For those 40 videos also seen by the untrained observers, there was also a significant correlation ($r(39) = 0.356$, $p < 0.024$).

There was no significant difference in the strength of correlations between trained and untrained observers ($p > 0.05$) for any of the correlations reported above (correlations between convergence duration and ratings). However, for the 40 videos seen by both sets of observers, mean convergence period durations were longer for the CCTV operators than the untrained observers

($t(39) = 3.540$, $p < 0.01$) indicating that CCTV operators were more likely to spend longer periods of time consistently looking at the same part of the screen as one another than was the case for the untrained observers.

DISCUSSION

For this complex task, the DAF analysis reveals a temporal relationship between eye movements and subsequent manual responses. These results indicate that as found previously for a sports monitoring task (Howard et al., 2010), this DAF method is a powerful one for examining eye movement behavior towards moving stimuli. The method circumvents the need to analyse events within the video, nor low-level physical properties of the stimulus in order to examine goal-directed attention.

We found a significant negative correlation between eye position spread data and manual responses. In other words, observers tended to push the joystick to indicate perceived suspiciousness at times shortly after between-observer eye position differences decreased, and tended to give low suspiciousness ratings shortly after eye position spread increased. The time lag between eye position spread changes and corresponding suspiciousness judgement responses is relatively long and in the order of hundreds of milliseconds to seconds. This lag between eye movements and corresponding suspiciousness responses was longer in CCTV operators than in untrained observers. CCTV operators show reduced between-observer eye position spread and longer periods of eye position convergence than untrained observers. They also showed a greater degree of between-observer consistency in terms of suspiciousness ratings. We also show a relationship between the mean durations of eye position convergence events and the suspiciousness assigned to different videos. As discussed above regarding the lag analysis between eye movement spread and responses, both salience and goal relevance are likely to drive eye movements to different extents depending on the nature of the events depicted at any given time. However, we find a relationship between eye position convergence duration and ratings. Hence, like eye position spread, eye position convergence duration appears to be a strong enough indicator of goal relevance to show up over and above any effects of salient but not task-irrelevant events or the effects of multiple simultaneous relevant events.

This task was simpler than that of monitoring many screens at once as is the case in real CCTV control rooms. However, for multiple screens, untrained observers are able to perform this task since their eye movements are driven by goal relevance (i.e., suspiciousness) to a much greater extent than they are influenced by low-level image properties (Howard et al., 2011). The current study shows that both trained and untrained observers are able to respond to a single screen in such a way that their eye movements are related to goal relevance. Therefore, it seems likely that trained CCTV operators would be able to perform multiple screen monitoring to the same level or to a superior extent than untrained individuals and this deserves future investigation. In real CCTV control rooms, operators will need to monitor very many screens at once for suspicious activity and this method provides a starting point for understanding such a complex task. One way in which this method captures some of the processes involved in real CCTV monitoring is the use of joystick pushing/pulling as

the manual response since in the real control room a similar joystick is used for operators to zoom into areas of interest or suspicious activity. The use of real CCTV footage from the urban areas familiar to operators and a realistic suspiciousness judgement are also very close to the demands of real CCTV monitoring in the control room. For these reasons, there should be a good degree of generalisability from our findings here to real CCTV tasks in terms of the relationship between eye movements and manual responses.

The task of judging perceived suspiciousness was an inherently ambiguous one. For example, footage of individuals “loitering” in a car park at night may be judged as suspicious to a greater or lesser extent by different individuals depending on their interpretation of the events depicted. In fact, it has been previously shown for the same task that mood state can alter these judgements whilst monitoring CCTV (Cooper et al., 2013) reinforcing the subjective nature of these judgements. Therefore, even for the trained CCTV operators, there can be no objectively “correct” rating. We did not attempt to provide a benchmark of “correct” responses for this reason, though anecdotally while watching the videos, higher ratings of suspiciousness were associated with behavior such as that mentioned above in a car park, similar “loitering” around the entrance to nightclubs after dark or in an urban shopping area pedestrian underpass. A formal analysis of the content of video that is judged to be more or less suspicious is possible though it is beyond the scope of the work presented here. One can identify those periods of time in different videos that were given the highest suspiciousness ratings, and locate areas of the screen that were fixated just before the ratings were given, allowing for the time lag between eye movements and ratings. Whilst the spread measure is relative in that it describes eye positions only in terms of how close they are to other observers’ fixations, one could identify the location of the centroid of these between-observer fixations to locate the most goal relevant areas of the screen whilst suspiciousness ratings are high. Characterising the content of such activity might be done either qualitatively by coding different behavior-environment interactions, or more quantitatively by looking for physical qualities of these video events.

The fact that analysis of video content is not necessary (though it is possible) for this technique makes this a powerful new tool for examining eye movements to dynamic scenes. One of the main challenges in studying eye movements to complex moving stimuli, is how to associate eye movements with different aspects of the stimuli, and therefore how to compare eye movements between observers. For simple stimuli such as a single target or a relatively small number of moving targets, dynamic areas of interest are a common method of analysis. One can use such a moving area to calculate fixations and dwell times etc., to particular stimuli of interest. However, this technique becomes unwieldy when there are very many targets, complex motion, shape changes, occlusion events or where stimuli are complex enough (such as in real-world scenes) that defining what is a target becomes non-trivial. There are additional problems with the use of dynamic areas of interest, such as how to define saccadic overshooting or undershooting, catch-up saccades and extrapolatory eye movements. The current technique circumvents all these problems.

We find that perception of dynamic scenes in this task proceeds relatively slowly: observers' responses lag behind eye movement convergence by a minimum of several hundred milliseconds. This is considerably longer than the typical time periods required for rapid evaluations of static scenes such as the "gist" which can be extracted effectively in around 100 ms (e.g., Biederman et al., 1974; Rousselet et al., 2005; Fei-Fei et al., 2007). The task here was very different from typical gist perception studies in several respects. Typical gist perception studies present stimuli only for a limited time and often measure a threshold of simple scene judgements. Here, however, the judgement was continuous and required semantic processing beyond simple scene-type judgements. More complex perceptual representations of scenes have been studied in the context of encoding images into memory. For example, Tatler et al. (2003) showed that memory representations of gist formed very rapidly. However, other judgements about more detailed aspects of the scene, like shapes, colours and positions of the scene elements benefitted from very many more seconds exposure up to 10 seconds. From this and later similar findings (Melcher, 2006) one might assume that the time course of semantic scene perception is as slow as this. However it is entirely possible that the limit in these memory studies may have occurred only at the stage of encoding and not perceptual processing. The results of the current study indicate a slow time course for semantic perception of dynamic scenes that ranges from several hundred milliseconds to several seconds.

Our findings here are somewhat consistent with earlier findings for a similar task but with a different stimulus (Howard et al., 2010) where observers watched a real videotaped football match and made continuous judgements about imminent goal likelihood. In this sports evaluation task observers' manual responses lagged behind gaze convergence by 1360 ms (non-experts) or 2260 ms (experts). The reason for the longer lags seen in the sports task than the CCTV task is not clear, but there are several differences between the two studies. The sports task is more constrained in terms of likely events. The CCTV task, by contrast, contains several different scenes, and several different types of events that are relevant to the suspiciousness judgement e.g., loitering in a car park, activity in a city shopping street, around a cashpoint etc. The CCTV task involved viewing ten different urban scenes with frequent changes between scenes. In contrast, the sports task stimulus was a single football match with a continuous shot from the same camera. Hence, the CCTV task contains more uncertainty in terms of what counts as the relevant perceptual variable, "suspiciousness", than does the sports task where the relevant variable is "goal likelihood". There may also be a greater social perception component inherent in the CCTV stimulus since the task involves making judgements about individuals' intentions and interactions with one another. Nonetheless, some comparison can be made between the two tasks since both require continuous semantic evaluation of moving scenes and appear to incur processing delays over several hundred milliseconds.

Two factors are likely to make our estimates here for the time course of dynamic scene perception longer than that previously reported for static scene evaluations. First, our method includes the time it takes to prepare and execute a response to the visual stimulus. However, reaction times to produce a manual response

to stimuli tend to be in the order of 200–250 ms (Goldstone, 1968; Green and von Gierke, 1984) and the magnitude of the lags here implicates additional contributing processes. Second, and most interestingly, the nature of the continuous task itself is likely to have caused these large time lags. Here we used a continuous video stimulus within which events unfold over time. One reason why perception of dynamic scenes may lag behind visual events is that the visual system often integrates information over a temporal window of at least 100 ms (e.g., Gorea, 1986; Watamaniuk and Sekuler, 1992) which is a physical necessity for information with a temporal component such as stimulus change or motion. Hence any temporal averaging may serve to increase these lags. Lags are also likely to be increased by the complex perceptual demands inherent in making decisions about these dynamic stimuli. For example, attending to the biological motion of humans and making judgements about their intentions is attentionally demanding and particularly important when the signal is degraded, ambiguous or subject to competition from other attentionally demanding stimuli (Thompson and Parasuraman, 2012). We also know that attending to multiple regions of a scene in terms of their visual features is attentionally demanding and can incur costs in terms of temporal lags (Howard and Holcombe, 2008; Lo et al., 2012), especially under conditions of competition for attention by different stimuli with similar features. In addition, whilst attending to video stimuli, one must use sustained attention, the nature of which is known to be different from that of transient attention (Ling and Carrasco, 2006) and it is possible that processing using sustained attention proceeds relatively slowly compared to the more transient attention that can be used for more briefly presented stimuli. Hence the complex and continuous nature of the stimulus and the task likely comprise a large component of the time course of this type of dynamic scene processing.

Howard et al. (2011) showed that observers could monitor four CCTV screens at once for suspicious events. Whilst this is very different from a traditional visual search task where the search array is typically static and the target is typically very well defined, these results and the results in the current study can be considered evidence that observers can perform visual search for extremely high level semantic targets. In the current study, this high level semantic target is "suspiciousness" which may take many different visual forms. This extends the literature from more traditional visual search tasks showing that observers need not be given complete or fully determined target template information to perform visual search in scenes (e.g., Malcolm and Henderson, 2009; Schmidt and Zelinsky, 2009; Bravo and Farid, 2012). In this task, for moving, complex scenes, attention can be guided by very high-level semantic interpretation of scenes.

Trained observers showed a greater lag between eye movements and manual responses than untrained observers. This is consistent with previous data for a similar task but when making judgements about a sports match (Howard et al., 2010). In the sports task, it appeared that expert observers were more able than non-experts to move their eyes to the goal relevant areas of the scene earlier, thus allowing them more time to produce their response. A similar mechanism could be operating here if CCTV operators "know what to look for" in the scenes. This would also explain the fact that CCTV operators showed

greater between-observer consistency in eye position and longer periods of between-observer eye position convergence than our untrained observers. The fact that trained observers' ratings were more similar to one another at any given time than was the case for the untrained observers is also consistent with this picture. It is worth noting that the CCTV operators were familiar with the scenes presented in the videos and hence their expertise may lie both in the task of CCTV monitoring itself and also in their specific knowledge of the scenes and environments depicted in the video footage. The CCTV operators' reduced level of eye movement variability compared to untrained observers may account for why the relationship between eye position spread and responses was less highly correlated than it was for untrained observers. It might also help explain why the lag curve for expert observers was flatter and less pronounced than for untrained observers.

There is evidence that expertise affects visual processing efficiency in a range of different tasks. Hershler and Hochstein (2009) examined the influence of expertise during visual search. They found that experts in specific recognition of either categories of "cars" or "birds" appeared to be able to process visual information in their area of expertise from a wider portion of the search display with each fixation. This is consistent with an explanation of expertise on the grounds of a greater capacity for information processing across the spatial domain. These results are reminiscent of similar results in the vision-for-action literature. For example, Underwood et al. (2002) report that expert drivers scan a wider portion of the road scene than novices. In musical sight reading, Furneaux and Land (1999) find that experts and non-experts tended to look at positions in the musical text that are approximately one second ahead of the notes they are currently playing. However the expert musicians appeared to be able to store more information in this visual information memory buffer. In their information reduction hypothesis of skill acquisition, Haider and Frensch (1996) point towards selective attention to goal relevant information as the cause of improved performance in experts. This selection of task relevant information over non-relevant information could be operating here and elsewhere, though it is also possible that processing is more efficient even once the most critical areas of the scene are selected by attention. These two aspects of efficiency i.e., selection and post-selection processing, are difficult to tease apart in the data presented here. However, the effect of expertise in these very different types of tasks, visual search and vision for action, are likely to reside in efficiency of visual information processing, albeit at potentially different cognitive and perceptual stages.

Any visual processing efficiency differences may plausibly reduce the cognitive or mnemonic load for experts. Since the task presented here involves a high degree of perceptual, cognitive and mnemonic load, this may be especially beneficial to the experts in the current study. Cognitive complexity and memory load have been shown to influence fixation patterns with observers using fixations to regain goal relevant information under conditions of high load (Droll and Hayhoe, 2007; Hardiess et al., 2008). Hence experts may have needed to use fewer re-fixations in this manner, contributing to the difference in eye movement patterns observed between our two groups.

One further possibility for the locus of expertise in this and in our previously reported sports task (Howard et al., 2010) is that experts are better able to anticipate upcoming visual events. Some evidence that this may be the case is given by Didierjean and Marmèche (2005) who showed anticipatory representations in expert basketball players. This was evidenced by the fact that experts' comparisons between pairs of gameplay configurations was poorer when making comparisons about pairs that moved forwards in time rather than backwards. It appeared that their representations had already moved the events on in time when presented with the future configuration. Perhaps the experts in the current study and in our previous sports monitoring task were more able to predict near future events and hence use their eye movements more efficiently.

At first glance it is not clear how these visual processing differences might account for the longer lags reported here for CCTV operators than untrained observers between eye movements and manual responses. However in the football task, experts appeared to move their eyes to the relevant parts of the scene earlier, and this could have been facilitated by superior visual processing. The longer lag between eye movement convergence and manual responses may be a result of experts deliberately adopting an accuracy-over-speed strategy, perhaps as a direct result of more confidence about making goal relevant fixations. Trained operators may choose to undertake more processing before reaching a decision about manual responses. Additional time observing events is likely to result in increased visual information and decreased ambiguity about events being displayed, and it is possible that operators use a waiting strategy to minimise the number of false alarms. Indeed in the real CCTV control room, operators use a similar type of joystick to zoom in to events in real time, zooming in and out as desired depending on the unfolding events. Zooming in to more closely examine a particular stimulus incurs some information cost since it narrows the field of view in that particular camera and carries the risk of missing events occurring at other locations in the scene. Hence experts may have learned to use a conservative criterion for making suspiciousness judgements. One factor to note here is that our two groups of observers differed along many dimensions including training and experience, knowledge and expectations of the scenes presented, gender, age, socio-economic background and specific instruction in the task. Of course any or all of these factors could have contributed to this difference.

In the CCTV task presented here, our experts were trained professional CCTV operators, compared to untrained psychology undergraduates. In the football task, all the observers were undergraduate psychology students, but they differed in their level of self-reported experience watching football matches. Hence, there was a greater difference in expertise level in this CCTV monitoring task than in the football task and this may account for some of the differences in the results. Other differences between the CCTV and football tasks include the greater level of constraint about events in the football match (i.e., events typical of a football match such as passes, tackles, goal attempts, etc.) than the CCTV task, which is video footage of several different types of urban scene. Additionally, the CCTV video is potentially much

more of a task of social perception than the football task, since it requires judgements of intentions, potential future behavior and interactions between individuals. Therefore the data we present here are a second example of the successful application of this DAF method of measuring eye-hand lags in two very different contexts. The method enables the use of tasks with moving video

stimuli from real-life scenarios, as well as on-line continuous judgements about these stimuli. We demonstrate that cognitive evaluation of these moving scenes is a somewhat slow process. The ramifications of this processing time when multiple screens must be monitored, as in CCTV monitoring, may be particularly severe.

REFERENCES

- Biederman, I., Rabinowitz, J. C., Glass, A. L., and Stacy, E. W. (1974). On the information extracted from a glance at a scene. *J. Exp. Psychol.* 103, 597–600. doi: 10.1037/h0037158
- Bravo, M. J., and Farid, H. (2009). The specificity of the search template. *J. Vis.* 9, 31–39. doi: 10.1167/9.1.34
- Bravo, M. J., and Farid, H. (2012). Task demands determine the specificity of the search template. *Atten. Percept. Psychophys.* 74, 124–131. doi: 10.3758/s13414-011-0224-5
- Bruce, V., Henderson, Z., Newman, C., and Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *J. Exp. Psychol. Appl.* 7, 207–218. doi: 10.1037/1076-898x.7.3.207
- Cooper, R., Howard, C. J., Attwood, A. S., Stirland, R., Rostant, V., Renton, L., et al. (2013). Acutely induced anxiety increases negative interpretations of events in a closed-circuit television monitoring task. *Cogn. Emot.* 27, 273–282. doi: 10.1080/02699931.2012.704352
- Didierjean, A., and Marmèche, E. (2005). Anticipatory representation of visual basketball scenes by novice and expert players. *Vis. Cogn.* 12, 265–283. doi: 10.1080/13506280444000021
- Droll, J. A., and Hayhoe, M. M. (2007). Trade-offs between gaze and working memory use. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1352–1365. doi: 10.1037/0096-1523.33.6.1352
- Fei-Fei, L., Iyer, A., Koch, C., and Perona, P. (2007). What do we perceive in a glance of a real-world scene? *J. Vis.* 7, 1–29. doi: 10.1167/7.1.10
- Findlay, J. M., and Gilchrist, I. D. (2003). *Active Vision: The Psychology of Looking and Seeing*. Oxford: Oxford University Press.
- Furneaux, S., and Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proc. Biol. Sci.* 266, 2435–2440. doi: 10.1098/rspb.1999.0943
- Gegenfurtner, A., Lehtinen, E., and Säljö, R. (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educ. Psychol. Rev.* 23, 523–552. doi: 10.1007/s10648-011-9174-7
- Goldstone, S. (1968). Reaction time to onset and termination of lights and sounds. *Percept. Mot. Skills* 27, 1023–1029. doi: 10.2466/pms.1968.27.3f.1023
- Gorea, A. (1986). Temporal integration characteristics in spatial frequency identification. *Vision Res.* 26, 511–515. doi: 10.1016/0042-6989(86)90194-x
- Green, D. M., and von Gierke, S. M. (1984). Visual and auditory choice reaction times. *Acta Psychol.* 55, 231–247. doi: 10.1016/0001-6918(84)90043-x
- Greene, M. R., and Oliva, A. (2009). The briefest of glances: the time course of natural scene understanding. *Psychol. Sci.* 20, 464–472. doi: 10.1111/j.1467-9280.2009.02316.x
- Haider, H., and Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cogn. Psychol.* 30, 304–337. doi: 10.1006/cogp.1996.0009
- Hardiess, G., Gillner, S., and Mallot, H. A. (2008). Head and eye movements and the role of memory limitations in a visual search paradigm. *J. Vis.* 8, 1–13. doi: 10.1167/8.1.7
- Hershler, O., and Hochstein, S. (2009). The importance of being expert: Top-down attentional control in visual search with photographs. *Atten. Percept. Psychophys.* 71, 1478–1486. doi: 10.3758/app.71.7.1478
- Howard, C. J., and Holcombe, A. O. (2008). Tracking the changing features of multiple objects: progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Res.* 48, 1164–1180. doi: 10.1016/j.visres.2008.01.023
- Howard, C. J., Gilchrist, I. D., Troscianko, T., Behera, A., and Hogg, D. C. (2011). Task relevance predicts gaze in videos of real moving scenes. *Exp. Brain Res.* 214, 131–137. doi: 10.1007/s00221-011-2812-y
- Howard, C. J., Troscianko, T., and Gilchrist, I. D. (2010). Eye-response lags during a continuous monitoring task. *Psychon. Bull. Rev.* 17, 710–717. doi: 10.3758/pbr.17.5.710
- Land, M. (1999) Motion and vision: why animals move their eyes. *J. Comp. Physiol.* 185, 341–352. doi: 10.1007/s003590050393
- Ling, S., and Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Res.* 46, 1210–1220. doi: 10.1016/j.visres.2005.05.008
- Lo, S.-L., Howard, C. J., and Holcombe, A. O. (2012). Feature-based attentional interference revealed in perceptual errors and lags. *Vision Res.* 63, 20–33. doi: 10.1016/j.visres.2012.04.021
- Malcolm, G. L., and Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: evidence from eye movements. *J. Vis.* 9, 1–13. doi: 10.1167/9.11.8
- Melcher, D. (2006). Accumulation and persistence of memory for natural scenes. *J. Vis.* 6, 8–17. doi: 10.1167/6.1.2
- Reingold, E. M., Charness, N., Pomplun, M., and Stampe, D. M. (2001). Visual span in expert chess players: evidence from eye movements. *Psychol. Sci.* 12, 48–55. doi: 10.1111/1467-9280.00309
- Rousselet, G. A., Joubert, O. R., and Fabre-Thorpe, M. (2005). How long to get to the gist of real-world natural scenes? *Vis. Cogn.* 12, 852–877. doi: 10.1080/13506280444000553
- Schmidt, J., and Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Q. J. Exp. Psychol.* 62, 1904–1914. doi: 10.1080/17470210902853530
- Stainer, M. J., Tatler, B. W., and Scott-Brown, K. (2011). “Viewing multiplex displays: effects of continuity of content and spatial contiguity on fixation selection”, in *Abstracts of the 16th European Conference on Eye Movements*, Vol. 4, Journal of Eye Movement Research, eds F. Vitu, E. Castet and L. Goffart, Marseille, 255.
- Tatler, B., Gilchrist, I., and Rusted, J. (2003). The time course of abstract visual representation. *Perception* 32, 579–593. doi: 10.1068/p3396
- Thompson, J., and Parasuraman, R. (2012). Attention, biological motion, and action recognition. *Neuroimage* 59, 4–13. doi: 10.1016/j.neuroimage.2011.05.044
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., and Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception* 33, 87–101. doi: 10.1068/p3402
- Underwood, G., Chapman, P., Bowden, K., and Crundall, D. (2002). Visual search while driving: skill and awareness during inspection of the scene. *Transp. Res. Part F. Traffic Psychol. Behav.* 5, 87–97. doi: 10.1016/s1369-8478(02)00008-6
- Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., and Crundall, D. (2003). Visual attention while driving: sequence of eye fixations made by experienced and novice drivers. *Ergonomics* 46, 629–646. doi: 10.1080/0014013031000090116
- Watamaniuk, S., and Sekuler, R. (1992). Temporal and spatial integration in dynamic random-dot stimuli. *Vision Res.* 32, 2341–2347. doi: 10.1016/0042-6989(92)90097-3
- Yang, H., and Zelinsky, G. J. (2009). Visual search is guided to categorically defined targets. *Vision Res.* 49, 2095–2103. doi: 10.1016/j.visres.2009.05.017

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 12 April 2013; accepted: 18 July 2013; published online: 22 August 2013.

Citation: Howard CJ, Troscianko T, Gilchrist ID, Behera A and Hogg DC (2013) Suspiciousness perception in dynamic scenes: a comparison of CCTV operators and novices. *Front. Hum. Neurosci.* 7:441. doi: 10.3389/fnhum.2013.00441

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Howard, Troscianko, Gilchrist, Behera and Hogg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Looking for trouble: a description of oculomotor search strategies during live CCTV operation

Matthew J. Stainer^{1,2}, Kenneth C. Scott-Brown^{3*} and Benjamin W. Tatler¹

¹ Active Vision Lab, School of Psychology, University of Dundee, Dundee, Angus, UK

² Department of Optometry and Vision Science, University of Melbourne, Melbourne, VIC, Australia

³ Department of Psychology, University of Abertay Dundee, Dundee, Angus, UK

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

Evan F. Risko, University of Waterloo, Canada

Michele Furlan, Royal Holloway University of London, UK

Matteo Valsecchi, Justus-Liebig Universität Giessen, Germany

*Correspondence:

Kenneth C. Scott-Brown,
Department of Psychology,
University of Abertay Dundee, Bell
Street, Dundee, Angus, DD1 1HG,
UK

e-mail: k.c.scott-brown@
abertay.ac.uk

Recent research has begun to address how CCTV operators in the modern control room attempt to search for crime (e.g., Howard et al., 2011). However, an often-neglected element of the CCTV task is that the operators have at their disposal a multiplexed wall of scenes, and a single spot-monitor on which they can select any of these feeds for inspection. Here we examined how 2 trained CCTV operators used these sources of information to search from crime during a morning, afternoon, and night-time shift. We found that they spent surprisingly little time viewing the multiplex wall, instead preferentially spending most of their time searching on the single-scene spot-monitor. Such search must require a sophisticated understanding of the surveilled environment, as the operators must make their selection of which screen to view based on their prediction of where crime is likely to occur. This seems to be reflected in the difference in the screens that they selected to view at different times of the day. For example, night-clubs received close monitoring at night, but were seldom viewed in mid-morning. Such narrowing of search based on a contextual understanding of an environment is not a new idea (e.g., Torralba et al., 2006), and appears to contribute to operator's selection strategy. This research prompts new questions regarding the nature of representation that operators have of their environment, and how they might develop expectation-based search strategies to countermand the demands of the large influx of visual information. Future research should ensure not to neglect examination of operator behavior "in the wild" (Hutchins, 1995a), as such insights are difficult to gain from laboratory based paradigms alone.

Keywords: CCTV, surveillance, visual search, spatial selection, eye guidance, multiplex

INTRODUCTION

The task of the CCTV operator is to find and, if possible, prevent crime in public spaces. Research has shown that when asked to predict whether the events presented in a single video will turn violent, naive observers perform as well as trained CCTV operators (Troschianko et al., 2004; Grant and Williams, 2011). In control rooms, however, the CCTV operator is tasked with searching for such undefined targets across not one, but a vast number of screens displaying dynamic information from locations across a wide geographical area. For example, in a survey of 11 local authority and private security CCTV control rooms, operators were faced with a range of 27–520 cameras per operator, with up to 175 feeds displayed simultaneously across a bank of monitors (Gill and Spriggs, 2005; Gill et al., 2005). As such, the visually rich layout of the modern CCTV control room seems unnatural, complex, and ill-suited to the perceptual and cognitive constraints of the human operator (Scott-Brown and Cronin, 2008). It is well characterized that when searching for a target, the number of distractors that are present can dramatically influence search time (see review by Wolfe, 1998), including when a target's identity is not known (Rensink, 2000). Thus, performance skill in CCTV operation may be better characterized by their ability

to find a "target" scene (i.e., containing information for the task) amongst a large number of "distractor" scenes (e.g., see Howard et al., 2009).

MULTIPLE SCENE SURVEILLANCE

Tickner and Poulton (1973) demonstrated the behavioral costs when faced with increasing numbers of scenes in a surveillance-based task. These authors showed that when monitoring simultaneous feeds from cameras in a prison, the accuracy with which participants detected suspicious events was lower when the number of simultaneously-viewed camera feeds was high; with 83% for 4 monitors, 84% for 9 monitors, and 64% for 16 monitors. Wallace et al. (1997) examined observers' target detection across multiple scenes and found decreases in performance when increasing the number of town center scenes in the display. Correspondingly, this difficulty is reflected in CCTV operators confidence of multi-scene detection. When interviewed, 82% of CCTV operators interviewed only reported confidence with monitoring up to sixteen screens, with 50% reporting that they felt comfortable monitoring up to four screens simultaneously (Wallace and Diffley, 1998). This is considerably less than the number of screens that can be displayed in the modern control

room. In another study, Howard et al. (2011) presented participants with a series of quadruplex displayed CCTV clips and recorded their perceived suspiciousness of the video by means of a joystick. Participants moved the joystick forward to indicate their belief that an event was likely to happen. Viewers eye-gaze in these conditions, where multiple different video streams compete for attention, was allocated according to the relative suspiciousness of each video clip.

The overriding message from what is known about visual information load and visual search performance (e.g., Wolfe, 1998), and the performance in multiple-scene detection tasks (Tickner and Poulton, 1973; Howard et al., 2011) is that efficient search for crime among a large number of scenes is likely to be poor. However, while simultaneous display of a large number of camera feeds in multiplexes is an integral feature of CCTV control rooms, operators also have at their disposal individual spot monitors that can be used to selectively view the information from a single camera at a time (Figure 1). The selection of content in this way is an often-neglected element in studies of the CCTV task and it is important to characterize the relative use of multiplex and spot monitor for real surveillance situations.

Not only is it important to understand the manner in which the multiplex and spot monitor are used in surveillance, but it is also important to consider the different cognitive demands associated with the use of each of these display formats. In the multiplex, all visual content is displayed at one time. However, skilled and strategic use of the spot monitor relies on an understanding of the camera array and geographical area under surveillance, (Hillstrom et al., 2008). For example, tracking a suspect across an extended area of space requires selection of geographically adjacent cameras, even though they may neither be spatially adjacent in the multiplex nor visually continuous in content. Thus, selection on the spot monitor is not simply based on visual guidance, but rather by the representation of the environment or a mixture of the two. The multiplex and the spot monitor therefore present rather different challenges and opportunities for the operator and potentially rely on rather different underlying knowledge. Moreover, these two display types may be differentially suited to particular aspects of the surveillance task: the multiplex might be well suited for detecting unexpected or suspicious events as

these might occur in any of a number of different locations in the environment at any time. On the other hand, detailed information of unfolding events at a particular location might be better accessed via the spot monitor, where potential distraction from other camera feeds can be avoided.

One relatively unexplored aspect of the surveillance task is the extent to which the task demands vary over a 24-h period and how this impacts on operator behavior. For example, flash-point outbreaks of violence are a more prominent feature of the task at night than during the day in many urban settings (Felson and Poulson, 2003). Not only do the likely types of events differ over the 24-h period, but also the likely locations at which these events occur changes: night-clubs are a likely venue for fights at night, but not during the day. During a visual search task, when people are told the area of a scene that contains the target, performance is related to the size of that area, rather than the whole scene (Zelinsky and Schmidt, 2009). In the control room, and idea of where to look for different targets would likely serve to reduce the load of the observer. Similarly contextual understanding of scenes has been shown to influence where people search for items (Torralba et al., 2006). Given the intimate link between vision and task demands in real world activity (see Land and Tatler, 2009) it seems likely that visual strategies of CCTV operators will vary depending upon the time of day or night during which they are working.

The cognitive ethology (Kingstone et al., 2008; or ethnography e.g., Hutchins, 1995a; Hollan et al., 2000) approach to understanding how a system functions can provide otherwise hidden insights into how tasks are completed, such as how drivers navigate corners (Land and Lee, 1994). The purpose of this paper is to offer a first step toward understanding the nature of the surveillance task as it exists in a real CCTV control room. While in doing so we sacrifice some of the control which laboratory paradigms afford, such studies are essential to ensure that the questions we can ask in the lab are valid to the task (see also Hutchins, 1995a and Weibel et al., 2012 for a recent example including eye-tracking).

The first question we examined was to look at what extent the operators use the multiplex or the spot monitor. Research has addressed both single scene (e.g., Troscianko et al., 2004) and multiplex surveillance (e.g., Tickner and Poulton, 1973) viewing conditions, but a systematic analysis of their use in day-to-day Control Room operation has yet to be conducted. The second question that this paper addresses is to what extent is selection based on the monitoring task, and, by extension, to what extent is selection based on the viewing preference of the individual operator? If the task dictates spatial selection, then we would expect there to be larger differences in selection of content between shifts of operation. However, if selection is more related to the preferences of the individual operator, we would expect selection to be more different between the operators, and similar across different sessions.

MATERIALS AND METHODS

PARTICIPANTS

The observers were two trained CCTV Control Room operators from Tayside Police (now "Police Scotland") Control Room.

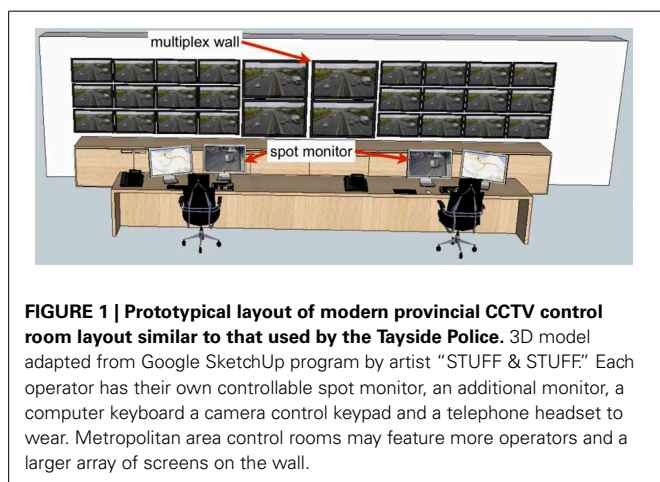


FIGURE 1 | Prototypical layout of modern provincial CCTV control room layout similar to that used by the Tayside Police. 3D model adapted from Google SketchUp program by artist "STUFF & STUFF." Each operator has their own controllable spot monitor, an additional monitor, a computer keyboard a camera control keypad and a telephone headset to wear. Metropolitan area control rooms may feature more operators and a larger array of screens on the wall.

Operator 1 had been working as an operator for approximately 10 years, whereas Operator 2 had been in the position for approximately 2 years (and was trained by Operator 1).

TAYSIDE CONTROL ROOM

Tayside Police Control Room receives live feeds from around 100 CCTV cameras in the Dundee City area at any one time. These camera feeds are displayed on a multiplexed bank of 47 CRT monitors (Figure 2). Several of the monitors are used to simultaneously display four camera feeds in split-screen (usually low-activity scenes such as car parks), and some automatically scroll through up to five different cameras, showing each one at a time for a period of several seconds. Many of the cameras are also on a set “walk” pattern, whereby they automatically pan across the area in a pre-set manner. Both operators that we recorded reported being able to comfortably see detail on the multiplex from their viewing position.

Operators in Tayside Control room work in teams of two (although they may occasionally be joined by a third person who will review footage on a separate station). This research was authorized by the Force Executive of Tayside Police.

EYE MOVEMENT RECORDING

Eye movements were recorded using a lightweight Positive Science LLC mobile eye tracking system built by Jason Babcock (Babcock and Pelz, 2004). The system samples eye position at a 30 Hz and creates a video overlay of the scene viewed from a first person perspective with a gaze-cursor cross. Two cameras were mounted on a spectacle frame, simultaneously recording the scene and the observer's eye. The key benefit of this system is its unobtrusive qualities. Thanks to its small visual footprint and low-weight construction, operators can enjoy full freedom of movement in their normal seated position. As viewing behavior may be influenced by the process of wearing an eye-tracker (e.g., an “eye-tracker awareness”; Risko and Kingstone, 2011), operators were given no instruction other than to carry out their task as usual to attempt to minimize experimenter effects.

The video from the cameras was captured live into the Yabus software package (version 2.2.2) on a MacBook Pro (4 GB Memory, 2.4 GHz Intel Core 2 Duo), where eye position was estimated based on detection of the pupil (with accuracy within a degree of visual angle). Observers calibrated live using a 9-point

grid made up of the corners of monitors on the data wall, and the four corners around their spot monitor.

Data were exported as videos from the scene camera overlaid with eye position (Figure 3). The videos were then hand-coded to extract where the operators were looking throughout each session in terms of the type of display (multiplex and spot-monitor), and the camera feed that was shown on that display. Data during blinks were excluded from analysis.

PROCEDURE

Recordings were made during live system operation from each observer at each of their three shifts of work (afternoon, morning, and late night). Each recording session for each operator was 15 min in duration. Care was taken to ensure fitting and removal of the equipment from the operator was performed at convenient times within the surveillance task so as not to interrupt actions *en train*. Operators were told that they could remove the glasses at any time if they felt it was hindering their work (although neither chose to at any point).

ANALYSIS

Our approach to examining the question of how operators search for crime is not a traditional experimental design, but rather an observational approach. There are potential issues with over-generalizing the data from such observations (particularly given the low number of operators). However, our aim is to describe behavior as it occurs. Thus, we applied traditional quantitative techniques of analysis to attempt to quantify this behavior, and describe the operators' use of the multiplex and the spot monitor in their search for crime. As such, some data presented are simply numerical (such as the number of cameras that an operator viewed on a particular session).

When examining differences between operators based on continuous variables, we used linear mixed-effect modeling (for example, to examine the difference in scanning time per scene between operators). Linear-mixed effects models have become increasingly used to examine non-normally distributed data (e.g., see Druker and Anderson, 2010). They allow for modeling of fixed factors, and random factors, with all data included (rather than condensing the data to a single mean). Thus, it considers the variance within a random factor (such as participant), as well as the variance between fixed factors. However, here we consider operator as a fixed factor. Conventionally the fixed factor in an analysis must be repeatable (Baayen, 2007, p263). However, we include



FIGURE 2 | Tayside Police control room.

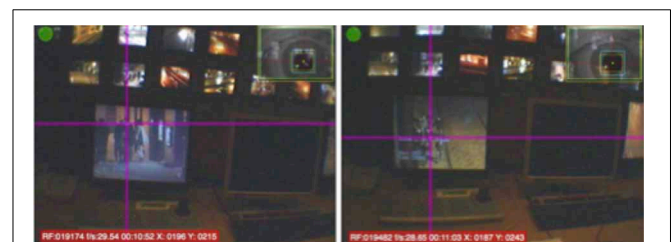


FIGURE 3 | Examples of eye gaze videos with gaze position crosshair overlaid.

operator here as a fixed factor, as we do not intend to generalize our data beyond differences between our operators (and simply to try to quantify if they *were* different).

Here, we analyzed the data using the *lme4* (Bates, 2005) and *languageR* (Baayen, 2007) packages in *R* (R Development Core Team, 2009). We follow the reporting style of Druker and Anderson (2010), who used similar modeling, to report the mean difference between conditions with highest 95% posterior density intervals from Markov Chain Monte Carlo mean estimates, with approximated *p* values generated with the *pval.fnc* function (Baayen, 2007; Baayen et al., 2008).

When looking at categorical differences between operators, we employed Kullback-Leibler divergence analysis (for example, to analyze whether there was a difference in the cameras selected between operators, and between sessions). Kullback-Leibler divergence is an information theoretic measure that allows us to quantify the difference between two probability distributions in terms of the number of bits of code that is required to describe one distribution based on another. We present these probability distributions in graph form, with camera number being a categorical factor, plotted against probability of fixation. Thus, the Kullback-Leibler divergence score can be used as a measure of the difference between two categorical distributions (for similar use see Tatler et al., 2005). This allows us to quantify whether differences in selection are greater between shifts of operation, or between operators, with higher scores representing larger differences.

RESULTS

Across all sessions, we found that operators spent the majority of their time selecting content on their individual spot monitor (>90% across both observers in all sessions; Table 1). Operator 1 did not use the multiplex at all in the morning, or evening sessions, with the highest proportion of time spent on the multiplex being the afternoon session for both operators.

SPOT MONITOR SCANNING

We looked at four principle measures of spot monitor use that are summarized in Figure 4. First, Figure 4A reveals that in the afternoon and morning sessions, Operator 1 viewed around half the total number of scenes compared to Operator 2. However, in the night session Operator 1 viewed more scenes in total than Operator 2 (although this total was less than the number of scenes viewed by Operator 2 other two sessions).

Per scan (a viewing session on the spot monitor that was uninterrupted by looks at the multiplex), Operator 2 was relatively

consistent, viewing around 2–3 scenes between looks to the multiplex in all three shifts (Figure 4B). However, as Operator 1 did not view the multiplex at all in the morning and evening recording sessions, they viewed more scenes per scan than Operator 2. In the only session that Operator 1 did use the multiplex (the afternoon session), the number of scenes per scan was similar to Operator 2 (2–3 scenes). Correspondingly, Figure 4C shows that the Operator 1 had longer periods of spot monitor scanning than Operator 2 in all sessions.

Finally, we looked at how long Operators would view each scene before selecting to view content from a different camera. To examine scanning time per scene, a linear mixed effect model was carried out with operator included as a fixed factor, and session included as a random factor. Figure 4D demonstrates that Operator 2 inspected each scene for significantly less time than Operator 1 (Markov-Chain Monte Carlo (MCMC) mean difference = −35.19 s, 95% CI = −49.86 to −21.55 s, $p < 0.0001$).

SPOT MONITOR SELECTION

The amount of time that operators spent on each selected scene viewed on the spot monitor across the three recording sessions is illustrated in Figure 5. Operators' selection of content was most similar between the afternoon and night sessions (Figure 6 center bar of panels 1 and 2). Operators showed the greatest difference in the scenes that they chose to view on the spot monitor in the morning compared to the night shift (right bar of panels 1 and 2). The scenes that were selected at night were most similar between operators (right bar of panel 3), and least similar in the afternoon.

MULTIPLEX SCANNING

As discussed previously, Operator 1 used their spot monitor for the entire morning and night session. Figure 7A reveals that Operator 2 viewed just over 3 times as many scenes in the afternoon session compared to Operator 1. Operator 2 also viewed more scenes per scan (Figure 7B), and had longer periods of multiplex scanning (Figure 7C). However, Figure 7D reveals that when Operator 1 did look at scenes on the multiplex, the operator spent more time on average viewing each scene before moving to another.

MULTIPLEX SELECTION

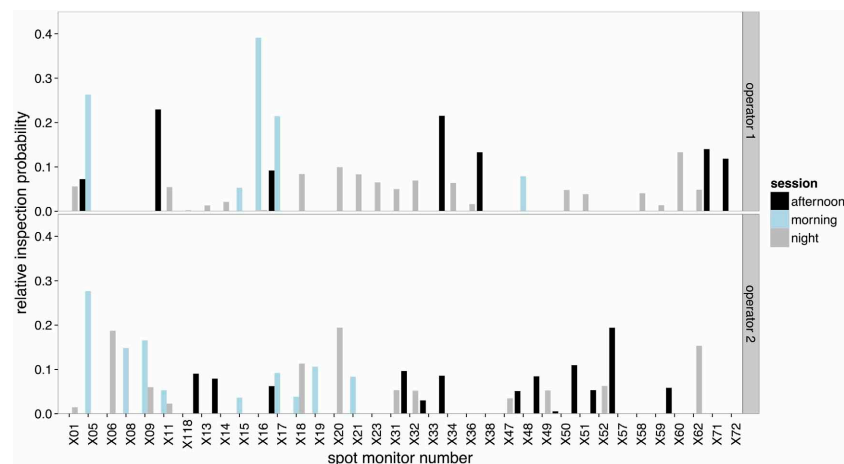
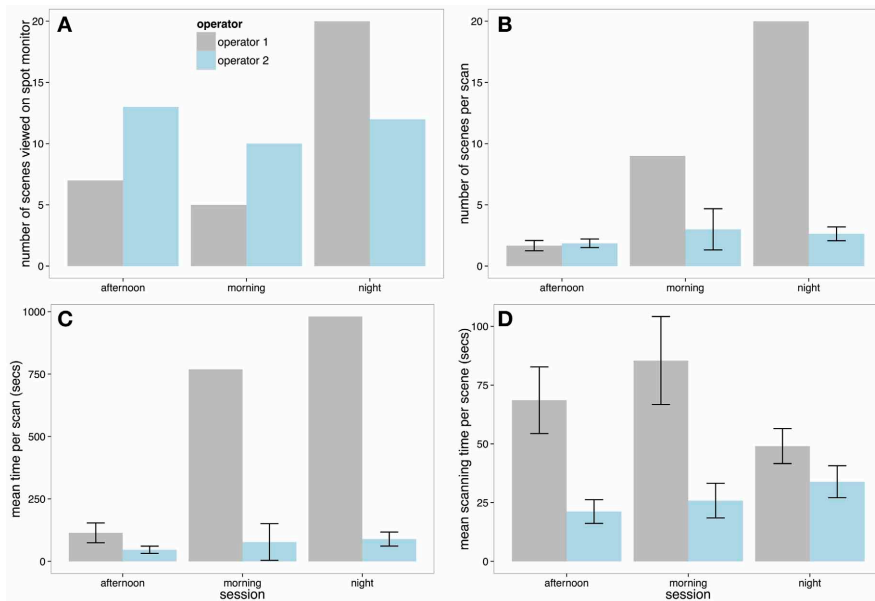
The distributions of time spent viewing scenes on the multiplex can be seen in Figure 8. As Operator 1 did not use the multiplex on either the morning or afternoon session, only selection by Operator 2 was examined using Kullback-Leibler divergence. Figure 9 shows that there was much less variance in selection on the multiplex between sessions compared to the content viewed on the spot monitor (which yielded higher Kullback-Leibler scores). However, when compared across sessions, selection followed a similar pattern as on the spot monitor. Selection of content was most similar between the afternoon and night sessions.

DISCUSSION

In what we believe to be the first study of visual strategies for expert CCTV surveillance in a public space control room under normal working conditions, we report the results of a mobile eye-tracking study of CCTV operator performance during day

Table 1 | The proportion of time spent by each operator viewing their spot monitor, and the multiplex.

		Afternoon (%)	Morning (%)	Evening (%)
Operator 1	Spot monitor	98.5	100	100
	Multiplex	1.5		
Operator 2	Spot monitor	91.51	97.57	98.47
	Multiplex	8.31	4.43	1.53



and night shift team-based surveillance. Spot monitor scanning and selection was compared with multiplex scanning and selection data along with a comparison of inter-operator differences in screen inspections.

For the operators we studied, spot monitor observation took up more than 90% of inspection time in the control room during the periods of observation (afternoon, morning, and evening). The data demonstrate that during our recording spatial selection in the control room differed dramatically both between operators, and between different shifts of operation. For example, Operator

1 spent more time viewing content on the spot monitor than Operator 2, and spent longer on each scene before transitioning. These differences between operators may reflect different idiosyncratic styles for surveillance or the differing experience of the two operators. However, the operators work as a team and these differences may reflect the different roles that each operator took in their collaborative effort. For example, Operator 1 might take the role of monitoring the night clubs, while Operator 2 monitors at the suburbs. Such distribution of cognition has been previously demonstrated, for example, between pilots in the cockpit of

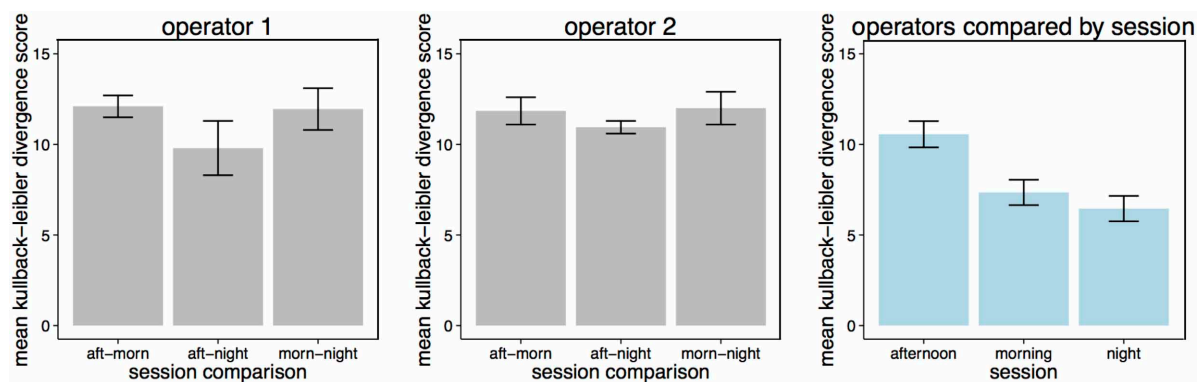


FIGURE 6 | Kullback-Leibler divergence score in screens selected for viewing on the spot monitor by session for Operator 1 (panel 1), Operator 2 (panel 2) and across all sessions (panel 3). \pm SE are included,

and represent that Kullback-Leibler divergence analysis gives two scores for each comparison (the probability of distribution A/B, and the probability of distribution B/A).

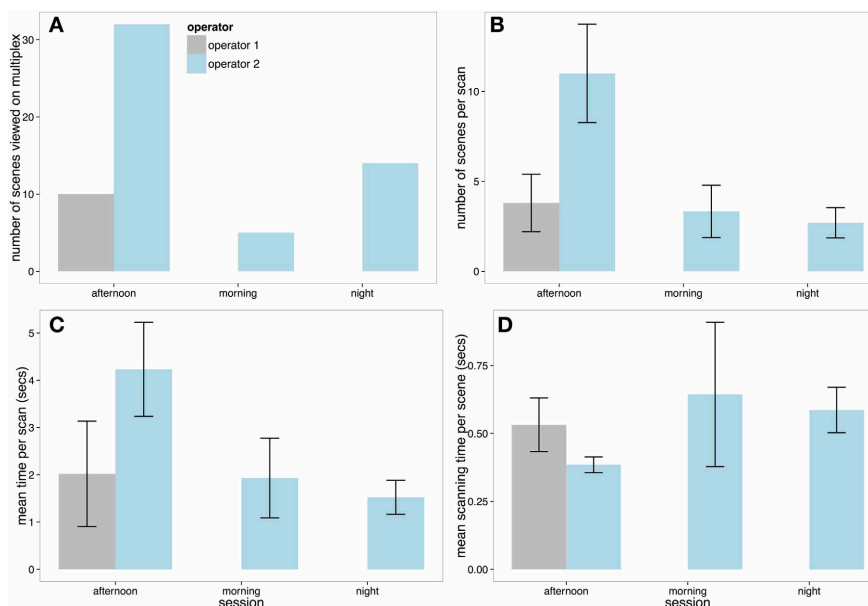


FIGURE 7 | (A) Total number of scenes viewed on the multiplex on each session by each operator. (B) The number of scenes selected by each operator per scan. (C) Mean length of each multiplex scanning session and (D) Mean scanning time per scene on the multiplex (with \pm SE).

an airplane (Hutchins, 1995b). While the question of how operators work together to efficiently detect crime was not the aim of this study, this would likely be an informative and interesting direction for future research.

Despite the data showing that during three 15-min recording sessions the operators spent little time viewing content on the multiplex, when operators did use the multiplex, they were more similar to each other in what they chose to view compared to their selections for inspection on the spot monitor. Short scans of the multiplex lasting approximately 1–4 s punctuate the longer spot monitor views, and inspection times for individual scenes are extremely short when viewed on the multiplex. Thus, it appears that anything worth further inspection is probably brought to the spot monitor, and multiplex viewing may be

used primarily to help identify content that should receive more detailed scrutiny. Content selection in the multiplex appears most similar in afternoon and night conditions.

These findings indicate that approaches to understanding surveillance that are based solely on multiplex detection (Tickner and Poulton, 1973) or single screen detection (Troschianko et al., 2004) may provide insights into aspects of the task. However, given the dynamic interplay between multiplex viewing and selecting single camera feeds for further inspection, these two modes of viewing need to be considered together. Moreover, single screen viewing is a very active process in which content from different cameras is actively selected, with new camera feeds being selected on average every 26.94 (Operator 2) to 62.44 (Operator 1) s while using the spot monitor. Selection of

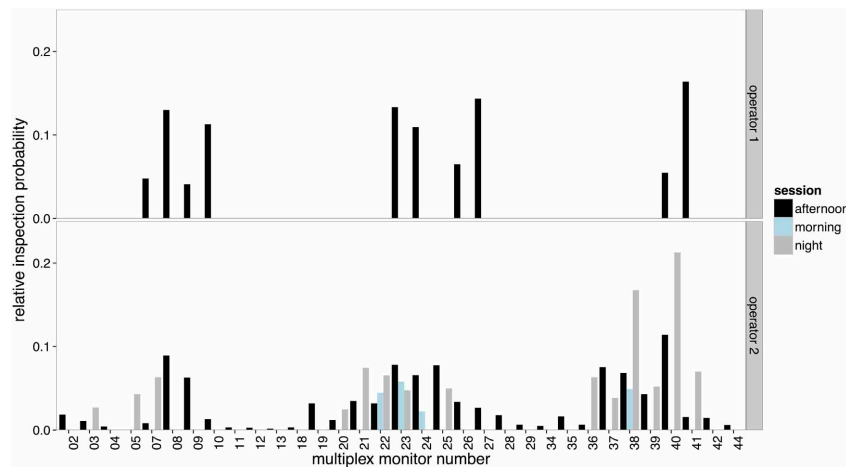


FIGURE 8 | The relative inspection probabilities for each scene selected on the multiplex by Operator 1 (upper panel) and Operator 2 (lower panel).

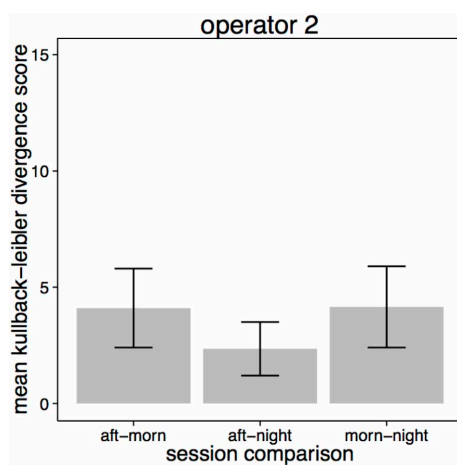


FIGURE 9 | Kullback-Leibler divergence score in screens selected for viewing on the multiplex by session for Operator 2.

content during spot monitor use necessarily reflects considerable use of the internal representation of the surveilled environment, including an understanding of the camera locations in external space.

STRATEGIES IN SEARCHING FOR CRIME

When searching for crime, we found that the CCTV operators spent very little time searching the multiplex. In accordance to this finding, operators of multiplex systems reported low confidence in their ability to monitor several scenes (Wallace and Diffley, 1998). This would be entirely consistent with what is understood about search of complex displays (e.g., see Wolfe, 1998). Increasing the amount of visual information in a display increases search time (with visual information measured in several methods; Rosenholtz et al., 2007; Henderson et al., 2009; Beck et al., 2010; Bi et al., 2010; Wolfe et al., 2011; Asher et al., 2013).

Given the likelihood of a bottleneck of attention at some-point in the visual system (for example, see limits on the number of objects we can simultaneously track; Alvarez and Franconeri, 2007), the multiplex might present a daunting task to the visual system. Performance drops have been shown at four screens (Tickner and Poulton, 1973, or Rousselet et al., 2004), which was less than 1/10th of the screens in the multiplex of the control room examined here.

One way that operators might effectively be able to increase confidence is to use the spot monitor (i.e., reduce the task to a single scene load). If operators conduct the majority of work on their single spot monitor, it is important that they select the appropriate scenes to view. While previous research has found no effect of training on single scene detection tasks (such as Troscianko et al., 2004), it may be that expertise in the control room serves to guide operators' search for crime within the large number of scenes that they could potentially select and view. Accordingly, Howard et al. (2010) demonstrated that the difference between experts and novices watching a five-a-side football match is that experts look at the most informative locations earlier than novices. In the surveillance context, we found that operators appear to select content differently at different times of day and this seems likely to be based on both their knowledge of the environment and their experience of where events are likely to occur at different times of the day.

It is important to consider how operators are able to select a subset of appropriate content from the large array of camera feeds available. It is possible that this is based on reactive selection to events unfolding in each camera feed. However, the proactive nature of surveillance and the often subtle events that are selected for detailed monitoring suggests that the selection processes are likely to be strategic, based on prior knowledge and expectation. One plausible possibility is that operators have an understanding of the likely locations at which events will occur at different times and that they use this to constrain much of their surveillance effort to the cameras that depict these locations. In this way, suspicious events will be monitored primarily within expected

locations in the surveilled environment. This suggestion is similar to the contextual selection that has been demonstrated in scene viewing paradigms, where observers appear to combine expectations of where things are likely to be in the world with low level feature information (Torralba et al., 2006; Ehinger et al., 2009). In such paradigms, it has been shown that observers primarily search regions in which targets are expected to occur, with search time being related to the area the observer has to search, rather than the whole display (Zelinsky and Schmidt, 2009). Some cameras facing night-clubs (e.g., feed X20 and X62) were not viewed at all in the morning and afternoon session, but made up a large proportion of the night-time surveillance. How operators develop their criteria for selecting appropriate content is a question that further research should seek to address.

We propose four potential ways that expectation might develop: First, expectation may simply be based on general associations of social factors (e.g., areas associated with drug use are more likely to be high violence areas, Lum, 2011). Second, expectation might be built up via reinforcement, as operators successfully experience or detect events in certain scenes (similar to the development of spatial bias in visual search, e.g., Carpenter and Williams, 1995). Third, it may be based on how the amount of activity (and hence content and motion within the camera feeds, e.g., see Howard and Holcombe, 2010) changes throughout the day. There are simply more people around night-clubs at night than anywhere else. Fourth, strategic selections may arise as a result of explicit instruction about where to look and when during operator training (e.g., Wallace and Diffley, 1998, Appendix A). We might speculate that the fourth possibility does not account for aspects of our findings because the two operators differed in the scenes they viewed, however, as previously suggested this difference might be an active choice for efficient collaboration of efforts across the control room.

CONCLUSIONS

Research has shown that when observers attempt to detect criminal activity in one scene, untrained observers are as good as trained CCTV operators (Trosianko et al., 2004; Grant and

Williams, 2011). However, this situation only captures one part of the CCTV operator's task. First, operators have to correctly select the scene to view from a large number of possibilities. As such, the task of CCTV operation is not simply a case of looking at the right place at the right time, but rather of looking at the right place at the right time *in the right scene*. To complete this task, we found that two trained CCTV operators spent more time searching for crime using a single scene spot monitor, rather than the multiplex data wall, despite the latter giving the operator more information at one time. This may, in part, reflect the difficulty of search across large amounts of visual information (e.g., Wolfe, 1998 among others). However, to be able to search with the spot monitor, operators must select screens based on their representation of the surveilled world. Moreover, this understanding of the environment seems to incorporate the monitoring demands associated with different shifts of operation, with operators selecting different screens at day compared to night, for example. This may reflect the locations of high event likelihood being different at night, compared to during the morning, which would be consistent with using contextual understanding to guide visual search to areas likely to contain a target (such as Torralba et al., 2006).

Using cognitive ethology, we can gain a more comprehensive, ecologically valid idea of how cognition functions "in the wild." We echo the sentiments of Kingstone et al. (2008) that observation of naturally occurring behavior can provide an essential complement to laboratory-based studies in generating valid hypotheses and questions, as neither alone can provide a complete picture of complex cognitive tasks such as CCTV operation.

ETHICS STATEMENT

This research was carried out in accordance with, and approval of the University of Dundee Ethics Committee.

ACKNOWLEDGMENTS

The authors would like to thank Tayside Police (now Police Scotland) and their CCTV Control Room staff for their collaboration and participation in this research. Thanks also to Sharon Scrafton for comments on the manuscript.

REFERENCES

- Alvarez, G. A., and Franconeri, S. L. (2007). How many objects can you track?: evidence for a resource-limited attentive tracking mechanism. *J. Vis.* 7, 1–10. doi: 10.1167/7.13.14
- Asher, M. F., Tolhurst, D. J., Trosianko, T., and Gilchrist, I. D. (2013). Regional effects of clutter on human target detection performance. *J. Vis.* 13, 1–15. doi: 10.1167/13.5.25
- Baayen, R. H. (2007). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.
- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005
- Babcock, J. S., and Pelz, J. B. (2004). "Building a lightweight eyetracking headgear," in *Presented at the ETTRA '04*, (San Antonio, Texas: ACM), 95–100.
- Bates, D. (2005). *Fitting Linear Mixed Models in R*, Madison: R news.
- Beck, M. R., Lohrenz, M. C., and Trafton, J. G. (2010). Measuring search efficiency in complex visual search tasks: global and local clutter. *J. Exp. Psychol. Appl.* 16, 238–250. doi: 10.1037/a0019633
- Bi, X., Bae, S. H., and Balakrishnan, R. (2010). "Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection," in *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, (Atlanta, GA), 65–74.
- Carpenter, R. H. S., and Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377, 59–62. doi: 10.1038/377059a0
- Druker, M., and Anderson, B. (2010). Spatial probability aids visual stimulus discrimination. *Front. Hum. Neurosci.* 4:63. doi: 10.3389/fnhum.2010.00063
- Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A., and Oliva, A. (2009). Modelling search for people in 900 scenes: a combined source model of eye guidance. *Vis. Cogn.* 17, 945–978. doi: 10.1080/13506280902834720
- Felson, M., and Poulsen, E. (2003). Simple indicators of crime by time of day. *Int. J. Forecast.* 19, 595–601. doi: 10.1016/S0169-2070(03)00093-1
- Gill, M., Spriggs, A., Allen, J., Hemming, M., Jessiman, P., Kara, D., et al. (2005). "Control room operation: findings from control room observations," in *Home Office Online Report*, (London).
- Gill, M., and Spriggs, A. (2005). *Assessing the impact of CCTV*, (London: Home Office Research, Development and Statistics Directorate), 1–176.
- Grant, D., and Williams, D. (2011). The importance of perceiving social contexts when predicting crime and antisocial behaviour in CCTV images. *Legal Criminol.*

- Psychol.* 16, 307–322. doi: 10.1348/135532510X512665
- Henderson, J. M., Chanceaux, M., and Smith, T. J. (2009). The influence of clutter on real-world scene search: evidence from search efficiency and eye movements. *J. Vis.* 9, 1–8. doi: 10.1167/9.1.32
- Hillstrom, A., Hope, L., and Nee, N. (2008). “Applying psychological science to the CCTV review process: a review of cognitive and ergonomic literature,” in *Project Report* (London: HMSO).
- Hollan, J., Hutchins, E., and Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput. Hum. Interact.* 7, 174–196. doi: 10.1145/353485.353487
- Howard, C., Troscianko, T., Gilchrist, I. D., Behera, A., and Hogg, D. C. (2009). “Searching for threat: factors determining performance during CCTV monitoring,” in *Human Factors, Security and Safety*, eds D. de Waard, J. Godthelp, F. L. Kooi and K. A. Brookhuis (Maastricht: Shaker Publishing), 1–7.
- Howard, C. J., Gilchrist, I. D., Troscianko, T., Behera, A., and Hogg, D. C. (2011). Task relevance predicts gaze in videos of real moving scenes. *Exp. Brain Res.* 214, 131–137. doi: 10.1007/s00221-011-2812-y
- Howard, C. J., and Holcombe, A. O. (2010). Unexpected changes in direction of motion attract attention. *Atten. Percept. Psychophys.* 72, 2087–2095. doi: 10.3758/APP.72.8.2087
- Howard, C. J., Troscianko, T., and Gilchrist, I. D. (2010). Eye-response lags during a continuous monitoring task. *Psychon. Bull. Rev.* 17, 710–717. doi: 10.3758/PBR.17.5.710
- Hutchins, E. (1995a). *Cognition in the Wild*. Cambridge, MA: The MIT Press.
- Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cogn. Sci.* 19, 265–288. doi: 10.1207/s15516709cog1903_1
- Kingstone, A. A., Smilek, D., and Eastwood, J. D. (2008). Cognitive ethology: a new approach for studying human cognition. *Br. J. Psychol.* 99, 317–340. doi: 10.1348/000712607X251243
- Land, M. F., and Lee, D. N. (1994). Where we look when we steer. *Nature* 369, 742–744. doi: 10.1038/369742a0
- Land, M. F., and Tatler, B. W. (2009). *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198570943.001.0001
- Lum, C. (2011). Violence, drug markets and racial composition: challenging stereotypes through spatial analysis. *Urban Stud.* 48, 2715–2732. doi: 10.1177/0042098010388953
- R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rensink, R. A. (2000). Visual search for change: a probe into the nature of attentional processing. *Vis. Cogn.* 7, 345–376. doi: 10.1080/135062800394847
- Risko, E. F., and Kingstone, A. A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Atten. Percept. Psychophys.* 73, 291–296. doi: 10.3758/s13414-010-0042-1
- Rosenholtz, R., Li, Y., and Nakano, L. (2007). Measuring visual clutter. *J. Vis.* 7, 17–22. doi: 10.1167/7.2.17
- Rousselet, G. A., Thorpe, S. J., and Fabre-Thorpe, M. (2004). Processing of one, two or four natural scenes in humans: the limits of parallelism. *Vision Res.* 44, 877–894. doi: 10.1016/j.visres.11.014
- Scott-Brown, K. C., and Cronin, P. D. J. (2008). Detect the unexpected: a science for surveillance. *Policing* 31, 395–414. doi: 10.1108/13639510810895777
- Tatler, B. W., Baddeley, R. J., and Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Res.* 45, 643–659. doi: 10.1016/j.visres.2004.09.017
- Tickner, A. H., and Poulton, E. C. (1973). Monitoring up to 16 synthetic television pictures showing a great deal of movement. *Ergonomics* 16, 381–401. doi: 10.1080/00140137308924529
- Torralba, A., Oliva, A., Castelano, M. S., and Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol. Rev.* 113, 766–786. doi: 10.1037/0033-295X.113.4.766
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., and Wilson, A. (2004). What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception* 33, 87–101. doi: 10.1068/p3402
- Wallace, E., and Diffley, C. (1998). *CCTV: Making It Work*. CCTV Control Room Ergonomics. Sandridge, St Albans: Police Scientific Development Branch.
- Wallace, E., Diffley, C., Baines, E., and Aldridge, J. (1997). “Ergonomic design considerations for public area CCTV safety and security applications,” in *Presented at the International Ergonomics Association Congress*, (Sandridge, St Albans).
- Weibel, N., Fouse, A., Emmenegger, C., Kimmich, S., and Hutchins, E. (2012). “Let’s look at the cockpit: exploring mobile eye-tracking for observational research on the flight deck,” in *Proceedings of the Symposium on Eye Tracking Research and Applications*, (Santa Barbara, CA: ACM), 107–114.
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychol. Sci.* 9, 33–39. doi: 10.1111/1467-9280.00006
- Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., and Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Atten. Percept. Psychophys.* 73, 1650–1671. doi: 10.3758/s13414-011-0153-3
- Zelinsky, G. J., and Schmidt, J. (2009). An effect of referential scene constraint on search implies scene segmentation. *Vis. Cogn.* 17, 1004–1028. doi: 10.1080/13506280902764315

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 27 May 2013; accepted: 07 September 2013; published online: 30 September 2013.

Citation: Stainer MJ, Scott-Brown KC and Tatler BW (2013) Looking for trouble: a description of oculomotor search strategies during live CCTV operation. *Front. Hum. Neurosci.* 7:615. doi: 10.3389/fnhum.2013.00615

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Stainer, Scott-Brown and Tatler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Cues derived from facial appearance in security-related contexts: a biological and socio-cognitive framework

Christopher D. Watkins*

Division of Psychology, School of Social and Health Sciences, University of Abertay, Dundee, UK

*Correspondence: c.watkins@abertay.ac.uk

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

S. Craig Roberts, University of Stirling, UK

Failures in the security process can have profound costs for both the individual and organizations (e.g., fraud costs the British economy approximately £72 billion; NFA, 2012). A biological and socio-cognitive framework may enhance our understanding of the security process, as the two perspectives collectively acknowledge that (i) competition for resources is/was an important factor in human social behavior and evolution (e.g., Bowles, 2009) and (ii) individuals differ in the ways in which they interpret information given their own traits and circumstances. Both levels of explanation (Mayr, 1963; Tinbergen, 1963) could generate novel hypotheses. For example, proximate-level explanations may clarify *how* resources are defended and extorted, and the cognitive processes underlying the “chess game” between gatekeepers and “gate crashers.” Ultimate-level explanations may clarify *why* some individuals are more likely than others to succeed at securing or gaining access to resources and whether certain security-related outcomes can be reliably predicted given specific contexts or ecological conditions.

DECISION-MAKING UNDER UNCERTAINTY

Humans make many decisions (consciously or otherwise) based on uncertain outcomes. Error management theory proposes that cognition has evolved so that when faced with two alternate strategies, we pick the strategy that would result in the least-costly errors (Haselton and Buss, 2000; Haselton and Nettle, 2006). This “cost-benefit” approach to decision-making is of value in describing the nature of human conflict. For example, sex differences in aggression are said to reflect the greater net “pay-off” to the reproductive

fitness of males who engage in potentially risky competition for resources (see Archer, 2009 for discussion). Local differences in income inequality are also an important predictor of violent male-male competition (Daly et al., 2001), and may “pay-off” if harsh environments promote risky behavior in light of future economic uncertainty (e.g., Wilson and Daly, 2006). Research on environmental differences in behavior could provide an evidence-base for effective investment in crime-prevention (e.g., examining the local distribution of CCTV cameras), given that current strategies may be sub-optimal (see, e.g., Webster, 2009).

Psychological mechanisms also play an important role in conflict. Overconfidence, the illusion of thinking you are better than you are, is an important cause of warfare (Johnson et al., 2006) and is more likely to evolve in contexts where the perceived benefits of competition outweigh their perceived costs (Johnson and Fowler, 2011). Indeed, this is neatly illustrated by George Bush’s “mission accomplished” speech aboard the USS Abraham Lincoln in May 2003. Thus, given knowledge of context, ultimate levels of explanation can aid our understanding of the maladaptive practice of warfare.

Error management theories suggest that we will tolerate “false alarms” in circumstances where they are much less costly than having no alarm in place when really needed. Given the costs of security failure, to what extent will a gatekeeper tolerate false alarms (e.g., risk making a false conviction) given their own personality or immediate environment? These issues are clearly very current, as commentators debate the “trade off” between security and civil liberties. Indeed, the extent to which human error accounts

for the false conviction of suspects (e.g., the innocence project; see Jenkins and Burton, 2011 for related discussion) is a neat illustration that demonstrating scientific evidence for a given behavior (e.g., cognitive errors/biases) is not the same as morally-endorsing that behavior (see, Greene, 2003).

Contextual cues may alter the nature of the trade-off between the perceived costs and benefits of identifying, controlling and monitoring perceived threats to the security of one’s resources. Across species, evidence for the “winner effect” suggests that future decisions to engage in competition are modulated by recent experience such that winners more likely to escalate a future confrontation (even with a rival of higher rank than themselves) and losers are more likely to withdraw from future confrontation [reviewed in Hsu et al. (2006)]. Recent work suggests that confrontation outcomes modulate competition-related perceptions in men in a similar way as it appears to do in other species. Men who are primed to imagine having lost a confrontation find facial cues of dominance in other men to be more salient than men who are primed to imagine having won a confrontation (Watkins and Jones, 2012). These effects may be adaptive if they function as a compensatory response to the increased vulnerability of loss of resources in light of recent experience, and are consistent with other work which demonstrates how a lack of power can predict general inhibition in behavior and greater orientation toward threat [reviewed in Keltner et al. (2003)]. Contextual factors relevant to competition may predict systematic variation in judgments toward other cues of threat, such as facial expression or movement. Differential treatment toward others based

on their appearance suggests an underlying biological basis to social interactions that might be important for effective competition.

A BIOLOGICAL BASIS TO SOCIAL JUDGMENTS

Information provided by the face plays an important role in social interaction (Bruce and Young, 1986), and the categorization (e.g., Hugenberg and Bodenhausen, 2003; Mason et al., 2006) and identification (e.g., Hancock et al., 2000) of other people. We appear to be very quick to make our mind up about the character of an individual based on his or her facial appearance; trait judgments of faces made after just 100 ms of exposure are highly correlated with judgments made at longer exposure intervals (Willis and Todorov, 2006). A principal components analysis of trait judgments made toward faces revealed that differences in human face shape can be modeled on two primary dimensions, reflecting the extent to which an individual appears *intent* on causing harm to others (their perceived trustworthiness) and the extent to which an individual appears *capable* of causing harm to others [their perceived dominance; (Oosterhof and Todorov, 2008)]. Rapid judgments of traits that are important for personal safety are functionally adaptive if the costs of erring on the side of optimism are much greater than the costs of erring on the side of caution—the speed of social judgments at zero acquaintance may be more important than their accuracy [reviewed in Todorov et al. (2008)]. For example, given the potential costs of competition (Manson and Wrangham, 1991; Bowles, 2009), a rapid attribution of “threat” that turns out to be inaccurate is much less costly than an attribution of “no threat” that turns out to be inaccurate.

Perceptions of dominance and trust appear to have an underlying biological basis and are of obvious relevance to security scientists. Although the relationship between hormones and facial appearance is complex (see Pound et al., 2009), sex differences in the human face are thought to depend on exposure to gonadal steroids (see Puts et al., 2012 for discussion). Masculine physical characteristics in men are positively correlated with their perceived dominance (e.g., Perrett et al., 1998;

Puts et al., 2006; Jones et al., 2010) and untrustworthiness (e.g. Perrett et al., 1998; Boothroyd et al., 2007). These attributions toward physically dominant individuals may have a “kernel of truth.” For example, physically dominant men are more likely to endorse the use of physical force to resolve conflict (Sell et al., 2009), are more aggressive in certain contexts (Carré and McCormick, 2008; Carré et al., 2009) and are less likely to share resources fairly with others (Stirratt and Perrett, 2010; Price et al., 2011) than their less dominant peers. From a biological perspective, physically dominant individuals should express less concern for the welfare of others than their less dominant peers, given that dominant individuals are better-placed to exploit or forcefully acquire resources with impunity (Sell et al., 2009; Puts, 2010; Stirratt and Perrett, 2010). Indeed, the costs of conflict are rarely symmetric between two parties (Maynard Smith and Price, 1973), and recent work suggests that facial cues of dominance in potential rivals are more salient to those who are less well-equipped to “offset” these costs (Watkins et al., 2010a,b). Systematic variation in dominance perceptions may be adaptive if it functions to minimize the costs of conflict in light of the perceiver’s own dominance (Watkins et al., 2010a,b; Watkins and Jones, 2012). Exploring the extent to which the gatekeeper’s own dominance predicts security-related outcomes may be a practical application for this line of reasoning.

Other aspects of facial appearance may predict trusting behavior in the exchange of resources. While attractive individuals are more likely to be trusted in economic exchanges (Solnick and Schweitzer, 1999; Hancock and DeBruine, 2003; Wilson and Eckel, 2006; Andreoni and Petrie, 2008), particularly attractive individuals are more likely than their less attractive peers to “shift” toward more trusting behavior when they believe that others’ have the opportunity to take their appearance into account (Smith et al., 2009). Given that attractiveness is associated with a suite of positive attributions (Langlois et al., 2000) and that a positive reputation can benefit one’s reproductive fitness (Fehr, 2004; Nowak and Sigmund, 2005), strategic economic behavior in light of a beautiful appearance is to be expected,

particularly given the severe penalties incurred when individuals are perceived as having used their looks for nefarious purposes (e.g., in cases of fraud; see Mazzella and Feingold, 1994 for a meta-analytic review; see also Wilson and Eckel, 2006).

If visible cues play an important role in trusting behavior and the exchange of resources, the context in which we interact with others may be important for security-related outcomes. For example, while direct face-to-face combat could be described as the “traditional method” of resource competition, online theft presents an evolutionary-novel challenge that strategists might only just be coming to terms with (see Anderson et al., 2012 for discussion). Given the potential for anonymity in the extortion of resources online, individuals may be better-placed to exploit others with impunity in these contexts. Thus, overconfidence may be expected to “evolve” among hackers, and this may be particularly pronounced among those who are less physically-equipped to inflict *immediate* costs on others during face-to-face competition. Future research could explore the relationship between personality and hacking behavior using a behavioral measure of “persistence” in “code-cracking” tasks.

PRACTICAL APPLICATIONS

Understanding how individual and environmental differences predict security outcomes could generate practical solutions to problems. The extent to which personality and appearance influence social judgments and behavior at key “barriers” to entry may enhance the overall quality of professional recruitment and training. For example, work has already demonstrated that self-rated attention to detail is predictive of security screening performance (Rusconi et al., 2012). In a high-risk, high-reliability industry, stress within the immediate environment may affect the performance of some more than others, even at basic levels of cognition. For example, while experimentally-activating feelings of power has a positive effect on performance in executive-function tasks (Smith et al., 2008) it may also promote abstract thinking at a potential cost of false recognition—broadly speaking, focussing on the bigger picture at the expense of the

finer details (Smith and Trope, 2006). The possibility that both transient (following a security breach) and stable (promotion) changes in perceived power within a security role may alter task performance is worthy of further research.

Competition for resources could be investigated at the neural level by exploring the neural basis of individual differences in morality and risk-taking in contexts related to resource acquisition and defence. Testosterone is associated with both financial risk-taking (Apicella et al., 2008; Coates and Herbert, 2008; Stanton et al., 2011) and strict endorsement of utilitarian morals (Carney and Mason, 2010), and increases as feelings of power are primed experimentally (Carney et al., 2010). Individual differences in state and trait levels of testosterone may predict the nature of the “trade-off” between the costs and benefits of monitoring and controlling perceived threats to security. Imaging studies could shed light on this, given that recent work suggests a complementary role for dopamine and noradrenalin in the evaluation of benefit and cost respectively (Bouret et al., 2012).

CONCLUSION

Biology provides a unifying framework with which to understand human behavior in light of differences between individuals and their surrounding environment. An understanding of the biological basis of strategic “biases” in social judgments can potentially increase the quality of security decision-making in light of greater awareness of the contexts and environments that might mitigate or exacerbate the risk of lost resources.

ACKNOWLEDGMENTS

I would like to thank the topic Editors for encouraging me to submit to this special issue. Thanks also to Finlay Smith for helpful comments on an earlier draft.

REFERENCES

- Anderson, R., Barton, C., Böhm, R., Clayton, R., Van Eeten, M. J. G., Levi, M., et al. (2012). *Measuring the Cost of Cybercrime*. Available online at: http://weis2012.econinfocsec.org/papers/Anderson_WEIS2012.pdf
- Andreoni, J., and Petrie, R. (2008). Beauty, gender and stereotypes: evidence from laboratory experiments. *J. Econ. Psychol.* 29, 73–93.
- Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., and Little, A. C. (2008). Testosterone and financial risk preferences. *Evol. Hum. Behav.* 29, 384–390.
- Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behav. Brain Sci.* 32, 1–63.
- Boothroyd, L. G., Jones, B. C., Burt, D. M., and Perrett, D. I. (2007). Partner characteristics associated with masculinity, health and maturity in male faces. *Pers. Individ. Differ.* 43, 1161–1173.
- Bouret, S., Ravel, S., and Richmond, B. J. (2012). Complementary neural correlates of motivation in dopaminergic and noradrenergic neurons of monkeys. *Front. Behav. Neurosci.* 6:40. doi: 10.3389/fnbeh.2012.00040
- Bowles, S. (2009). Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science* 324, 1293–1298.
- Bruce, V., and Young, A. (1986). Understanding face recognition. *Br. J. Psychol.* 77, 305–327.
- Carney, D. R., Cuddy, A. J. C., and Yap, A. J. (2010). Power posing. *Psychol. Sci.* 21, 1363–1368.
- Carney, D. R., and Mason, M. F. (2010). Decision making and testosterone: when the ends justify the means. *J. Exp. Soc. Psychol.* 46, 668–671.
- Carré, J. M., and McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proc. Biol. Sci.* 275, 2651–2656.
- Carré, J. M., McCormick, C. M., and Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychol. Sci.* 20, 1194–1198.
- Coates, J. M., and Herbert, J. (2008). Endogenous steroids and financial risk taking on a London trading floor. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6167–6172.
- Daly, M., Wilson, M., and Vasdev, S. (2001). Income inequality and homicide rates in Canada and the United States. *Can. J. Criminol.* 43, 219–236.
- Fehr, E. (2004). Don't lose your reputation. *Nature* 432, 449–450.
- Greene, J. (2003). From neural ‘is’ to moral ‘ought’: what are the moral implications of neuroscientific moral psychology? *Nat. Rev. Neurosci.* 4, 847–850.
- Hancock, P. J., Bruce, V., and Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends Cogn. Sci.* 4, 330–337.
- Hancock, P. J., and DeBruine, L. M. (2003). What's a face worth: noneconomic factors in game playing. *Behav. Brain Sci.* 26, 162–163.
- Haselton, M. G., and Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *J. Pers. Soc. Psychol.* 78, 81–91.
- Haselton, M. G., and Nettle, D. (2006). The paranoid optimist: an integrative evolutionary model of cognitive biases. *Pers. Soc. Psychol. Rev.* 10, 47–66.
- Hsu, Y., Earley, R. L., and Wolf, L. L. (2006). Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biol. Rev. Camb. Philos. Soc.* 81, 33–74.
- Hugenberg, K., and Bodenhausen, G. V. (2003). Facing prejudice: implicit prejudice and the perception of facial threat. *Psychol. Sci.* 14, 640–643.
- Jenkins, R., and Burton, A. M. (2011). Stable face representations. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 1671–1683.
- Johnson, D. D. P., and Fowler, J. H. (2011). The evolution of overconfidence. *Nature* 477, 317–320.
- Johnson, D. D. P., McDermott, R., Barrett, E. S., Cowden, J., Wrangham, R., McIntyre, M. H., et al. (2006). Overconfidence in wargames: experimental evidence on expectations, aggression, gender and testosterone. *Proc. Biol. Sci.* 273, 2513–2520.
- Jones, B. C., DeBruine, L. M., Main, J. C., Little, A. C., Welling, L. L. M., Feinberg, D. R., et al. (2010). Facial cues of dominance modulate the short-term gaze-cuing effect in human observers. *Proc. Biol. Sci.* 277, 617–624.
- Keltner, D., Gruenfeld, D., and Anderson, C. (2003). Power, approach, and inhibition. *Psychol. Rev.* 110, 265–284.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., and Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* 126, 390–423.
- Manson, J., and Wrangham, R. (1991). Intergroup aggression in chimpanzees and humans. *Curr. Anthropol.* 32, 369–390.
- Mason, M. F., Cloutier, J., and Macrae, C. N. (2006). On construing others: category and stereotype activation from facial cues. *Soc. Cogn.* 24, 540–562.
- Maynard Smith, J., and Price, G. R. (1973). The logic of animal conflict. *Nature* 246, 15–18.
- Mayr, E. (1963). *Animal Species and Evolution*. Cambridge, MA: Harvard University Press.
- Mazzella, R., and Feingold, A. (1994). The effects of physical attractiveness, race, socioeconomic status, and gender of defendants and victims on judgments of mock jurors: a meta-analysis. *J. Appl. Soc. Psychol.* 24, 1315–1338.
- National Fraud Authority. (2012). *Annual Fraud Indicator*. London: Home Office.
- Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* 437, 1291–1298.
- Oosterhof, N. N., and Todorov, A. (2008). The functional basis of face evaluation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11087–11092.
- Perrett, D. I., Lee, K. J., Penton-Voak, I. S., Rowland, D. R., Yoshikawa, S., Burt, D. M., et al. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature* 394, 884–887.
- Pound, N., Penton-Voak, I. S., and Surridge, A. K. (2009). Testosterone responses to competition in men are related to facial masculinity. *Proc. Biol. Sci.* 276, 153–159.
- Price, M. E., Kang, J., Dunn, J., and Hopkins, S. (2011). Muscularity and attractiveness as predictors of human egalitarianism. *Pers. Individ. Differ.* 50, 636–640.
- Puts, D. A. (2010). Beauty and the beast: mechanisms of sexual selection in humans. *Evol. Hum. Behav.* 31, 157–175.
- Puts, D. A., Gaulin, S. J. C., and Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* 27, 283–296.
- Puts, D. A., Jones, B. C., and DeBruine, L. M. (2012). Sexual selection on human faces and voices. *J. Sex Res.* 49, 227–243.
- Rusconi, E., McCrory, E., and Viding, E. (2012). Self-rated attention to detail predicts threat detection performance in security X-ray images. *Secur. J.* 25, 356–371.

- Sell, A., Tooby, J., and Cosmides, T. (2009). Formidability and the logic of human anger. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15073–15078.
- Smith, F. G., DeBruine, L. M., Jones, B. C., Krupp, D. B., Welling, L. L. M., and Conway, C. A. (2009). Attractiveness qualifies the effect of observation on trusting behavior in an economic game. *Evol. Hum. Behav.* 30, 393–397.
- Smith, P. K., Jostmann, N. B., Galinsky, A. D., and van Dijk, W. W. (2008). Lacking power impairs executive functions. *Psychol. Sci.* 19, 441–447.
- Smith, P. K., and Trope, Y. (2006). You focus on the forest when you're in charge of the trees: power priming and abstract information processing. *J. Pers. Soc. Psychol.* 90, 578–596.
- Solnick, S. J., and Schweitzer, M. E. (1999). The influence of physical attractiveness and gender on ultimatum game decisions. *Organ. Behav. Hum. Decis. Process.* 793, 199–215.
- Stanton, S. J., Lienen, S. H., and Schultheiss, O. C. (2011). Testosterone is positively associated with risk taking in the Iowa gambling task. *Horm. Behav.* 59, 252–256.
- Stirrat, M., and Perrett, D. I. (2010). Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychol. Sci.* 21, 349–354.
- Todorov, A., Said, C. P., Engell, A. D., and Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends Cogn. Sci.* 12, 455–460.
- Tinbergen, N. (1963). On the aims and methods of ethology. *Z. Tierpsychol.* 20, 410–433.
- Watkins, C. D., Fraccaro, P. J., Smith, F. G., Vukovic, J., Feinberg, D. R., DeBruine, L. M., et al. (2010a). Taller men are less sensitive to cues of dominance in other men. *Behav. Ecol.* 21, 943–947.
- Watkins, C. D., Jones, B. C., and DeBruine, L. M. (2010b). Individual differences in dominance perception: dominant men are less sensitive to facial cues of male dominance. *Pers. Individ. Differ.* 49, 967–971.
- Watkins, C. D., and Jones, B. C. (2012). Priming men with different contest outcomes modulates their dominance perceptions. *Behav. Ecol.* 23, 539–543.
- Webster, C. W. R. (2009). CCTV policy in the UK: reconsidering the evidence base. *Surveill. Soc.* 6, 10–22.
- Willis, J., and Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychol. Sci.* 17, 592–598.
- Wilson, M., and Daly, M. (2006). Are juvenile offenders extreme future discounters? *Psychol. Sci.* 17, 989–994.
- Wilson, R. K., and Eckel, C. C. (2006). Judging a book by its cover: beauty and expectations in the trust game. *Polit. Res. Q.* 59, 189–202.

Received: 08 April 2013; accepted: 01 May 2013; published online: 16 May 2013.

Citation: Watkins CD (2013) Cues derived from facial appearance in security-related contexts: a biological and socio-cognitive framework. *Front. Hum. Neurosci.* 7:204. doi: 10.3389/fnhum.2013.00204

Copyright © 2013 Watkins. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



The reasoning criminal vs. Homer Simpson: conceptual challenges for crime science

Noémie Bouhana*

Department of Security and Crime Science, University College London, London, UK

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

Elliot Berkman, University of Oregon, USA

Nicola Lettieri, Isfol, Italy

***Correspondence:**

Noémie Bouhana, Department of Security and Crime Science, University College London, 35 Tavistock Square, WC1H 9EZ, London, UK
e-mail: n.bouhana@ucl.ac.uk

A recent disciplinary offshoot of criminology, crime science (CS) defines itself as “the application of science to the control of crime.” One of its stated ambitions is to act as a cross-disciplinary linchpin in the domain of crime reduction. Despite many practical successes, notably in the area of situational crime prevention (SCP), CS has yet to achieve a commensurate level of academic visibility. The case is made that the growth of CS is stifled by its reliance on a model of decision-making, the Rational Choice Perspective (RCP), which is inimical to the integration of knowledge and insights from the behavioral, cognitive and neurosciences (CBNs). Examples of salient developments in the CBNs are provided, as regards notably multiple-system perspectives of decision-making and approaches to person-environment interaction. Short and long-term benefits of integration for CS are briefly outlined.

Keywords: crime science, situational crime prevention, rational choice, decision-making, theory

A VERY SHORT INTRODUCTION TO CRIME SCIENCE

A recent disciplinary offshoot of criminology, crime science (CS) defines itself as “the application of science to the control of crime” (Laycock et al., 2005; Laycock, 2008:149). Problem-driven, CS is chiefly concerned with the design of social and technological systems in service to the needs of stakeholders and end-users—be they industry, government, security agencies, or the general public. Underpinning CS and its preferred approach to crime reduction, situational crime prevention (SCP), is the premise that crime is best tackled by targeting its immediate causes. This focus on proximate factors is intentionally lopsided. While the necessary conditions of crime are defined as the intersection in time and space of a motivated offender and a suitable target in the absence of a capable guardian, relatively little attention has been paid to the “offender” part of the equation. CS digs its philosophical roots in the 18th Century Classical School, whereby Man is understood as an essentially self-interested animal driven by desires which he seeks to fulfill while incurring the least amount of effort. Susceptibility to temptation is thus taken as a given and CS looks to situational control—the removal of temptations—as the most promising crime reduction strategy. “Opportunity makes the thief”: remove the opportunity, increase the effort and reduce the rewards of offending, and the crime will be prevented (Clarke, 2012).

The effectiveness of this approach has been demonstrated against a diverse range of crime problems. The promise of technological solutions and an emphasis on practical problem-solving have been popular with law enforcement agencies, and the claim was made that CS would soon eclipse criminology departments within universities (Clarke, 2004). However, CS has yet to achieve commensurate visibility in the academic sphere. This paper contends that the conceptual

limitations of CS’s standard model of decision-making, the Rational Choice Perspective (RCP), as well as the discipline’s largely “bottom-up” research programme, hold it back from fulfilling its stated ambition to act as a cross-disciplinary linchpin (Laycock et al., 2005). The case is made that CS must look to developments in the cognitive, behavioral and neurosciences (henceforth, Cognitive and Neurosciences (CBNs)) to address RCP’s shortcomings. Examples of developments which suggest potential for integration are provided. In conclusion, the benefits of integration are further outlined.

THE CASE FOR “BOUNDED” PARSIMONY

It is not possible to leave the offender out of crime prevention altogether. In order to “increase effort” and “reduce rewards”, a model of criminal decision-making is needed. For this purpose, the fathers of SCP adopted the RCP (Clarke et al., 1985; Cornish et al., 2008). As presented, RCP is not a theory *per se*, but a heuristic device, a “good enough” conceptual model which provides a schematic understanding of how offenders make decisions—evaluating, to the best of their abilities, the costs and benefits of their actions. Armed with this basic understanding, the crime controller can design an array of situational techniques to influence the offender’s decisional process away from crime (Smith and Clarke, 2012).

While RCP has met with notable success as an engineering heuristic, it has fallen short as a model of offender decision-making (Wortley et al., 2013). Although the framework acknowledges, on the one hand, the less-than-rational aspects of offender decision-making—criminal rationality is described as “bounded”—it implies, on the other, that the problem isn’t worth agonizing over: a parsimonious, *as-if* model, unencumbered

by the vagaries of human affect and cognition, should serve the crime controllers well enough (Smith and Clarke, 2012). As Wortley et al. (2013) observes, this state of affairs has had the consequence of stifling theoretical development in CS, so much so that RCP has remained essentially static since the 1980s. One may take Wortley's critique further and observe that other theoretical perspectives within the "family" of opportunity theories—notably, the Routine Activity Approach (Cohen and Felson, 1979)—have likewise remained relatively untouched. Opportunity theories are still, to a large extent, axiomatic statements rather than explanations of the causal processes which bring crime about (Wikström et al., 2011). This is illustrated by the oft-repeated claim that opportunities cause crime (Felson and Clarke, 1998); for it is not, of course, the opportunity which causes the crime, but its perception by the offender (the Thomas Theorem in action), among other processes: opportunities, whether provocations or temptations, are not criminal in themselves. To address this problem, some have proposed that the ecological concept of *affordance* (Gibson, 1979) should replace opportunity in CS parlance (Pease et al., 2006). However, affordance has yet to be integrated into the wider opportunity control framework. To take affordance on board, a model of criminal action is required which explains motivation in terms of the interaction between individual and situation, instead of postulating it as a given.

The move towards a more dynamic, interactionist model has been resisted, for fear that it would compromise RCP's radical parsimony, a condition of its heuristic usefulness. Faced with evidence of the non-rational features of offender decision-making, the strategy has been to stretch the concept of "rationality" to encompass the new phenomena. Drives to criminal action are restated as factors in a cost-benefit analysis. Psychological rewards (e.g., excitement), moral emotions (e.g., guilt, shame), social inducements (e.g., status), psychobiological factors (e.g., addiction), and so on, are reinterpreted in "rational" terms (e.g., Clarke et al., 1997). This approach renders the model impregnable, but runs roughshod over Einstein's admonition that theory should "make the irreducible basic elements as simple and as few as possible *without having to surrender the adequate representation of a single datum of experience*" (Einstein, 1934:165, emphasis added). The construct which explains everything explains nothing: the more phenomena is stuffed into the construct, the emptier it becomes. "Bounded" rather than radical parsimony would seem the more reasonable option.

DRAWBACKS OF "BOTTOM-UP" RESEARCH

Calls to overhaul RCP and bring the offender back into SCP have been sounded in the past (Ekblom and Tilley, 2000; Wortley, 2001; Wortley et al., 2013), but have fallen on reluctant ears. New SCP techniques concerned with situational precipitators have been added to the catalogue (Cornish et al., 2003), falling far short of a conceptual shake-up. CS's continuing identity struggle may explain this inertia: "science" moniker aside, CS is fundamentally an engineering discipline, with a self-confessed preference for short-term problem-solving (Laycock et al., 2005). At the outset, SCP was established as the technological framework most likely to deliver returns. A number of technological rules

and design principles, most of them implicated in opportunity control, were identified, which produced reliable results. The discipline's scientific programme was thus largely circumscribed to those research activities which provided a knowledge-base for the design of opportunity control technologies (broadly defined), or contributed to the testing, validation and refinement of those technological rules and design principles at the heart of the discipline.

Arguably, the crime scientist's trademark question is, "So what?" (Laycock, 2012). If the topic is not self-evidently useful to crime control, it is not worth investigating. On the upside, this instrumental approach, whereby CS's engineering ambitions dictate the discipline's research activity, has produced reliable analytical tools and prevention technologies, which have achieved concrete gains in terms of crime reduction. On the downside, this relatively narrow research agenda has done little to encourage inquiry driven by "big questions". Indeed, crime scientists have been known to take criminologists to task for studying the "wrong" kinds of causes and failing to be more problem-oriented (Clarke, 2004), as if only a finite number of scientific questions about crime were worth asking.

The concern is that this "bottom-up" research agenda has insularised CS from a wealth of knowledge in other disciplines, notably the CBNs, as much as it has impeded theoretical growth from within. Yet a field which looks to medicine as a desirable model of cross-disciplinarity (Laycock et al., 2005) needs a conceptual framework which *affords* (in Gibson's sense of the word) disciplinary integration. Medicine and its parent disciplines share the foundations of a systemic (chemical, biological, psychosocial, ecological, and so on) understanding of the human organism and its environment. To achieve its stated goal, CS needs, if not a unified framework, then conceptual models which are not inimical to neighboring research programmes. As a first step, opportunity perspectives should clarify what they mean by "bounded rationality" and formulate explicit mechanisms of person-situation interaction (which will also necessitate a clear definition of "situation"; Snyder, 2013). Examples of developments in the CBNs may illustrate the value of integration.

ENTERS HOMER SIMPSON, STAGE RIGHT

The outsider looks on with envy at the effervescence which has characterized the growth and, increasingly, the integration of the CBNs in recent years. Given the breakneck speed of research in these domains, an overview isn't attempted, but it is noteworthy that the surge of activity has often been accompanied, if not triggered, by an empirical challenge to single-factor (notably rationalist) models and theories.

In social psychology, dual-process models (Evans, 2003; Mischel et al., 2004; Kahneman et al., 2005; Kahneman, 2011) followed from observations that departures from classical rationality are an ubiquitous feature of human thinking (Kahneman et al., 1982; Kahneman, 2011). In moral psychology, dual models of moral judgment have likewise emerged which call into question the Kohlbergian view of moral development, adopting instead an adaptationist perspective in which moral

intuitions underpin moral judgment as much as moral reasoning, if not more so (Haidt, 2001; Greene and Gazzaniga, 2009; Cushman et al., 2010).

Of particular interest, given SCP's original borrowing of the rational perspective from economics, has been the development of behavioral economics, which built upon social psychology's insights to address commonly observed violations of the standard neo-classic model (Thaler, 1991; Mullainathan et al., 2001). As Camerer et al. (2004) put it, "At the core of behavioral economics is the conviction that increasing the realism of the psychological underpinnings of economic analysis will improve economics on its own terms—generating theoretical insights, making better predictions of field phenomena, and suggesting better policy." The scientific gain, behavioral economists feel, is worth renouncing the seductive (i.e., simple and clear-cut), but ultimately misleading, solutions proposed by standard models. While neo-classical economics would like people to think like Mr. Spock, the average human being is rather closer to Homer Simpson (Thaler and Sunstein, 2008). Policies aimed at improving anything from individual health to personal finances, road safety, energy savings, and so on, are better designed while keeping Springfield's most famous resident in mind. Boosted by these developments in behavioral economics, neuroeconomics has set out to open the "black box" of the economic brain (Camerer et al., 2005), progressively adding detail to an "emotional" organ (Oullier et al., 2010) constituted of neural systems so enmeshed it makes little sense to study decision-making without reference to emotional states (Sanfey et al., 2006), or—another fundamental revision to the standard models—without reference to the socio-physical environment.

THE FUTURE'S BRIGHT, THE FUTURE'S INTERACTIVE

The emphasis on system interaction within the organism has been accompanied by growing attention to organism-environment interaction. Given the importance of self-control to the explanation of criminal behavior (Tooby and Cosmides, 2007), research on self-regulation is particularly instructive, revealing self-control to be less of a fixed "trait" than a complex situational mechanism. How much of this resource individuals may draw on in any given circumstance is influenced by situational features, as well as individual factors. Self-control can be depleted by the prior exercise of self-control (Baumeister et al., 2007) and by the exercise of choice between alternatives (Vohs et al., 2008), with implications for the subsequent ability to self-monitor, cope with stress, control aggression, think logically, and so on. It can be depleted vicariously by watching others exercise restraint (Ackerman et al., 2009), but can also be restored vicariously by taking on the perspective of others engaged in self-control replenishing activities (Egan et al., 2012). Relevantly, self-regulatory depletion is associated with unethical behavior in well-intentioned individuals, though much less so in individuals with highly internalized moral standards, plausibly because they do not need to engage in higher cognitive processes, but automatically disregard the opportunity to behave unethically (Gino et al., 2011). This observation would seem to support situational action models of moral rule-breaking (Svensson et al., 2010).

More generally, self-regulation is sensitive to cognitive load. Decisions-making in environments which impose a high cognitive burden on individuals can lead to greater reliance on (more economical) automated decision-making, which in turn can lead to cognitive shortcuts, such as racial stereotyping (Burgess, 2010). Research into the causes of self-defeating decision-making among the poor suggests that the very conditions that define poverty, such as scarcity, impact decision-making through biosocial mechanisms which produce attentional shifts, self-control depletion, and reduce cognitive capacity generally (Spears, 2010; Shah et al., 2012; Mani et al., 2013). Self-regulation depletion also appears affected by self-belief, whereby individuals' implicit theories of willpower moderate self-control depletion (Job et al., 2010). Overall, modern research offers an increasingly sophisticated picture of self-control as a fluctuating resource subject to the interaction of an array of individual and socio-contextual factors (see Inzlicht and Schmeichel, 2012). It also suggests avenues to integrate mechanistically so-called "root causes" (e.g., poverty) and situational choice perspectives, traditionally at odds in the context of crime studies.

Interaction is, naturally, a chief concern of those disciplines working within an adaptationist framework. In the context of evolutionary psychology, "rationality" is not portrayed as a universal construct; rather, processes are understood as domain-specific and may produce "faulty" choices when considered from another behavioral domain's point-of-view. In this sense, rationality is not so much bounded as *ecological* (Tooby and Cosmides, 2007). This perspective suggests a framework for the continued development of still-rare ecological studies of criminal decision-making (Snook et al., 2011). It might be worthwhile in that context to explore how domain-specific processes relate (or not) to domain-general processes (Chiappe and MacDonald, 2005), as well as to niche construction (Laland and Brown, 2006).

Beyond functional explanations, evolutionary perspectives of human development have yielded constructs such as "differential susceptibility to the environment" and "biological sensitivity to context", which add to an understanding of the role of individual differences in the outcome of person-environment interactions (Ellis et al., 2011). They suggest that heightened vulnerability to context runs both ways—some individuals are more susceptible to *both* negative *and* positive influences—and raise intriguing questions as to the persistent effect, if any, of this susceptibility into adulthood. Even these exceedingly brief examples suggest significant potential to progress CS's take on person-situation interaction beyond its (relatively) primitive state.

SO WHAT?

The preceding should not be taken as an entreaty for crime scientists to give up their preferred methods and reach for the fMRI—though, as with previous successful imports from epidemiology (e.g., Bowers and Johnson, 2004), greater integration will likely result in substantial methodological gains. Nor is it a demand to adopt any given approach wholesale. Indeed, the most onerous part of the conceptual shift advocated here will be to keep up with fundamental debates internal to other disciplines (e.g., Bolhuis et al., 2011). It should, however, be taken as a plea for scientific realism, for the development of

theories of human behavior which go beyond axiomatic, “as-if” theoretical frameworks to specify the constellation of biosocial mechanisms which account for the phenomenon (Bunge et al., 2006). As it stands, CS’s standard model, RCP, isolates it from a wealth of knowledge in contemporary disciplines. This is a major obstacle to the development of a modern science of crime prevention.

This proposal for a more modern approach to conceptual development should not be interpreted, either, as a request to relinquish the problem-solving side of the business. Tackling practical problems generates hypotheses and throws up invaluable challenges to theoretical assumptions. Furthermore, embracing the CBN knowledge-base is bound to open up short-term avenues for crime prevention engineering. Research on the deleterious effects of cognitive load on healthcare decision-making already suggests that environmental changes, learned routines and “reflective practice” could improve the performance of crime controllers working in stressful settings (Burgess, 2010). Understanding the rewards associated with automated brain processes hints at strategies to tackle resistance to change in law enforcement organizations (Becker and Cropanzano, 2010). Experiments which elicit moral emotions such as disgust, combined with eye-tracking studies of anti-smoking warnings, could inform the design and evaluation of crime prevention publicity campaigns (see Oullier and Sauneron, 2010). Likewise, neuroimaging studies of the Ultimatum Game—which investigate why participants “irrationally” turn down money when faced with offers perceived as unfair—might help crime controllers understand why “rational” crime prevention advice is sometimes spiritedly rejected by potential victims (such as advice which suggests women should alter their behavior to prevent sexual assault).

More ambitiously, the convergence of cognitive neuroscience, social psychology, architecture (e.g., Sternberg and Wilson,

2006), consumer studies (e.g., Mick et al., 2004), and crime prevention might inspire interdisciplinary research into the design of “neurocognitively sustainable” environments, which would aim to minimize deleterious interaction (in terms of cognitive overload, depletion of self-control, and so on), with the prospect of benefit diffusion across multiple categories of social problems. The perspective of a wide-ranging contribution from evolutionary psychology has already captured the imagination of crime scientists (Roach and Pease, 2013), though reminders that adaption is an onerous explanatory concept, and that accounts of ultimate (evolutionary) causes must be accompanied by an understanding of proximal (e.g., neuropsychological) mechanisms, should be heeded (de Waal, 2002). In criminology, embryonic comparative research into the executive functioning of white collar criminals (Raine et al., 2012) hints at the possibility of tailoring prevention technologies by offending type. Executive functioning—self-regulation, but also the functions which underpin cognitive adaptability and flexibility—is likely to be a fruitful area of research for CS should it seek to account more deeply for the failure of many criminals to displace. When explaining human behavior, evaluating causal factors in isolation makes poor sense. A science of crime prevention should become comfortable with multilevel theorizing.

This paper proceeded from a simple premise: that a scientific discipline which aims to capture the imagination of future generations of researchers cannot exist only to solve practical problems; it must also set out to answer fundamental questions. While technology must be simple enough for end-users to implement, the science which is the bedrock of these technologies should be as complex as it needs to be. “Good enough” theory surrenders too much of experience to be worth the short-term benefits to any scientific discipline.

REFERENCES

- Ackerman, J. M., Goldstein, N. J., Shapiro, J. R., and Bargh, J. A. (2009). You wear me out: the vicarious depletion of self-control. *Psychol. Sci.* 20, 326–332. doi: 10.1111/j.1467-9280.2009.02290.x
- Baumeister, R. F., Vohs, K. D., and Tice, D. M. (2007). The strength model of self-control. *Curr. Dir. Psychol. Sci.* 16, 351–355. doi: 10.1111/j.1467-8721.2007.00534.x
- Becker, W. J., and Cropanzano, R. (2010). Organizational neuroscience: the promise and prospects of an emerging discipline. *J. Organ. Behav.* 31, 1055–1059. doi: 10.1002/job.668
- Bolhuis, J. J., Brown, G. R., Richardson, R. C., and Laland, K. N. (2011). Darwin in mind: new opportunities for evolutionary psychology. *PLoS Biol.* 9:e1001109. doi: 10.1371/journal.pbio.1001109
- Bowers, K., and Johnson, S. (2004). Who commits near repeats? A test of the boost explanation. *West. Crim. Rev.* 5, 12–24.
- Bunge, M. (2006). “A systemic perspective on crime,” in *The Explanation of Crime: Context, Mechanisms, and Development*, eds P.-O. Wikström and R. J. Sampson (Cambridge: Cambridge University Press), 8–30.
- Burgess, D. J. (2010). Are providers more likely to contribute to healthcare disparities under high levels of cognitive load? How features of the healthcare setting may lead to biases in medical decision making. *Med. Decis. Making* 30, 246–257. doi: 10.1177/0272989X09341751
- Camerer, C. F., and Loewenstein, G. (2004). “Behavioral economics: past, present and future,” in *Advances in Behavioral Economics*, eds C. Camerer, G. Loewenstein and M. Rabin (New York: Princeton University Press), 3–51.
- Camerer, C., Loewenstein, G., and Prelec, D. (2005). Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* 43, 9–64. doi: 10.1257/0022051053737843
- Chiappe, D., and MacDonald, K. (2005). The evolution of domain-general mechanisms in intelligence and learning. *J. Gen. Psychol.* 132, 5–40. doi: 10.3200/GENP.132.1.5-40
- Clarke, R. V. (2004). Technology, criminology and crime science. *Eur. J. Crim. Policy Res.* 10, 55–63. doi: 10.1023/B:CRIM.0000037557.42894.f7
- Clarke, R. V. (2012). Opportunity makes the thief. Really? And so what? *Crime Sci.* 1:3. doi: 10.1186/2193-7680-1-3
- Clarke, R. V., and Cornish, D. B. (1985). “Modeling offenders’ decisions: a framework for research and policy,” in *Crime and Justice: An Annual Review of Research* Vol. 6, eds M. Tonry and N. Morris (Chicago: University of Chicago Press), 147–185.
- Clarke, R. V., and Homel, R. (1997). “A revised classification of situational crime prevention techniques,” in *Crime Prevention at the Crossroads*, ed S. P. Lab (Cincinnati, OH: Anderson), 17–30.
- Cohen, L. E., and Felson, M. (1979). Social change and crime rate trends: a routine activities approach. *Am. Sociol. Rev.* 44, 588–608. doi: 10.2307/2094589
- Cornish, D. B., and Clarke, R. V. (2003). “Opportunities, precipitators and criminal decisions: a reply to Wortley’s critique of situational crime prevention,” in *Theory for Practice in Situational Crime Prevention*, eds M. Smith and D. B. Cornish (Monsey, NY: Criminal Justice Press), 41–96.

- Cornish, D. B., and Clarke, R. V. (2008). "The rational choice perspective," in *Environmental Criminology and Crime Analysis*, eds R. Wortley and L. Mazerolle (Cullompton: Willan), 21–47.
- Cushman, F., Young, L., and Greene, J. D. (2010). "Our multi-system moral psychology: towards a consensus view," in *The Oxford Handbook of Moral Psychology*, eds J. Doris, G. Harman, S. Nichols, J. Prinz, W. Sinnott-Armstrong and S. Stich (New York: Oxford University Press), 47–71.
- de Waal, F. (2002). Evolutionary psychology: the wheat and the chaff. *Curr. Dir. Psychol. Sci.* 11, 187–191. doi: 10.1111/1467-8721.00197
- Egan, P., Hirt, E., and Karpen, S. (2012). Taking a fresh perspective: vicarious restoration as a means of recovering self-control. *J. Exp. Soc. Psychol.* 48, 457–465. doi: 10.1016/j.jesp.2011.10.019
- Einstein, A. (1934). On the method of theoretical physics. *Philos. Sci.* 1, 163–169. doi: 10.1086/286316
- Eklom, P., and Tilley, N. (2000). Going equipped: criminology, situational crime prevention and the resourceful offender. *Br. J. Criminol.* 40, 376–398. doi: 10.1093/bjc/40.3.376
- Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., and Van Ijzendoorn, M. H. (2011). Differential susceptibility to the environment: an evolutionary-neurodevelopmental theory. *Dev. Psychopathol.* 23, 7–28. doi: 10.1017/S0954579410000611
- Evans, J. S. (2003). In two minds: dual-process accounts of reasoning. *Trends Cogn. Sci.* 7, 454–459. doi: 10.1016/j.tics.2003.08.012
- Felson, M., and Clarke, R. V. (1998). *Opportunity Makes the Thief: Practical Theory for Crime Prevention*. Police Research Series, Paper 98. London: Home Office.
- Gibson, J. J. (1979). *The Ecological Approach to Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gino, F., Schweitzer, M., Mead, N., and Arieli, D. (2011). Unable to resist temptation: how self-control depletion promotes unethical behavior. *Organ. Behav. Hum. Decis. Process* 115, 191–203. doi: 10.1016/j.obhdp.2011.03.001
- Greene, J. D. (2009). "The cognitive neuroscience of moral judgment," in *The Cognitive Neurosciences*, 4th Edn. ed M. S. Gazzaniga (Cambridge, MA: MIT Press), 987–1002.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol. Rev.* 108, 814–834. doi: 10.1037/0033-295x.108.4.814
- Inzlicht, M., and Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspect. Psychol. Sci.* 7, 450–463. doi: 10.1177/1745691612454134
- Job, V., Dweck, C. S., and Walton, G. M. (2010). Ego depletion—is it all in your head? implicit theories about willpower affect self-regulation. *Psychol. Sci.* 21, 1686–1693. doi: 10.1177/0956797610384745
- Kahneman, D. (2011). *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kahneman, D., and Frederick, S. (2005). "A model of heuristic judgment," in *The Cambridge Handbook of Thinking and Reasoning*, eds K. J. Holyoak and R. G. Morrison (New York: Cambridge University Press), 267–293.
- Kahneman, D., Slovic, P., and Tversky, A. (eds.) (1982). *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Laland, K. N., and Brown, G. R. (2006). Niche construction, human behaviour and the adaptive-lag hypothesis. *Evol. Anthropol.* 15, 95–104. doi: 10.1002/evan.20093
- Laycock, G. (2005). "Defining crime science," in *Crime Science: New Approaches to Preventing and Detecting Crime*, eds M. J. Smith and N. Tilley (Cullompton: Willan), 3–25.
- Laycock, G. (2008). Introduction: special edition on crime science. *Policing* 2, 149–153. doi: 10.1093/police/pan028
- Laycock, G. (2012). Editorial introduction to 'Opportunity makes the thief. Really? And so what?'. *Crime Sci.* 1:2. doi: 10.1186/2193-7680-1-2
- Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty impedes cognitive function. *Science* 341, 976–980. doi: 10.1126/science.1238041
- Mick, D. G., Broniarczyk, S. M., and Haidt, J. (2004). Choose, choose, choose, choose, choose, choose: emerging and prospective research on the deleterious effects of living in consumer hyper-choice. *J. Bus. Ethics* 52, 207–211. doi: 10.1023/B:BUSI.0000035906.74034.d4
- Mischel, W., and Ayduk, O. (2004). "Willpower in a cognitive-affective processing system: the dynamics of delay of gratification," in *Handbook of Self-Regulation: Research, Theory, and Applications*, eds R. F. Baumeister and K. D. Vohs (New York: Guilford), 99–129.
- Mullainathan, S., and Thaler, R. H. (2001). "Behavioral economics," in *International Encyclopedia of the Social and Behavioral Sciences*, eds N. Smelser and P. Bates (New York: Pergamon), 1094–1110.
- Oullier, O. (2010). "The useful brain: why neuroeconomics might change our views on rationality and a couple of other things," in *The Irrational Economist: Making Decisions in a Dangerous World*, eds E. Michel-Kerjan and P. Slovic (New York: Public Affairs), 88–96.
- Oullier, O., and Sauneron, S. (eds.) (2010). *Improving Public Health Prevention with Behavioural, Cognitive and Neuroscience*. Paris: La Documentation Française.
- Pease, K. (2005). "Science in the service of crime reduction," in *Handbook of Crime Prevention and Community Safety*, ed N. Tilley (Cullompton: Willan Press), 171–197.
- Pease, K. (2006). "No through road: closing pathways to crime," in *Crime Reduction and the Law*, eds K. Moss and M. Stephens (London: Routledge), 50–66.
- Raine, A., Laufer, W. S., Yang, Y., Narr, K. L., Thompson, P., and Toga, A. W. (2012). Increased executive functioning, attention, and cortical thickness in white-collar criminals. *Hum. Brain Mapp.* 33, 2932–2940. doi: 10.1002/hbm.21415
- Roach, J., and Pease, K. (2013). *Evolution and Crime*. London: Routledge.
- Sanfey, A. G., Loewenstein, G., McClure, S. M., and Cohen, J. D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends Cogn. Sci.* 10, 108–116. doi: 10.1016/j.tics.2006.01.009
- Shah, A. K., Mullainathan, S., and Shafir, E. (2012). Some consequences of having too little. *Science* 338, 682–685. doi: 10.1126/science.1222426
- Smith, M. J., and Clarke, R. V. (2012). "Situational crime prevention: classifying techniques using 'good enough' theory," in *The Oxford Handbook of Crime Prevention*, eds B. C. Welsh and D. P. Farrington (New York: Oxford University Press), 291–315.
- Snook, B., Dhimi, M. K., and Kavanagh, J. M. (2011). Simply criminal: predicting burglars' occupancy decisions with a simple heuristic. *Law Hum. Behav.* 35, 316–326. doi: 10.1007/s10979-010-9238-0
- Snyder, M. (2013). B = f (P, S): perspectives on persons and situations, from Lewin to Bond and beyond. *Asian J. Soc. Psychol.* 16, 16–18. doi: 10.1111/ajsp.12013
- Spears, D. (2010). *Economic Decisionmaking in Poverty Depletes Behavioral Control* (Working Paper No. 213). Princeton: Center for Economic Policy Studies. Retrieved from <http://www.princeton.edu/ceps/workingpapers/213spears.pdf>.
- Sternberg, E. M., and Wilson, M. A. (2006). Neuroscience and architecture: seeking common ground. *Cell* 127, 239–242. doi: 10.1016/j.cell.2006.10.012
- Svensson, R., Pauwels, L., Weerman, F. (2010). Does the effect of self-control on adolescent offending vary by level of morality? *Crim. Justice Behav.* 37, 732–743. doi: 10.1177/0093854810366542
- Thaler, R. H. (1991). *The Winner's Curse: Paradoxes and Anomalies of Economic Life*. Princeton: Princeton University Press.
- Thaler, R. H., and Sunstein, C. (2008). *Nudge: Improving Decisions About Health, Wealth and Happiness*. London: Penguin.
- Tooby, J., and Cosmides, L. (2007). Evolutionary psychology, ecological rationality, and the unification of the behavioral sciences. *Behav. Brain Sci.* 30, 42–43. doi: 10.1017/s0140525x07000854
- Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., and Tice, D. M. (2008). Making choices impairs subsequent self-control: a limited resource account of decision making, self-regulation, and active initiative. *J. Pers. Soc. Psychol.* 94, 883–898. doi: 10.1037/0022-3514.94.5.883
- Wikström, P.-O. (2011). "Does everything matter? Addressing the problem of causation and explanation in the study of crime," in *When Crime Appears: The Role of Emergence*, eds J. M. McGloin, C. J. Sullivan and L. W. Kennedy (London: Routledge), 53–72.
- Wikström, P.-O., and Trieber, K. (2007). The role of self-control in crime causation: beyond Gottfredson and Hirschi's general theory of crime. *Eur. J. Criminol.* 4, 237–264. doi: 10.1177/1477370807074858
- Wortley, R. (2001). A classification of techniques for controlling situational precipitators of crime. *Secur. J.* 14, 63–82. doi: 10.1057/palgrave.sj.8340098

Wortley, R. (2013). "Rational choice and offender decision making: lessons from the cognitive sciences," in *Cognition and Crime: Offender Decision Making and Script Analysis*, eds B. Leclerc and R. Wortley (London: Routledge), 237–251.

Conflict of Interest Statement: The authors declare that the research

was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 August 2013; accepted: 28 September 2013; published online: 23 October 2013.

Citation: Bouhana N (2013) The reasoning criminal vs. Homer Simpson:

conceptual challenges for crime science. Front. Hum. Neurosci. 7:682. doi: 10.3389/fnhum.2013.00682

This article was submitted to the journal Frontiers in Human Neuroscience.

Copyright © 2013 Bouhana. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY).

The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A biological security motivation system for potential threats: are there implications for policy-making?

Erik Z. Woody^{1*} and Henry Szechtman²

¹ Department of Psychology, University of Waterloo, Waterloo, ON, Canada

² Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada

Edited by:

Andrea Szymkowiak, University of
Abertay Dundee, UK

Reviewed by:

Jacqueline M. Archibald, University
of Abertay Dundee, UK
Aaron Elkins, Imperial College
London, UK

*Correspondence:

Erik Z. Woody, Department of
Psychology, University of Waterloo,
200 University Ave West, Waterloo,
ON N2L 3G1, Canada
e-mail: ewoody@uwaterloo.ca

Research indicates that there is a specially adapted, hard-wired brain circuit, the security motivation system, which evolved to manage potential threats, such as the possibility of contamination or predation. The existence of this system may have important implications for policy-making related to security. The system is sensitive to partial, uncertain cues of potential danger, detection of which activates a persistent, potent motivational state of wariness or anxiety. This state motivates behaviors to probe the potential danger, such as checking, and to correct for it, such as washing. Engagement in these behaviors serves as the terminating feedback for the activation of the system. Because security motivation theory makes predictions about what kinds of stimuli activate security motivation and what conditions terminate it, the theory may have applications both in understanding how policy-makers can best influence others, such as the public, and also in understanding the behavior of policy-makers themselves.

Keywords: potential danger, precautionary behavior, security motivation, risk, decision making, obsessive-compulsive disorder (OCD)

INTRODUCTION

The world in which we currently live confronts people responsible for making decisions about security with very challenging issues. These issues call for sophisticated logical and statistical analysis, detection and forecasting systems, cost-benefit analysis, and the like. However, the crux of security is the necessity of dealing with the prospect of potential danger. Because potential dangers have had very substantial consequences for reproductive fitness for many thousands of years, evolution has shaped brain systems specially adapted for managing them. Thus, in addition to the logical armamentarium that present-day decision-makers bring to issues of security, they inevitably bring the intuitions and motivations that are generated by a biologically ancient, “hard-wired” system.

This potential-threat system in the brain has been termed the defense system (Trower et al., 1990) and the hazard-precaution system (Boyer and Lienard, 2006). In our own work, we have called it the security motivation system (Szechtman and Woody, 2004). Our research investigating this system has focused on its role in everyday circumstances, such as behavior to manage threats of contagion due to dirt and germs, and in pathological variants of these behaviors, such as the compulsive hand-washing seen in obsessive-compulsive disorder (OCD). However, it is likely that the influence of the security motivation system extends well beyond such relatively mundane circumstances. The purpose of this perspective article is to explain briefly what we know about the security motivation system and to advance the following question: Does this biological system affect policy-making about security in important ways? We hope to stimulate the thinking of researchers who investigate security-related decision-making, in particular by sketching some of the kinds of hypotheses that could be examined in such research.

PROPERTIES OF THE SECURITY MOTIVATION SYSTEM

The security motivation system is hypothesized to be a reasonably distinct module in the brain, which evolved to be specially adapted for handling potential threats (Tooby and Cosmides, 1990, 1992, 2006; Trower et al., 1990; Pinker, 1997). Such a module has several key characteristics. First, it is dedicated to the detection of particular types of stimuli as input, rapidly processing a special class of information of particular relevance for survival. Second, when activated, it functions as a motivational system, driving relevant responses (Kavaliers and Choleris, 2001). Third, its output consists of a characteristic set of species-typical behaviors, and engagement in these behaviors plays a crucial role in terminating the activation of the module.

TYPE OF STIMULI THAT ACTIVATE THE SYSTEM

Research on how animals manage the threat of predation illuminates the kinds of stimuli that activate the security motivation system. Animals use subtle, indirect cues of uncertain significance as indicators of potential danger (Blanchard and Blanchard, 1988; Lima and Bednekoff, 1999). Evaluating these indirect cues of potential danger is quite different from recognizing imminent danger, such as the actual presence of a predator, and has been characterized in terms of “labile perturbation factors” (Wingfield et al., 1998) and “hidden-risk mechanisms” (Curio, 1993). In short, the security motivation system is tuned to partial, uncertain cues of potential threat, rather than the recognition of imminent danger.

NATURE OF ACTIVATION OF THE SYSTEM

Studies of the threat of predation show that relatively weak cues readily activate vigilance and wariness (Brown et al., 1999). In addition, this activation ebbs only slowly (Wingfield et al., 1998),

even if no further, confirming cues follow (Masterson and Crawford, 1982; Curio, 1993; Marks and Nesse, 1994). This protracted activation motivates security-related behaviors. In short, weak cues can readily activate the security motivation system, and once activated, it has a protracted half-life and drives behavior.

OUTPUT BEHAVIORS AND TERMINATION OF ACTIVATION OF THE SYSTEM

The resulting acts consist of precautionary behaviors, which include probing the environment, checking, and surveillance to gather further information about any potential risks (Blanchard and Blanchard, 1988; Curio, 1993). They also include corrective or prophylactic behaviors, such as washing, that would lessen the effects of the danger if it were to eventuate. Of particular importance, we have characterized security-related behavior as “open-ended,” meaning that the environment does not normally provide a clear terminating stimulus to signal goal attainment (Szechtman and Woody, 2004). For example, if checking does not reveal the presence of a predator, this is not a clear indication of reduced risk (Curio, 1993); that is, the success of precautionary behavior is a non-event. Consequently, we proposed that it is the engagement in security-related behavior in itself that terminates security motivation. In short, activation of the security motivation system elicits precautionary behavior, and the system uses these actions themselves as the terminator of the motivation.

NEURAL AND PHYSIOLOGICAL BASIS AND EMPIRICAL EVIDENCE FOR THE SECURITY MOTIVATION SYSTEM

We have proposed a fairly detailed neuroanatomical-circuit model for the security motivation system, which is based on functional loops consisting of cascades of cortico-striato-pallido-thalamo-cortical connections (Alexander et al., 1986; Brown and Pluck, 2000), with feedback connections from the brainstem to terminate activity in these loops (Szechtman and Woody, 2004; Woody and Szechtman, 2011). We have also described the proposed physiological mechanisms of the security motivation system, which involve regulation of the parasympathetic nervous system and activation of the hypothalamic-pituitary-adrenocortical (HPA) axis (Woody and Szechtman, 2011).

We have demonstrated that activation and subsequent deactivation of the security motivation system can be tracked both with subjective ratings (e.g., anxiety and urge to engage in precautionary behavior) and also physiological changes, especially respiratory sinus arrhythmia (RSA; Porges, 2001, 2007a,b), based on heart-rate variability (Hinds et al., 2010). Using these measures, we have conducted a series of experiments that support the hypotheses that the security motivation system has the aforementioned characteristic properties. First, we have shown that the system is responsive to relatively weak, uncertain cues for potential danger (Hinds et al., 2010, Experiment 1). Second, we have shown that activation of the system, in the absence of subsequent precautionary behavior, is persistent over time, dissipating only very slowly (Hinds et al., 2010, Experiment 2). Third, we have shown that corrective behavior, such as hand washing in response to uncertain cues for contamination, deactivates the system (Hinds et al., 2010, Experiment 1). In contrast to the deactivating effect of corrective behavior, the security motivation

system, once it has been activated by uncertain cues, is relatively unresponsive to clear cognitive information that disconfirms the potential threat (Hinds et al., 2010, Experiment 3). This finding supports the hypothesis that the system is action-oriented, and engagement in some kind of precautionary behavior plays a crucial role in turning it off.

In a somewhat parallel series of experiments, we have tested our hypothesis that OCD represents a dysfunction of the security motivation system (Szechtman and Woody, 2004; Woody and Szechtman, 2005). It is well known that the content of OCD revolves around issues of potential danger, such as the threat of contamination or physical harm to oneself or close others (e.g., Reed, 1985; Wise and Rapoport, 1989). We hypothesized that in OCD patients, security motivation is activated in a manner that is reasonably similar to how it is activated in non-patients; however, in OCD patients, subsequent precautionary behaviors fail to turn this activation off in the usual fashion. Thus, once activated, OCD patients remain preoccupied with issues of potential danger for a protracted period of time and repeat the precautionary behaviors over and over again, in an attempt to deactivate the concerns. Our experimental data support this hypothesis that OCD is a stopping, rather than a starting, problem (Hinds et al., 2012). In particular, exposure to uncertain cues for contamination activates the security motivation system similarly in OCD patients and control non-patients, as indexed by both subjective measures and RSA. However, a subsequent fixed period of hand-washing, which returns the non-patients to baseline, has no significant effect on the activation levels of the OCD patients.

IMPLICATIONS OF THE SECURITY MOTIVATION SYSTEM

The security motivation system would be expected to have some important characteristics that are common to evolved, special-purpose modules. One important characteristic of such modules is that they tend to be encapsulated, operating relatively automatically and autonomously, and their internal computations are not accessible to introspection (Fodor, 1983). That is, they operate largely in the background, apart from the realm of volitionally directed formal logic, and their outputs become evident to the individual intuitively as feelings.

This distinction between a feeling-based system and rational analysis may not always be readily evident in everyday circumstances, because normally the two kinds of output are reasonably well aligned. However, the distinction becomes extremely striking in OCD. OCD patients feel driven to continue their obsessive concerns about potential danger and to repeat precautionary behaviors, such as checking or washing, even though at a rational level they find these concerns and behaviors to be excessive, illogical, and even absurd (Hollander et al., 1996). Indeed, OCD demonstrates that an intuitive, feeling-based module like the security motivation system is very powerful and can override the rational control of behavior.

The relatively automatic, intuitive, feeling-based operation of the security motivation system corresponds with what Kahneman (2011) has termed System 1, in contrast to the formal logic of System 2. What is important to appreciate is that even though the intuitive feelings generated by the security motivation system are vivid, immediate, and phenomenologically compelling

to the individual, they are not the same as objective reality, nor are they necessarily closely aligned to conclusions derivable from formal logic. They are, in essence, intuitions that worked well in our remote past but may have limited applicability to any specific, current set of circumstances.

DOES THIS BIOLOGICAL SYSTEM INFLUENCE POLICY-MAKING ABOUT SECURITY IN IMPORTANT WAYS?

The nature of the security motivation system may have important implications for policy makers wishing to involve others, such as the public, in the detection and appraisal of potential threat, as well as to shape their perceptions and get their support for policy initiatives. Even though the security motivation system is sensitive to the detection of slight, partial, uncertain cues, it evolved in such a way that it is tuned more to certain types of stimuli, but not others. It seems clear that the security motivation system is particularly sensitive to concrete and surprising or novel changes in the environment, and relatively insensitive to relatively abstract and gradual changes (which can become familiar and therefore lack novelty). Thus, for example, hearing some details of the latest terrorist attack, even if it occurred at a distant location, is likely to much more readily elicit activation of the security motivation system than is information about global warming, which is relatively abstract and involves very slow change. In addition, because activation of the security motivation system leads to probing for further information, there is a positive feedback cycle in which further concrete details are added, magnifying the initial difference.

Let's examine the case in which it seems relatively difficult influence others to take potential threats seriously, such as global warming. We would advance the hypothesis that for stimuli to be regarded as possible indicators of potential threat, they must elicit the feeling of a potential threat—that is, anxiety, and wariness, which is the indication that the security motivation system is activated. In other words, if the indicators of a putative potential threat fail to evoke the emotional resonance of potential threats, then the potential threat in question will not be perceived as credible. Because the cues for the potential threat of global warming are abstract, distant, and involve very gradual change, they do not resemble the types of cues the security motivation system is designed to respond to. We would suggest that this is why the issue strikes many people as “academic” or merely political—the relevant cues lack the feeling of potential threats, because they do not readily activate the security motivation system. One solution may be to use the arts to help supply the missing emotion. This is a possibility that is currently being explored in many ways by artists—film-makers, painters, writers, and so on—and directors of art museums; the idea, in the words of a director of New York's Museum of Modern Art, is to “touch and disturb” people and get them engaged (Economist, 2013, July 20–26).

The opposite type of case is one in which stimuli too readily activate the security motivation system, as with some terrorist incidents, in which the attention-grabbing qualities of some potential dangers may have little relation to and even interfere with objective analysis of their severity or likelihood. To inject these more abstract considerations into the operation of the security motivation system requires connecting System 2, which

handles abstract ideas, to System 1, which is based on concrete stimuli. We would hypothesize that to be effective, information putting potential threats into a broader critical perspective needs to come early, prior to exposure to the potential-threat stimuli. According to our model of the functional components of the security motivation system (Szechtman and Woody, 2004), such information can come into play at the stage of appraisal of potential danger, which integrates internal factors, such as plans, with external factors, such as concrete stimuli. In contrast, our work suggests that once the security motivation system is activated, it is not affected much by further cognitive information, but instead becomes highly action-oriented, driving, for example, checking and corrective behaviors rather than reappraisal (Hinds et al., 2010).

Of course, the security motivation system theory may have implications for policy-makers themselves, rather than simply those they hope to influence. For everyone, this system is intuitive and feeling-based, operating at least somewhat independently of rational analysis. Because the emotions that the system generates evolved to address crucial survival issues, they are powerful and strongly motivating. Thus, it is natural for decision-makers engaged with an issue of potential danger to be guided by their “gut feelings,” which are more vivid and pressing than the details of rational analysis. Unfortunately, feelings of potential threat (wariness and anxiety) are likely to map imperfectly onto the reality of potential threat. In a related vein, Schneier (2008) pointed out: “Security is both a feeling and a reality. And they are not the same.” A rational analysis of potential danger would need to take account of probabilities and other statistical information, but the intuitive operations of the security motivation system do not work this way. According to Suskind (2006, p. 62), as Vice President, Dick Cheney took the position that potential threats should not be evaluated according to “our analysis, or finding a preponderance of evidence,” but instead by a “one-percent doctrine”: if there is any chance of the reality of the threat, “we have to treat it as a certainty in terms of our response.” This position has a gut-intuitive appeal, in that fragmentary cues suggesting any potential of threat activate security motivation, which in turn naturally drives action. However, this is unlikely to be an adequate basis for making very difficult decisions about how to allocate resources to security-related behavior vs. other important goals.

We would also hypothesize that work circumstances that divide up the tasks involved in managing potential threats may tend to disrupt the stopping mechanism of the security motivation system—because, for example, the policy makers do not get to carry out any of the protective actions themselves. We would propose that this problem can lead some agencies working on security issues to function in a way that is analogous to our characterization of OCD—namely, too much can seem like too little (Hinds et al., 2012). Consider, for instance, that between 2001 and 2013, the Foreign Intelligence Surveillance Act (FISA) court of 14 judges in the USA approved 20,909 requests to monitor individuals or search properties, and turned down only 10. Recently, they apparently ruled that all American phone calls should be considered “relevant” to the investigation of terrorist threats (Economist, 2013, July 13–19). The reason why everything may come to seem relevant may be that the stopping function

of the security motivation system is based not on cognitive closure, but instead on concrete action, and those setting policy may not be involved in protective and corrective action at all (e.g., searching and evaluating records).

There are also other implications of the idea that the precautionary behaviors are crucial for turning off security motivation. The security motivation system operates according to what Kahneman (2011) terms System 1 processes. Unfortunately, as Kahneman has very convincingly demonstrated, System 1 is prone to substituting something that has only the form or appearance of a solution in place of a real solution, especially if the better solution would be more difficult. Thus, although turning off the anxiety of security motivation requires action, the details of what is done may not matter as much to the system. Possibly for this reason, policy-making responses to potential threats often seem only to be reactive, rather than proactive. For example, to prevent another shoe-bombing attempt, it is decided that all passengers' shoes must be inspected. Such a prescribed set of actions may be effective in calming security motivation for both policy-makers and the public. However, such a solution seems to ignore the fact

that biological agents (even germs) change strategies, so that what would have worked against them in the past may not do so in the future.

CONCLUSION

The foregoing hypotheses illustrate just a few of the ways in which the security motivation system theory could be used to generate interesting hypotheses for research on the psychology of security-related policy-making. Although these hypotheses need to be evaluated in future research, we hope they provide a convincing case that the security motivation system theory offers a novel, generative framework for advancing our understanding of policy-making processes related to security and potential danger.

ACKNOWLEDGMENTS

The authors' contributions were supported by grants from the Canadian Institutes of Health Research (MOP134450 and MOP-64424) and the Natural Sciences and Engineering Research Council of Canada (RGPIN A0544 and RGPGP 283352-04).

REFERENCES

- Alexander, G. E., Delong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381. doi: 10.1146/annurev.ne.09.030186.002041
- Blanchard, D. C., and Blanchard, R. J. (1988). Ethoexperimental approaches to the biology of emotion. *Annu. Rev. Psychol.* 39, 43–68. doi: 10.1146/annurev.ps.39.020188.000355
- Boyer, P., and Lienard, P. (2006). Why ritualized behavior? precaution systems and action parsing in developmental, pathological and cultural rituals. *Behav. Brain Sci.* 29, 595–613. doi: 10.1017/S0140525X06009332
- Brown, J. S., Laundre, J. W., and Gurung, M. (1999). The ecology of fear: optimal foraging, game theory, and trophic interactions. *J. Mammal.* 80, 385–399. doi: 10.2307/1383287
- Brown, R. G., and Pluck, G. (2000). Negative symptoms: the 'pathology' of motivation and goal-directed behaviour. *Trends Neurosci.* 23, 412–417. doi: 10.1016/S0166-2236(00)01626-X
- Curio, E. (1993). Proximate and developmental aspects of antipredator behavior. *Adv. Study Behav.* 22, 135–238. doi: 10.1016/S0065-3454(08)60407-6
- Economist. (2013, July 13–19). Surveillance: silence in court. *The Economist* 408, 8844, p. 29.
- Economist. (2013, July 20–26). Art about climate change: chilling. *The Economist* 408, 8845, pp. 72–73.
- Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
- Hinds, A. L., Woody, E. Z., Drandic, A., Schmidt, L. A., Van Ameringen, M., Coroneos, M., et al. (2010). The psychology of potential threat: properties of the security motivation system. *Biol. Psychol.* 85, 331–337. doi: 10.1016/j.biopsycho.2010.08.003
- Hinds, A. L., Woody, E. Z., Van Ameringen, M., Schmidt, L. A., and Szechtman, H. (2012). When too much is not enough: obsessive-compulsive disorder as a pathology of stopping, rather than starting. *PLoS ONE* 7:e30586. doi: 10.1371/journal.pone.0030586
- Hollander, E., Kwon, J. H., Stein, D. J., Broatch, J., Rowland, C. T., and Himelein, C. A. (1996). Obsessive-compulsive and spectrum disorders: overview and quality of life issues. *J. Clin. Psychiatry* 57(Suppl. 8), 3–6.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Toronto, ON: Doubleday Canada.
- Kavaliers, M., and Choleris, E. (2001). Antipredator responses and defensive behavior: ecological and ethological approaches for the neurosciences. *Neurosci. Biobehav. Rev.* 25, 577–586. doi: 10.1016/S0149-7634(01)00042-2
- Lima, S. L., and Bednekoff, P. A. (1999). Temporal variation in danger drives antipredator behavior: the predation risk allocation hypothesis. *Am. Nat.* 153, 649–659. doi: 10.1086/303202
- Marks, I. M., and Nesse, R. M. (1994). Fear and fitness: an evolutionary analysis of anxiety disorders. *Ethol. Sociobiol.* 15, 247–261. doi: 10.1016/0162-3095(94)90002-7
- Masterson, F. A., and Crawford, M. (1982). The defense motivation system - a theory of avoidance-behavior. *Behav. Brain Sci.* 5, 661–675. doi: 10.1017/S0140525X00014114
- Pinker, S. (1997). *How the Mind Works*. New York, NY: Norton.
- Porges, S. W. (2001). The polyvagal theory: phylogenetic substrates of a social nervous system. *Int. J. Psychophysiol.* 42, 123–146. doi: 10.1016/S0167-8760(01)00162-3
- Porges, S. W. (2007a). A phylogenetic journey through the vague and ambiguous Xth cranial nerve: a commentary on contemporary heart rate variability research. *Biol. Psychol.* 74, 301–307. doi: 10.1016/j.biopsycho.2006.08.007
- Porges, S. W. (2007b). The polyvagal perspective. *Biol. Psychol.* 74, 116–143. doi: 10.1016/j.biopsycho.2006.06.009
- Reed, G. F. (1985). *Obsessional Experience and Compulsive Behaviour: A Cognitive-Structural Approach*. Orlando, FL: Academic Press, Inc.
- Schneier, B. (2008). *The Psychology of Security*. Available online at <http://www.schneier.com/essay-155.html>
- Suskind, R. (2006). *The One Percent Doctrine*. New York, NY: Simon and Schuster.
- Szechtman, H., and Woody, E. (2004). Obsessive-compulsive disorder as a disturbance of security motivation. *Psychol. Rev.* 111, 111–127. doi: 10.1037/0033-295X.111.1.111
- Tooby, J., and Cosmides, L. (1990). The past explains the present: emotional adaptations and the structure of ancestral environments. *Ethol. Sociobiol.* 11, 375–424. doi: 10.1016/0162-3095(90)90017-Z
- Tooby, J., and Cosmides, L. (1992). "Psychological foundations of culture," in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds J. H. Barkow, L. Cosmides, and J. Tooby (New York, NY: Oxford University Press), 19–136.
- Tooby, J., and Cosmides, L. (2006). The evolved architecture of hazard management: Risk detection reasoning and the motivational computation of threat magnitudes. *Behav. Brain Sci.* 29, 631–633. doi: 10.1017/S0140525X06009538
- Trower, P., Gilbert, P., and Sherling, G. (1990). "Social anxiety, evolution, and self-presentation: an interdisciplinary perspective," in *Handbook of Social and Evaluation Anxiety*, ed H. Leitenberg (New York, NY: Plenum Press), 11–45.
- Wingfield, J. C., Maney, D. L., Breuner, C. W., Jacobs, J. D., Lynn, S., Ramenofsky, M., et al. (1998). Ecological bases of hormone-behavior interactions: The

- "Emergency Life History Stage". *Amer. Zool.* 38, 191–206. doi: 10.1093/icb/38.1.191
- Wise, S., and Rapoport, J. L. (1989). "Obsessive compulsive disorder - is it a basal ganglia dysfunction?" in *Obsessive Compulsive Disorder in Children and Adolescence*, ed J. Rapoport (Washington, DC: American Psychiatric Press), 327–344.
- Woody, E. Z., and Szechtman, H. (2005). Motivation, time course, and heterogeneity in obsessive-compulsive disorder: response to Taylor, McKay, and Abramowitz (2005). *Psychol. Rev.* 112, 658–661. doi: 10.1037/0033-295X.112.3.658
- Woody, E. Z., and Szechtman, H. (2011). Adaptation to potential threat: the evolution, neurobiology, and psychopathology of the security motivation system. *Neurosci. Biobehav. Rev.* 35, 1019–1033. doi: 10.1016/j.neubiorev.2010.08.003
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 24 April 2013; paper pending published: 27 June 2013; accepted: 22 August 2013; published online: 09 September 2013.
- Citation: Woody EZ and Szechtman H (2013) A biological security motivation system for potential threats: are there implications for policy-making? *Front. Hum. Neurosci.* 7:556. doi: 10.3389/fnhum.2013.00556
- This article was submitted to the journal *Frontiers in Human Neuroscience*. Copyright © 2013 Woody and Szechtman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Sensing, assessing, and augmenting threat detection: behavioral, neuroimaging, and brain stimulation evidence for the critical role of attention

Raja Parasuraman^{1*} and Scott Galster²

¹ Center of Excellence in Neuroergonomics, Technology and Cognition, George Mason University, Fairfax, VA, USA

² Applied Neuroscience Branch, Air Force Research Laboratory, Dayton, OH, USA

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

Raymond Nickerson, Tufts University, USA

Douglas D. Potter, University of Dundee, UK

*Correspondence:

Raja Parasuraman, Center of Excellence in Neuroergonomics, Technology and Cognition, George Mason University, 4400 University Drive, MS 3F5, Fairfax, VA 22030-4444, USA
e-mail: rparasur@gmu.edu

Rapidly identifying the potentially threatening movements of other people and objects—biological motion perception and action understanding—is critical to maintaining security in many civilian and military settings. A key approach to improving threat detection in these environments is to sense when less than ideal conditions exist for the human observer, assess that condition relative to an expected standard, and if necessary use tools to augment human performance. Action perception is typically viewed as a relatively “primitive,” automatic function immune to top-down effects. However, recent research shows that attention is a top-down factor that has a critical influence on the identification of threat-related targets. In this paper we show that detection of motion-based threats is attention sensitive when surveillance images are obscured by other movements, when they are visually degraded, when other stimuli or tasks compete for attention, or when low-probability threats must be watched for over long periods of time—all features typical of operational security settings. Neuroimaging studies reveal that action understanding recruits a distributed network of brain regions, including the superior temporal cortex, intraparietal cortex, and inferior frontal cortex. Within this network, attention modulates activation of the superior temporal sulcus (STS) and middle temporal gyrus. The dorsal frontoparietal network may provide the source of attention-modulation signals to action representation areas. Stimulation of this attention network should therefore enhance threat detection. We show that transcranial Direct Current Stimulation (tDCS) at 2 mA accelerates perceptual learning of participants performing a challenging threat-detection task. Together, cognitive, neuroimaging, and brain stimulation studies provide converging evidence for the critical role of attention in the detection and understanding of threat-related intentional actions.

Keywords: action understanding, attention, biological motion, brain stimulation, human performance augmentation, neuroimaging, security, threat detection

INTRODUCTION

Rapidly detecting and identifying the movements and actions of other people—biological motion perception—is an important function of many civilian and military operational settings involving surveillance and other security related tasks. For example, cameras mounted in prisons (Tickner and Poulton, 1975) and other sensitive locations (Stedmon et al., 2011), or on unmanned air (Cummings et al., 2007) and ground vehicles (Chen and Barnes, 2012), are increasingly used to provide video or infrared images to remotely located operators. Surveillance images typically show people or vehicles in motion and engaged in various activities. Such information can be used to identify individuals who pose potential threats or to determine the potential for danger in gatherings of large groups of people. The images are examined for possible threats by skilled human observers (Blake and Shiffrar, 2007), by automated systems (Cohen et al., 2008), or by a combination of the two.

Biological motion perception has typically been investigated in psychophysical studies using simple point-light “stick-figure”

movements of the type pioneered by Johansson (1973). More recent studies have examined more complex, naturalistic scenes of people moving or handling objects (Blake and Shiffrar, 2007; Ortigue et al., 2009; Parasuraman et al., 2009; Grafton and Tipper, 2012; Thompson and Parasuraman, 2012). Identifying the mechanisms and neural bases of action observation when people view naturalistic scenes can advance both the theory and practice in threat detection. More broadly, understanding the mechanisms of threat detection can contribute to scientific approaches to security based on human factors/ergonomics (Nickerson, 2010), neuroscience (National Research Council, 2008), and the intersection of these two fields, neuroergonomics (Parasuraman and Rizzo, 2008; Parasuraman, 2011; Parasuraman et al., 2012).

A key approach to improving threat detection is to *sense* when less than ideal conditions exist for the human operator in a particular security environment. The next step is to *assess* the threat detection performance of the human operator with respect to a standard baseline of required capability. Subsequently, and if necessary, methods can be implemented to *augment* the human

operator in case the standard is not met. In this paper we first describe this Sense—Assess—Augment framework in the context of security research and practice in the military. We then examine the perceptual and cognitive mechanisms of biological motion-based threat detection, focusing on the critical role of attention. Neuroimaging studies that point to the influence of attention are also discussed. Finally, given that humans have limited attentional capabilities, we discuss how non-invasive brain stimulation can be used to enhance threat detection and mitigate operator performance decrements.

SENSE-ASSESS-AUGMENT FRAMEWORK

A number of agencies in the military are planning how best to match personnel and emerging advanced technologies that are being rapidly implemented for use in both the civilian and defense sectors. For example, in 2010, the Chief Scientist of the US Air Force released a report outlining the science and technology needs in the 2010–2030 time frame (Dahm, 2010). A key conclusion of that report was that natural human capacities are becoming increasingly mismatched to the enormous data volumes, processing capabilities, and decision speeds that computer technologies either offer or demand. Although humans today remain more capable than machines for many tasks, particularly higher-order decision making and planning, by 2030 machine capabilities may increase to the point that human capabilities will be significantly challenged in a wide array of systems and processes. It is also the case that human operators are being overloaded *today* by data that the new technologies are able to provide at ever increasing speed. Both of these trends mean that humans and machines will need to become far more closely coupled, through improved human-machine interfaces and by direct augmentation of human performance. Focused research efforts over the next decade will permit significant practical instantiations of augmented human performance. These may come from increased use of autonomous systems, from novel human-machine interfaces to couple humans more closely and more intuitively with automated systems, or from direct augmentation of humans themselves. In this paper we focus on the last of these possibilities.

There are two primary questions to ask when deciding how to provide human performance augmentation: *when* to provide

the augmentation and *how* to provide it. The answers to these questions will likely determine if the right augmentation technique is being employed at the right time to produce the desired effects. The Sense-Assess-Augment taxonomy provides answers to these questions. **Figure 1** shows a representation of this taxonomy. The objective is to sense individual and team cognitive or functional state (using behavioral or neural measures, or both), assess the state relative to performance, and if necessary augment performance to optimize mission effectiveness. This taxonomy is being applied to improve human performance by leveraging the integration of several neurocognitive sensing technologies coupled with multiple assessment approaches to provide a robust understanding of the causes of operator performance decrements (Galster and Johnson, 2013). Given a better understanding of the causes for sub-optimal performance, targeted augmentation techniques can be employed to improve individual or team performance.

Threat detection has many features in common with other tasks the Air Force undertakes to defend capabilities in its air, space, and cyberspace operations. Many of the problems associated with information overload are exacerbated in threat detection tasks due to the exponential growth in the number of pictures or full motion videos that must be processed for actionable information. The use of the Sense-Assess-Augment taxonomy allows for the identification of specific bottlenecks that may occur during the information processing of the data that is required for accurate threat detection. It also allows for the augmentation of the operator based on the characteristics of the bottleneck, for example, whether due primarily to issues with divided or sustained attention, or a lack of the ability to discriminate threats regardless of the amount of training. The correct identification of the individual's source for sub-optimal performance will drive what augmentation method should be utilized to enhance and optimize human performance.

In this paper we describe the results of a number of studies of threat detection within the Sense-Assess-Augment framework. We begin with a discussion of behavioral studies that have investigated mechanisms of biological motion perception in relation to threat detection.

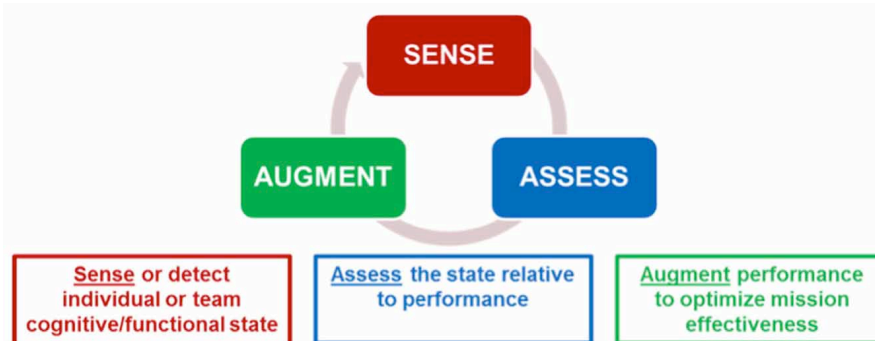


FIGURE 1 | Sense-Assess-Augment framework for optimizing threat detection performance. From Galster and Johnson (2013).

BEHAVIORAL STUDIES

AUTOMATICITY OF BIOLOGICAL MOTION PERCEPTION

Sensing the movement of other biological organisms has played an important role in the survival and evolution of species. Biological threats in natural environments—such as the movements of a predator—can help an animal in the “fight or flight” response. Similarly, the predator uses this ability to sense the motion of its prey. The capacity for biological motion perception appears to be largely present at birth. Infants as young as two days old, for example, show a preference for looking at point-light displays of biological motion as opposed to random motion (Simion et al., 2008). At the other end of the lifespan, older adults, who typically exhibit an age-related decline in the efficiency of processing non-biological moving objects (Gilmore et al., 1992; Jiang et al., 1999), nevertheless have been reported to be as efficient at processing biological motion stimuli as the young (Norman et al., 2004; Billino et al., 2008).

These different lines of evidence would seem to support the notion that biological motion perception is a “primitive” and largely automatic function, although it is also possible that it is learned very early in infancy (2 days). Thornton and Vuong (2004) provided evidence for automaticity in a study using the well-known “flanker” task (Eriksen and Eriksen, 1974). Participants had to determine the direction of movement (left or right) of a central moving human stick figure while flanking figures either moved in a direction congruent or incongruent with the central figure. The time to identify the direction of movement of the central figure was slowed in the incongruent condition, indicating that the flanking movements were processed even though they were outside the focus of attention and irrelevant to the task. Some computational models have also supported the view that biological motion perception occurs automatically through bottom-up visual mechanisms (Giese and Poggio, 2002).

Such a view would suggest that the detection of the threatening actions of people and objects should be rapid and efficient, which it often is under optimal viewing conditions. Yet this appears not to be always the case, as indicated by the performance of people watching complex moving images under challenging naturalistic conditions, such as an unmanned air vehicle operators watching for threat-related activity in video imagery. One possibility is that while simple movements and actions may be largely processed automatically in a bottom-up manner, biological motion may be influenced by top-down factors under more demanding viewing conditions, such as when images are degraded due to sensor or communication channel noise, partially obscured by other movements, when threats only occur rarely, and when other factors are present that place demands on operator attention. When viewed in ideal conditions, many movements and actions may be processed with little or no effort. However, under non-optimal or degraded viewing conditions, attention may be required to resolve ambiguity or enhance perceptual processing so that threats are detected.

EFFECTS OF DIVIDED ATTENTION

Nakayama and Joseph (1998) suggested that many perceptual processes—signal detection, pattern perception, object

recognition, etc.—require attentional resources (Norman and Bobrow, 1975), but typically only to a small degree. Consequently, in order to demonstrate that a perceptual process is attention sensitive, the observer’s attentional resources may need to be depleted to a large extent by a secondary task. Thornton et al. (2002) used this strategy in a dual-task study in which participants had to discriminate the direction of movement of point-light displays of human walkers while simultaneously performing a highly demanding secondary task—detecting changes in the orientation of four rectangles that surrounded the walkers. The secondary task was presented either with the moving walkers or with noise dots consisting of scrambled motion. The dual-task performance decrement was significantly greater for the walkers than for the noise stimuli, suggesting that determining the direction of movement of human walkers requires a global spatial integration process that is attentionally demanding. Moreover, increasing the interval between successive frames of the moving stimuli—thus making integration of motion information over time more challenging—also increased interference from the secondary task. Both sets of findings suggest that perception of biological motion requires attentional resources. In another study from the same group, accuracy in determining the orientation of point-light actions was found to be inversely correlated with the amount of interference participants exhibited on the Stroop color-word task, a well-known measure of the ability to control attention (Chandrasekharan et al., 2010).

The results of these behavioral studies indicate that attention plays a role in the perception of biological motion, such as determining the direction of movement of human walkers. But threat detection involves more than identifying movements: the *intent* behind the movements, or *action understanding*, must also be identified. Not all movements constitute a threat, only those associated with specific intentional actions aimed at other individuals. Biological motion-based threats could involve movements made by another person, actions performed on an object, or some combination. If attention is necessary to perceive biological motion, especially under less than optimal conditions, is it also required to understand these actions?

EFFECTS OF SUSTAINED ATTENTION

Parasuraman et al. (2009) examined the issue of the role of sustained attention in action understanding. Participants viewed videos of a person’s hand reaching to grasp either a gun or a similarly shaped object (a hairdryer) (**Figure 2**). The actor (whose face was not shown) grasped the gun or hairdryer in a manner compatible with using either object (utilization intent) or in such a way that it could not be used but only moved from one location to another (transport intent). The object could appear either in the left or right visual field and could point either left or right. Participants were asked to detect a particular target intentional action or threat that occurred infrequently—grasping the gun to use it to fire in a specific direction. All other movements were classified as non-targets or non-threatening events. Participants performed the task over a 22 min period under two conditions, with very low image degradation, so that movements were clearly perceptible, or with high image degradation that made the detection of intent more difficult.

When participants viewed videos that were minimally degraded, they were highly accurate in threat detection and showed no decline in performance over time on task. Thus, in this condition, performance was insensitive to any waning of attention over the 22-min duration of the task. However, when the images were degraded, there was a significant decline in hit rate as a function of time—that is, participants exhibited a vigilance decrement (**Figure 3**). This finding is consistent with other findings that vigilance decrement is increased for targets that are difficult to discriminate (Warm et al., 2008). Furthermore, analysis of the distribution of false alarms—incorrect threat present responses made to non-targets—showed that the requirement to sustain attention over a long period impaired participants' understanding of action intent: false alarms were more frequent for events associated with the wrong intention (e.g., grasping the gun to transport it instead of grasping the gun in order to fire it) than for other non-target events (e.g., using the hairdryer). Thus, the vigilance decrement could not be attributed to participants letting their minds wander (Robertson et al., 1997), as this would predict a random distribution of false alarms over non-target types. Rather, threat detection required effortful allocation of attentional resources, which became depleted over time (Warm et al., 2008).

These findings are consistent with the view that attending to the meaning of an observed action, such as the intention behind

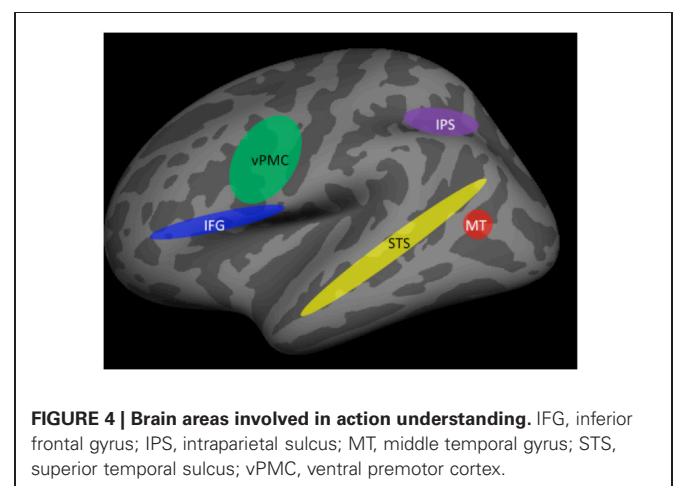
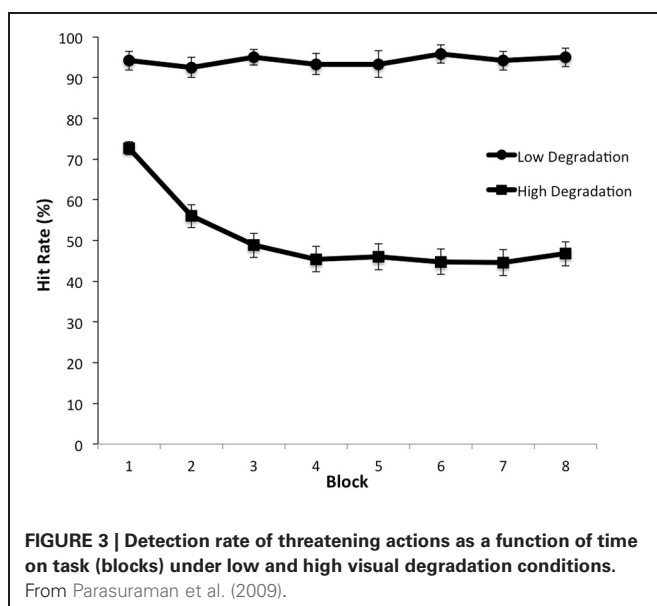
the action, is demanding if the stimuli are difficult to discriminate. The role of attention in action recognition is therefore not restricted solely to detecting or discrimination of a specific human action. Instead, if the movements and actions of people and objects occur under degraded viewing conditions, the decoding of inferences based on observed actions is also attentionally demanding. Attention is known to modulate neural activity in brain networks controlling different perceptual processes (Posner and Petersen, 1990; Petersen and Posner, 2012). Activity in brain regions responsible for biological motion perception should therefore also be affected by allocation or withdrawal of attention. We turn to such evidence next.

NEUROIMAGING STUDIES

THE ACTION UNDERSTANDING NETWORK

Neuroimaging studies using fMRI have revealed that a number of cortical regions within the dorsal and ventral visual processing pathways are associated with biological motion perception and action understanding. **Figure 4** shows the major components of the associated cortical networks. It should be noted that while the specific functions that each of these cortical areas mediate have been identified, the coordination and relative timing of neural activity between cortical areas is a continuing topic of research.

Initial encoding of biological motion occurs in regions of the posterior inferior temporal sulcus, in particular the medial temporal area (MT) (Grossman and Blake, 2002; Thompson et al., 2005). It is thought that the processing of features that compose an action may begin in this area, but that action recognition requires integration across features that is carried out in higher-order visual areas, in particular the superior temporal sulcus (STS). The STS is viewed as a critical brain structure for the recognition of human actions (Grossman and Blake, 2002). This was first demonstrated in single-unit recording studies in monkeys (Perrett et al., 1985) and subsequently confirmed in fMRI studies in humans (Grossman and Blake, 2002; Puce and Perrett, 2003). More importantly, the necessity of STS for action understanding was established in a large-sample study of stroke patients with unilateral lesions of STS and inferior frontal cortex but not of other brain regions (Saygin, 2007).



Additional evidence that the STS is important for the formation of action-specific representations comes from studies using the fMRI adaptation technique, in which neural responses to repeated stimuli that belong to the same category are compared to those to novel stimuli (Grill-Spector et al., 2006). Using this method, Grossman et al. (2010) showed that adaptation to repeated actions in STS was independent of the angle from which the actions were viewed, indicating that the STS is associated with the formation of higher-order representations of actions.

Additional processes must supplement the encoding of biological motion and the representation of actions for action *understanding* to occur. These include information about objects being used by another person [e.g., such as a gun or hairdryer as in the previously-described study by Parasuraman et al. (2009)]. Contextual information about the setting, or prior knowledge, are other factors that will influence understanding the meaning of another person's actions and of inferring their intent, including the possibility of threat. Brain regions outside the primary biological motion/action representation regions of the STS appear to be associated with such intention understanding. They include the inferior parietal and inferior frontal cortex. For example, in a study examining neural activity associated with grasping movements, Hamilton and Grafton (2006) showed that activation of the inferior parietal lobule (IPL) and the anterior intra-parietal sulcus (aIPS) was associated with processing the goal of the observed grasping action, rather than the movement kinematics of the action. In addition, the ventral premotor cortex (vPMC) and the inferior frontal gyrus (IFG) also appear to code the intention behind specific actions of others (de Lange et al., 2008; for a review, see Grafton and Tipper, 2012).

While there is good evidence that the brain regions shown in **Figure 4** are involved in the encoding, representation, and understanding of the actions of others, how these regions interact together, their relative timing of activation, and the effects of attention, are all not fully understood and are current topics of research. One approach to examining the coordination and relative timing issues is to supplement fMRI with electrophysiological methods such as EEG and MEG that have higher temporal resolution than fMRI. This method was used by Ortigue et al. (2009), who examined fMRI and high-density ERPs during performance of a version of the gun/hairdryer task described previously in the study by Parasuraman et al. (2009). Participants were instructed to attend to a series of 3 s-video-clips displaying a hand using or moving either a gun or hairdryer. They were required to respond rapidly (within 1 s) at the end of the last clip to indicate whether the action was consistent with an intention to use (e.g., fire the gun) or transport the object (e.g., move the hairdryer). The fMRI adaptation technique was also used, so that successive trials either repeated a hand-object interaction that reflected the same intention (e.g., use, use) or a different intention (e.g., use, transport). ERPs were recorded using the same event sequence in a separate session. Ortigue et al. (2009) found that compared to when the intentional action was repeated, novel intentions were associated with greater activation in the STS, IPS, and IFG, the main components of the action understanding network shown in **Figure 4**. The network for understanding intentions extends beyond earlier visual processing areas involved in feature detection (e.g., object

shape and size discrimination). In addition, ERP analysis showed that repeated and novel intentions differed in both early activity (~120 ms) that was localized to the STS and IPS and later activity (~350 ms) that was maximal in the IPS and IFG. These findings suggest that understanding the intent behind the movement and actions of another person, including determination of a threatening intent such as firing a gun, involves a distributed network of neocortical regions. The spatiotemporal dynamics of activation in this network can be specified to a degree. However, is it the case that attention has an influence on components of the network? We turn to this issue next.

EFFECTS OF ATTENTION

There are several ways that attention has been manipulated in neuroimaging studies to examine modulation of stimulus-processing cortical areas and the sources of such modulation (Posner and Petersen, 1990; Petersen and Posner, 2012). Previously we described two methods of increasing the attentional demands of biological motion perception, as suggested by Nakayama and Joseph (1998)—requiring participants to perform a challenging secondary task, or asking them to maintain attention to rarely occurring target stimuli in degraded visual images over a long period of time. Both of these attentional challenges are likely to occur in naturalistic threat detection environments. Another method, related to the dual-task technique, is to present other moving stimuli that do not need to be responded to but which compete for the participant's attention because they overlap with the movements that the participant has to process (O'Craven et al., 1999). The use of overlapping stimuli that appear in the same location also allows one to distinguish effects of attention on higher-order representations of biological motion and action from effects of spatial attention, which strongly modulates activity in widespread brain regions (Corbetta and Shulman, 2002).

Safford et al. (2010) used this method to examine whether attention modulated neural activity (using fMRI and ERPs) in the action understanding network when competing, non-biological motion was simultaneously present. Participants viewed videos of human point light motion (e.g., a person doing jumping jacks) that were superimposed on videos of tool motion. Participants were required to perform a 1-back task on either the biological motion or the tool motion, that is, to detect whenever one type of motion was repeated. Thus the task required participants to pay attention to one category of movement. fMRI revealed that activation of the STS was higher when participants attended to biological motion and was strongly suppressed when participants attended to the tool motion, even when biological motion was present but not task-relevant. The data suggested that attention acts on actions at the level of object-based representations, because the only way to select the human actions when they spatially overlapped the tool motion was by using the specific combination of form and motion that define that action. Safford et al. (2010) also recorded ERPs in the same participants and to the same stimuli in a separate session. Source localization analyses revealed that bilateral parietal and right lateral temporal cortices showed early activity at about 200 ms for both biological and tool motion. However, at about 450 ms, greater neural activity in the

right STS was observed for biological motion. Moreover, this later neural response to biological motion was strongly modulated by attention. The combined use of fMRI and EEG thus revealed the spatiotemporal characteristics of biological motion perception in the human brain.

A recent study by Hars et al. (2011) provided corroborative neuroimaging evidence for the modulating effects of attention on neural activity in brain regions subserving biological motion perception and action understanding. They had participants who were trained gymnasts watch either naturalistic videos of an expert perform acrobatic gymnastic movements or relatively impoverished point-light displays of the same movements, recorded with a motion capture system and from the same expert. EEG was recorded from 64 scalp sites and analyzed in three frequency bands (4–8, 8–10, and 10–13 Hz). Functional connectivity for the supplementary motor area in the 4–8 and 8–10 Hz frequency bands was greater during the less familiar and more attentional demanding point-light display than for the videos. The authors concluded that experts at understanding particular actions nevertheless require attention to understand those actions when they occur under unfamiliar viewing conditions, as with the point-light displays.

BRAIN STIMULATION STUDIES

The fMRI and ERP studies we have described have identified the key brain regions associated with biological motion perception and action understanding and, to a degree, the temporal dynamics of interactions between different parts of this network. We have also shown that attention modulates neural activity in key cortical regions, such as the STS. The source of such attentional modulation is the frontoparietal attention control network (Posner and Petersen, 1990; Petersen and Posner, 2012). This suggests that in cases where action understanding and threat detection is challenging and prone to error, such as those described previously—visually degraded images, obscured or overlapping movements, secondary tasks that must be performed, etc.—stimulation of the attention control network might be a possible method to boost performance, consistent with the sense-assess-augment framework described previously.

Skill in threat detection typically develops only after extensive training. For example, intelligence analysts looking at satellite imagery for threats or security officers examining surveillance videos of people for suspects may require many months or years to develop their expertise. At the same time, the number of operational settings that demand skilled surveillance operators is increasing day by day. Hence, validated augmentation methods that can accelerate learning and enhance performance in threat detection will meet a critical need.

There are many different techniques that are available for augmenting human perceptual and cognitive performance. These include neuropharmaceuticals or implants to improve alertness or memory (Mackworth, 1965; Warburton and Brown, 1972; Lynch, 2002). Selecting persons based on their genotype (Parasuraman, 2009), or even genetic modification itself, are other somewhat futuristic possibilities. While such methods raise many ethical issues (Farah et al., 2004) and may be questionable to some, potential adversaries may be entirely willing to make

use of them without reservation. Developing acceptable ways of using science and technology to augment human performance will become increasingly essential for realizing the benefits that many technologies afford. The current technical maturity of various approaches in this area varies widely, but significant steps to advance and develop early implementations are possible now and over the next decade.

A newly emerging augmentation method is to use non-invasive brain stimulation to modulate neuronal activity. There are many such brain stimulation methods, but the two that have received the greatest empirical scrutiny are Transcranial Magnetic Stimulation (TMS) and transcranial direct current stimulation (tDCS) (Utz et al., 2010; McKinley et al., 2012; Nelson et al., 2013). The latter method uses small DC electric currents (typically 1–2 mA) that are applied to the scalp of the participant either before or during the performance of a cognitive or motor task. Brain stimulation at these current levels is safe for use in healthy subjects for up to about 30 min of stimulation (Bikson et al., 2009). The mechanism by which tDCS influences brain function is not precisely known, but is thought to involve alteration of the electrical environment of cortical neurons, specifically small changes in the resting membrane potential of neurons, so that they fire more readily to input from other neurons (Bikson et al., 2004). A positive (anodal) polarity is typically used to stimulate neuronal function and enhance cognitive or motor performance. Conversely, a negative (cathodal) polarity is used to inhibit neuronal activity.

A number of tDCS studies have shown that it is possible to enhance human performance through the application of low-level DC current to the scalp while participants are engaged in simple perceptual, cognitive, and motor tasks (see Utz et al., 2010, for a review). Recently, Pavlidou et al. (2012) also reported improvement in discrimination of point-light stimuli depicting human and animal motion with tDCS of premotor cortex. However, they also reported that tDCS increased false alarms in their discrimination task, so that it is unclear whether tDCS can reliably enhance perceptual sensitivity (in the signal detection theory sense; Green and Swets, 1966), or whether it just lowers the threshold for detection. If the latter were true, it would not support the potential use of tDCS for augmenting threat detection. If both correct and false reports of threat increase with tDCS, threat detection efficiency would not be increased. Moreover, to evaluate whether tDCS can be an effective augmentation technique for threat detection, it should be examined in threat detection tasks with complex targets and naturalistic scenes. Finally, for tDCS to be a viable augmentation technique, its effects should not be transient but should last for some time, preferably for hours if not days.

A recent study by Falcone et al. (2012) addressed these issues. They examined whether tDCS would improve performance in a complex threat detection task and thereby accelerate learning. Signal detection theory analysis was used to examine effects of brain stimulation on perceptual sensitivity independently of bias. Furthermore, retention of any tDCS benefit on threat detection was assessed by testing participants immediately following and 24 h after brain stimulation. Participants were shown short videos of naturalistic scenes containing movements of soldiers and civilians that were taken from the “DARWARS Ambush”

virtual reality software (MacMillan et al., 2005). Half of the scenes included possible threats that participants had to detect, while the other half did not. Examples of threats and non-threats are shown in **Figure 5**. Participants were only told that they were to determine whether or not there was a threat present in the image, without being provided specific details as to what types of possible threats were present. For example, **Figure 5** (top) shows a threat involving a plainly clothed civilian with a concealed weapon behind his back (in his belt). The corresponding non-threat is shown in **Figure 5** (bottom). Other examples of threats were a sniper about to fire from a hidden location or a civilian sneaking up behind military personnel. In all cases, non-threats showed the same elements of the scene but without the critical object or movement that constituted the threat. Threat stimuli were subtle enough to be missed on first viewing but could be better identified with training.

During training participants were required to make a button press within 3 s of stimulus onset to indicate whether the scene contained a threat or a non-threat. After each response a short feedback video was presented for all four outcomes: hit, miss, false alarm, or correct rejection. If a threat was present

and the participant correctly reported it (a hit), the movie showed the scene progressing without harm and simultaneously a computer-generated voice-over complimented the participant. If a threat was present in the image but the participant missed it, the feedback movie showed the consequence of the failure to detect the threat (e.g., vehicle explosion, friendly casualty, building being destroyed). On a non-threat trial, if the participant responded that a threat was present (false alarm), the voice-over chastised the participant. Finally, if the participant correctly indicated that no threat was present on a non-threat trial (correct rejection), the voice-over praised the participant for correct response. None of these feedback videos provided specific information as to the identity of the threats. Participants were given four training blocks of 60 trials each. Each training block contained 60 trials, approximately half of which contained threats, and lasted 12 min. Test blocks were given before and after training and were similar to training blocks, except that no feedback was given after each response.

Anodal tDCS was applied to the electrode site F10 in the 10–10 EEG system, over the right sphenoid bone. The cathode was placed on the contralateral (left) upper arm. The site of the anode was selected based on previous fMRI results showing that this region of the frontal cortex was the primary locus of neural activity associated with performance of this task (Clark et al., 2012). This brain region is also part of the frontoparietal attention network. Hence, Falcone et al. (2012) reasoned that stimulation of this region with tDCS could serve to provide additional top-down attention control signals to the action understanding network and hence boost threat detection performance. Participants were randomly assigned to either active (2 mA current) or sham stimulation (0.1 mA) from the tDCS unit for a total of 30 min during the first two training blocks, beginning 5 min before the training started.

Figure 6 shows the results for the perceptual sensitivity measure d' . Compared to the 0.1 mA sham stimulation control, stimulation with 2 mA tDCS increased perceptual sensitivity in detecting targets and accelerated learning. As expected, performance was near chance in both groups at the beginning of training. However, the performance gain with tDCS

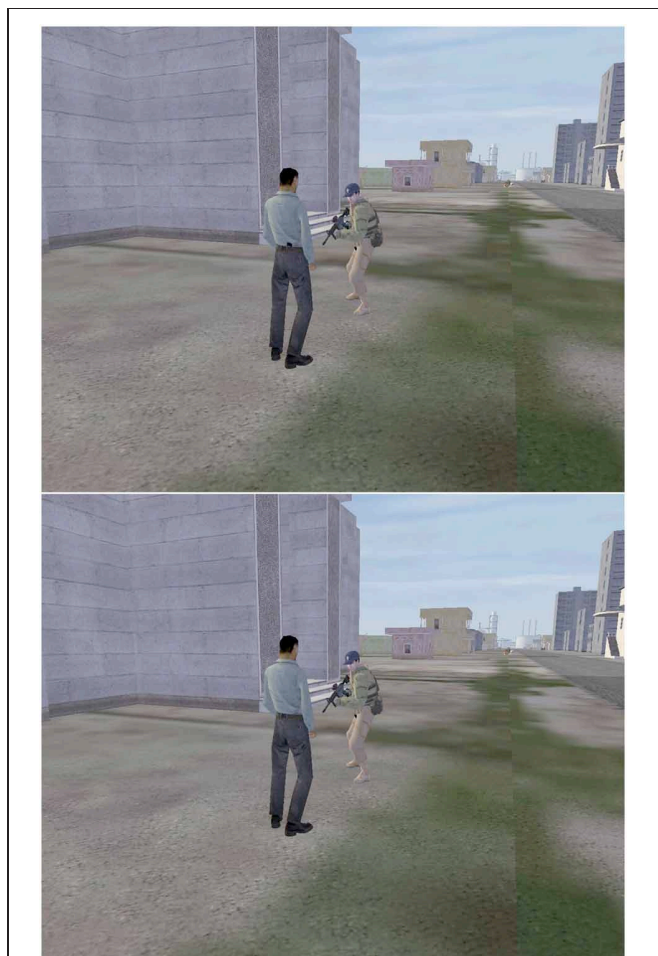


FIGURE 5 | Examples of images indicating threat (top) and non-threat (bottom) situations. From Falcone et al. (2012).

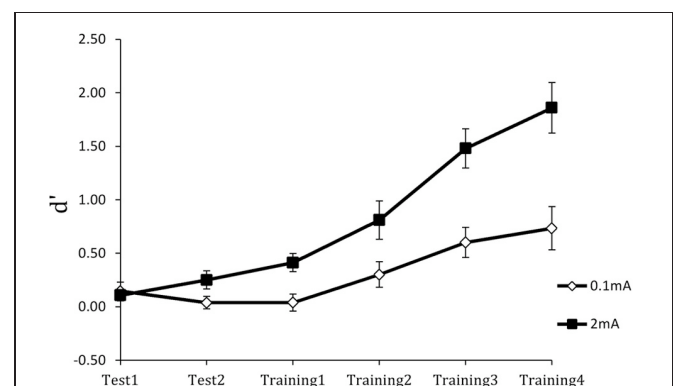
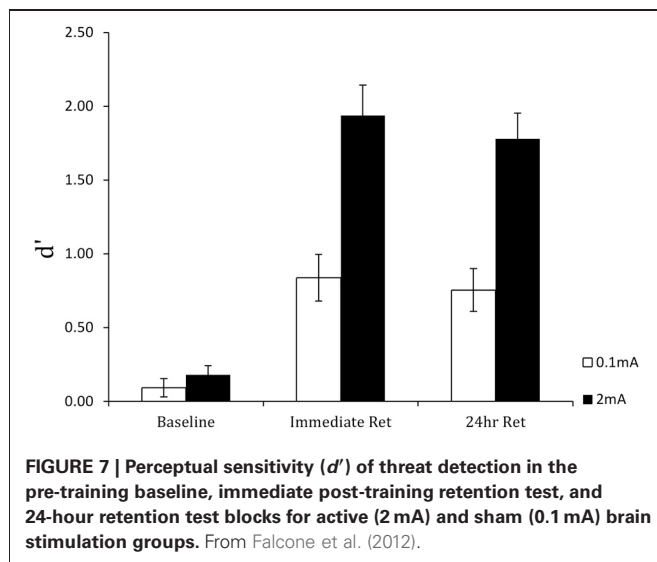


FIGURE 6 | Perceptual sensitivity (d') of threat detection across test and training blocks for active (2 mA) and sham (0.1 mA) brain stimulation groups. From Falcone et al. (2012).



was extensive: on completion of training, participants in the active stimulation group had more than double the perceptual sensitivity of the control group. Furthermore, there were no group or training effects on the response bias measure β , indicating that tDCS improved the actual efficiency of threat detection. Finally, the performance enhancement was maintained for 24 h, as shown in **Figure 7**. Following cessation of brain stimulation training, threat detection sensitivity remained at a high level (immediate retention). Furthermore, while there was some forgetting when participants returned for testing a day later, 24-h retention remained relatively high. This last finding bodes well for the use of tDCS as a training method with potentially lasting effects, although retention over longer periods of days and months will need to be demonstrated.

CONCLUSIONS

Civilian and military operations in the field of security depend on efficient interaction between technological systems and their human operators. Although the final decision in security-related tasks such as threat detection is typically placed in human hands, machine detection and analysis represent important inputs that are used by human decision makers. Thus, the overall efficiency of the human-machine system depends on the cognitive and affective characteristics of human operators. In this paper we

have proposed that improving threat detection in these environments requires a number of steps. First, analysts must sense when less than ideal conditions exist for the human operator in a threat detection task. Second, threat detection performance in that condition must be assessed relative to an expected standard. Third, augmentation methods must be applied if the standard is not met.

Behavioral and neuroimaging studies of sensing and assessment of humans performing threat detection tasks show that attention plays an important role in action identification and understanding. Attention is critically important when operators have to view images that are obscured by other objects or movements, or are visually degraded, when other tasks compete for the operator's attention, or when threats occur infrequently over a prolonged period of surveillance—all features that are characteristic of security-related operations.

Neuroimaging studies reveal that action understanding recruits the superior temporal cortex, intraparietal cortex, and inferior frontal cortex. Within this network, attention modulates activation of the STS and middle temporal gyrus. The dorsal frontoparietal network may provide the source of attention-modulation signals to action representation areas. If sensing and assessment of the human operator reveals attention to be a limiting factor in threat detection, stimulation of the attention network provides a method for augmenting performance. tDCS represents one such augmentation method. tDCS of the frontoparietal network boosts top-down attention control signals that can enhance the detection and identification of threat-related actions.

The cognitive, neuroimaging, and brain stimulation studies we have described provide converging evidence for the critical role of attention in threat detection. As such, these studies are a starting point for a deeper understanding of the neurocognitive mechanisms of threat detection. Although some of the studies we described used naturalistic scenes and videos, additional work needs to be done with even more realistic scenarios and under conditions that better approximate threat detection in real-world security operations.

ACKNOWLEDGMENTS

This work was supported by Air Force Office of Scientific Research grant FA9550-10-1-0385 to Raja Parasuraman and the Center of Excellence in Neuroergonomics, Technology, and Cognition (CENTEC), George Mason University.

REFERENCES

- Bikson, M., Datta, A., and Elwassif, M. (2009). Establishing safety limits for transcranial direct current stimulation. *Clin. Neurophysiol.* 120, 1033–1034. doi: 10.1016/j.clinph.2009.03.018
- Bikson, M., Inoue, M., Akiyama, H., Deans, J. K., Fox, J. E., Miyakawa, H., et al. (2004). Effects of uniform extracellular DC electric fields on excitability in rat hippocampal slices *in vitro*. *J. Physiol.* 557, 175–179. doi: 10.1113/jphysiol.2003.055772
- Billino, J., Bremmer, F., and Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: evidence against general perceptual decline. *Vis. Res.* 48, 1254–1261. doi: 10.1016/j.visres.2008.02.014
- Blake, R., and Shiffrar, M. (2007). Perception of human motion. *Ann. Rev. Psychol.* 58, 47–73. doi: 10.1146/annurev.psych.57.102904.190152
- Chandrasekharan, C., Turner, L., Bulthoof, H. H., and Thornton, I. (2010). Attentional networks and biological motion. *Psychologia* 43, 5–20. doi: 10.2298/PSI1001005C
- Chen, J. Y. C., and Barnes, M. J. (2012). Supervisory control of multiple robots: effects of imperfect automation and individual differences. *Hum. Factors* 54, 157–174. doi: 10.1177/0018720811435843
- Clark, V. P., Coffman, B. A., Mayer, A. R., Weisend, M. P., Lane, T. D. R., Calhoun, V. D., et al. (2012). TDCS guided using fMRI significantly accelerates learning to identify concealed objects. *Neuroimage* 59, 117–128. doi: 10.1016/j.neuroimage.2010.11.036
- Cohen, C. J., Morelli, E., and Scott, K. A. (2008). A surveillance system for the recognition of intent within individuals and crowds. *Proc. IEEE*

- Conf. Technol. Homeland Sec. 5, 559–565.
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Cummings, M. L., Brezinski, A. S., and Lee, J. D. (2007). Operator performance and intelligent aiding in unmanned aerial vehicle scheduling. *IEEE Intell. Syst.* 22, 52–59.
- Dahm, W. (2010). *Report on Technology Horizons: A Vision for Air Force Science, and Technology During 2010–2030. Technical Report.* Washington, DC: United States Air Force.
- de Lange, F. P., Spronk, M., Willems, R. M., Toni, I., and Bekkering, H. (2008). Complementary systems for understanding action intentions. *Curr. Biol.* 18, 454–457. doi: 10.1016/j.cub.2008.02.057
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- Falcone, B., Coffman, B. A., Clark, V. P., and Parasuraman, R. (2012). Transcranial direct current stimulation enhances perceptual sensitivity and 24-hour retention in a complex threat detection task. *PLoS ONE* 7:e34993. doi: 10.1371/journal.pone.0034993
- Farah, M. J., Illes, J., Cook-Deegan, R., Gardner, R., Kandel, E., King, P., et al. (2004). Neurocognitive enhancement: what can we do and what should we do? *Nat. Rev. Neurosci.* 5, 421–425. doi: 10.1038/nrn1390
- Galster, S. M., and Johnson, E. M. (2013). *Sense-Assess-Augment: A Taxonomy for Human Effectiveness. Technical Report.* Wright-Patterson Air Force Base, OH: United States Air Force Research Laboratory.
- Giese, M. A., and Poggio, T. (2002). Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* 4, 179–192. doi: 10.1038/nrn1057
- Gilmore, G. C., Wenk, H., Naylor, L., and Stuve, T. (1992). Motion perception and aging. *Psychol. Aging* 7, 654–660. doi: 10.1037/0882-7974.7.4.654
- Grafton, S. T., and Tipper, C. M. (2012). Decoding intention: a neuroergonomic perspective. *Neuroimage* 59, 14–24. doi: 10.1016/j.neuroimage.2011.05.064
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics.* New York, NY: Wiley.
- Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006
- Grossman, E. D., and Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron* 35, 1167–1175. doi: 10.1016/S0896-6273(02)00897-8
- Grossman, E. D., Jardine, N. L., and Pyles, J. A. (2010). fMR-adaptation reveals invariant coding of biological motion on the human STS. *Front. Hum. Neurosci.* 4:15. doi: 10.3389/neuro.09.015.2010
- Hamilton, A. F., and Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *J. Neurosci.* 26, 1133–1137. doi: 10.1523/JNEUROSCI.4551-05.2006
- Hars, M., Hars, M., Stam, C. J., and Calmels, C. (2011). Effects of visual context upon functional connectivity during observation of biological motions. *PLoS ONE* 6:e25903. doi: 10.1371/journal.pone.0025903
- Jiang, Y., Greenwood, P., and Parasuraman, R. (1999). Age-related reduction in 3-D motion priming. *Psychol. Aging* 14, 619–626. doi: 10.1037/0882-7974.14.4.619
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14, 201–211. doi: 10.3758/BF03212378
- Lynch, G. (2002). Memory enhancement: the search for mechanism-based drugs. *Nat. Rev. Neurosci.* 5, 1035–1038. doi: 10.1038/nrn935
- Mackworth, J. F. (1965). Effect of amphetamine on the detectability of signals in a vigilance task. *Can. J. Psychol.* 19, 104–110. doi: 10.1037/h0082897
- MacMillan, J., Alexander, A. L., Weil, S. A., Littleton, B., and Roberts, B. (2005). *DARWARS: An Architecture that Supports Effective Experiential Training.* Cambridge, MA: Technical Report. Waltham, MA: BBN Technologies.
- McKinley, R. A., Bridges, N., Walters, C. M., and Nelson, J. (2012). Modulating the brain at work using noninvasive transcranial stimulation. *Neuroimage* 59, 129–137. doi: 10.1016/j.neuroimage.2011.07.075
- Nakayama, K., and Joseph, J. S. (1998). “Attention, pattern recognition, and pop out in visual search,” in *The Attentive Brain*, ed R. Parasuraman (Cambridge, MA: MIT Press), 279–298.
- National Research Council. (2008). *Emerging Cognitive Neuroscience and Related Technologies.* Washington, DC: National Academies Press.
- Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., and Parasuraman, R. (2013). Enhancing vigilance in operators with prefrontal cortex transcranial direct stimulation. *Neuroimage*. doi: 10.1016/j.neuroimage.2012.11.061
- Nickerson, R. (2010). Role of human factors and ergonomics in meeting the challenge of terrorism. *Am. Psychol.* 66, 555–566. doi: 10.1037/a0024474
- Norman, D. A., and Bobrow, D. G. (1975). On data-limited and resource-limited processing. *Cogn. Psychol.* 7, 44–60. doi: 10.1016/0010-0285(75)90004-3
- Norman, J. F., Payton, S. M., Long, J. R., and Hawkes, L. M. (2004). Aging and the perception of biological motion. *Psychol. Aging* 19, 219–225. doi: 10.1037/0882-7974.19.1.219
- O’Craven, K. M., Downing, P. E., and Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature* 401, 584–587. doi: 10.1038/44134
- Ortigue, S., Thompson, J. C., Parasuraman, R., and Grafton, S. T. (2009). Spatio-temporal dynamics of human intention understanding in temporo-parietal cortex: a combined EEG/fMRI repetition suppression paradigm. *PLoS ONE* 4:e6962. doi: 10.1371/journal.pone.0006962
- Parasuraman, R. (2009). Assaying individual differences in cognition with molecular genetics: theory and application. *Theor. Iss. Erg. Sci.* 10, 399–416.
- Parasuraman, R. (2011). Neuroergonomics: brain, cognition, and performance at work. *Curr. Dir. Psychol. Sci.* 20, 181–186. doi: 10.1177/0963721411409176
- Parasuraman, R., Christensen, J., and Grafton, S. (2012). Neuroergonomics: the brain in action and at work. *Neuroimage* 59, 1–3. doi: 10.1016/j.neuroimage.2011.08.011
- Parasuraman, R., De Visser, E., Clarke, E., McGarry, W. R., Hussey, E., Shaw, T., et al. (2009). Detecting threat-related intentional actions of others: effects of image quality, response mode, and target cueing on vigilance. *J. Exp. Psychol. Appl.* 15, 275–290. doi: 10.1037/a0017132
- Parasuraman, R., and Rizzo, M. (2008). *Neuroergonomics: The Brain at Work.* New York, NY: Oxford University Press.
- Pavlidou, A., Edwards, M., Lange, J., Schnitzler, A., and Bell, J. (2012). “Anodal stimulation of premotor cortex facilitates the recognition of different forms of movements,” in *Proceedings of the NeuroVisionen 8 Conference*, 95, (Aachen).
- Perrett, D. I., Smith, P. A. J., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., et al. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behav. Brain Res.* 16, 153–170. doi: 10.1016/0166-4328(85)90089-0
- Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years later. *Annu. Rev. Neurosci.* 35, 73–89. doi: 10.1146/annurev-neuro-062111-150525
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42. doi: 10.1146/annurev.ne.13.030190.000325
- Puce, A., and Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 435–445. doi: 10.1098/rstb.2002.1221
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., and Yiend, J. (1997). “Oops!” Performance correlates of everyday attentional failures in traumatic brain injured and normal adults. *Neuropsychologia* 35, 215–226. doi: 10.1016/S0028-3932(97)00015-8
- Safford, A. S., Hussey, E. A., Parasuraman, R., and Thompson, J. C. (2010). Object-based attentional modulation of biological motion processing: spatiotemporal dynamics using functional magnetic resonance imaging and electroencephalography. *J. Neurosci.* 30, 9064–9073.
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain* 130, 2452–2461. doi: 10.1093/brain/awn162
- Simion, F., Regolin, L., and Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proc. Natl. Acad. Sci. U.S.A.* 105, 809–813. doi: 10.1073/pnas.0707021105
- Stedmon, A. W., Harris, S., and Wilson, J. R. (2011). Simulated multiplexed CCTV: the effects of screen layout and task

- complexity on user performance and strategies. *Secur. J.* 24, 344–356. doi: 10.1057/sj.2010.7
- Thompson, J., and Parasuraman, R. (2012). Attention, biological motion, and action recognition. *Neuroimage* 59, 4–13. doi: 10.1016/j.neuroimage.2011.05.044
- Thompson, J. C., Clarke, M., Stewart, T., and Puce, A. (2005). Configural processing of biological motion in human superior temporal sulcus. *J. Neurosci.* 25, 9059–9066. doi: 10.1523/JNEUROSCI.2129-05.2005
- Tickner, A. H., and Poulton, E. C. (1975). Watching for people and actions. *Ergonomics* 18, 35–51. doi: 10.1080/00140137508931438
- Thornton, I. M., Rensink, R. A., and Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception* 31, 837–853.
- Thornton, I. M., and Vuong, Q. C. (2004). Incidental processing of biological motion. *Curr. Biol.* 14, 1084–1089.
- Utz, K. S., Dimova, V., Oppenlander, K., and Kerkhoff, G. (2010). Electrified minds: transcranial direct current stimulation (tDCS) and Galvanic Vestibular Stimulation (GVS) as methods of non-invasive brain stimulation in neuropsychology. A review of current data and future implications. *Neuropsychologia* 48, 2789–2810. doi: 10.1016/j.neuropsychologia.2010.06.002
- Warburton, D. M., and Brown, K. (1972). The facilitation of discrimination performance by physostigmine sulphate. *Psychopharmacology* 3, 275–284. doi: 10.1007/BF00422808
- Warm, J. S., Parasuraman, R., and Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Hum. Factors* 50, 433–441. doi: 10.1518/001872008X312152
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 April 2013; accepted: 26 May 2013; published online: 12 June 2013.

Citation: Parasuraman R and Galster S (2013) Sensing, assessing, and augmenting threat detection: behavioral, neuroimaging, and brain stimulation evidence for the critical role of attention. *Front. Hum. Neurosci.* 7:273. doi: 10.3389/fnhum.2013.00273

Copyright © 2013 Parasuraman and Galster. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Non-invasive brain stimulation can induce paradoxical facilitation. Are these neuroenhancements transferable and meaningful to security services?

Jean Levasseur-Moreau¹, Jerome Brunelin^{1,2} and Shirley Fecteau^{1,3*}

¹ Faculté de Médecine, Centre Interdisciplinaire de Recherche en Réadaptation et en Intégration Sociale, Centre de Recherche de l'Institut Universitaire en Santé Mentale de Québec, Université Laval, Québec City, QC, Canada

² Centre Hospitalier le Vinatier, Université de Lyon, Université Claude Bernard Lyon I, Villeurbanne, Bron, France

³ Berenson-Allen Center for Noninvasive Brain Stimulation, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

Edited by:

Elena Rusconi, University College London, UK

Reviewed by:

Raja Parasuraman, George Mason University, USA

Antonio Oliviero, Hospital Nacional de Paraplejicos, Spain

*Correspondence:

Shirley Fecteau, Pavillon Ferdinand-Vandry, 1050, Avenue de la Médecine, Université Laval, Québec City, QC G1V 0A6, Canada
e-mail: shirley.fecteau@fmed.ulaval.ca

For ages, we have been looking for ways to enhance our physical and cognitive capacities in order to augment our security. One potential way to enhance our capacities may be to externally stimulate the brain. Methods of non-invasive brain stimulation (NIBS), such as repetitive transcranial magnetic stimulation (rTMS) and transcranial electrical stimulation (tES), have been recently developed to modulate brain activity. Both techniques are relatively safe and can transiently modify motor and cognitive functions outlasting the stimulation period. The purpose of this paper is to review data suggesting that NIBS can enhance motor and cognitive performance in healthy volunteers. We frame these findings in the context of whether they may serve security purposes. Specifically, we review studies reporting that NIBS induces paradoxical facilitation in motor (precision, speed, strength, acceleration endurance, and execution of daily motor task) and cognitive functions (attention, impulsive behavior, risk-taking, working memory, planning, and deceptive capacities). Although transferability and meaningfulness of these NIBS-induced paradoxical facilitations into real-life situations are not clear yet, NIBS may contribute at improving training of motor and cognitive functions relevant for military, civil, and forensic security services. This is an enthusiastic perspective that also calls for fair and open debates on the ethics of using NIBS in healthy individuals to enhance normal functions.

Keywords: non-invasive brain stimulation, motor function, cognitive function, transcranial magnetic stimulation, security, transcranial direct current stimulation (tDCS), neuroenhancement

INTRODUCTION

For centuries, we have been trying to improve our motor and cognitive performance in order to augment our security against predators, including our fellow human beings. Numerous ways have been explored to surpass limitations of the human body (e.g., physical training, education, technology, religion). The discovery of the electric neuronal transmission in the early 1800s' has reinforced the belief that one way to enhance motor and cognitive abilities may be to stimulate the brain using electric currents. Considerable progress in modifying electric neuronal activity non-invasively in living humans has been made in the recent years, making it now possible to modulate behaviors. Two of the modern non-invasive brain stimulation (NIBS) methods are the repetitive Transcranial Magnetic Stimulation (rTMS; for a review see Sandrini et al., 2011) and the recently re-discovered transcranial Electrical Stimulation (tES; for a review see Utz et al., 2010; Jacobson et al., 2012). They are now widely used in cognitive neuroscience to study and modulate human behaviors in pathological and normal conditions. Indeed NIBS can be used to characterize causal relationships between brain networks and behaviors. The brain region that is targeted with NIBS is often chosen based on

lesion work and imaging data (e.g., functional MRI) associating a given function with a specific brain network. The general hypothesis postulates that NIBS applied over a specific brain region will modulate level of performance of its associated underlying behavior(s). We can impair and improve normal behavioral performance in healthy individuals with NIBS. When we induce a deficit, this phenomenon is called virtual lesion. When such modulation leads to a functional enhancement, this phenomenon is called paradoxical facilitation. Paradoxical facilitation was first described in patients with brain lesions who performed better than normal subjects on certain tasks (for a review see Kapur, 1996). For example, it has been shown that patients with a right hemisphere lesion displayed shorter response time (RT) than healthy subjects at an attentional task (Ladavas et al., 1990). More recently, it has been reported that NIBS can induce paradoxical facilitation in healthy adults. For instance, normal behavioral performance of healthy subjects can be enhanced following a single session of rTMS or tES. The goal of this paper is to review data indicating that NIBS can promote motor and cognitive functions in healthy volunteers. Further, we frame these data in the context of whether they may benefit security purposes. Specifically,

we will discuss how modulation of motor and cognitive functions with NIBS may promote existing training in security services (e.g., military, police).

OVERVIEW OF NIBS TECHNIQUES

The principle of rTMS is based on Faraday's Law of electromagnetic induction. Brief current pulses are delivered through a coil placed on the subject's scalp (see **Figure 1A**). This generates a magnetic field that penetrates the scalp and skull, inducing a weak electrical current in the brain. rTMS can induce effects that outlast the stimulation period. Low frequencies rTMS ($= 1$ Hz) are known to decrease activity, whereas higher frequencies are assumed to increase activity of the targeted brain area. rTMS can also modulate activity of brain regions interconnected with the targeted area (Hoogendam et al., 2010). Specific mechanisms of these changes remain to be fully determined, but they are widely believed to reflect changes in synaptic potential by modulating depolarization or hyperpolarization states of neurons, leading to changes in long-term depression-like and long-term potentiation-like plasticity.

The principle of tES is quite different from that of rTMS. tES consists of applying electrodes on the subject's scalp (see **Figure 1B**). A weak transcranial Direct Current (tDCS), slow oscillatory Direct Current (so-tDCS, o-tDCS, or tSOS) or Alternating Current (tACS) flows through the brain between the anode and the cathode electrodes. This current flow modulates neural activity in the targeted area(s) as well as connectivity within an interconnected network (Keeser et al., 2011). The effect of tDCS can outlast the stimulation period. The anode is known to increase excitability of the targeted area and the

cathode to inhibit it (Nitsche et al., 2007). Although the exact mechanisms underlying tDCS effects remain unknown, pharmacological studies have highlighted changes in resting neuronal membrane potential and synaptic modifications linked to glutamatergic (NMDA-receptor) and GABAergic activity (for a review see Stagg and Nitsche, 2011). These findings were recently supported in a study using Magnetic Resonance Spectroscopy reporting GABA and Glutamate changes following NIBS (Clark et al., 2011).

The brain area can be targeted non-invasively with rTMS or tES on the subject's scalp based on the 10–20 EEG international system, an anatomical, or a functional MRI image (**Figure 1**).

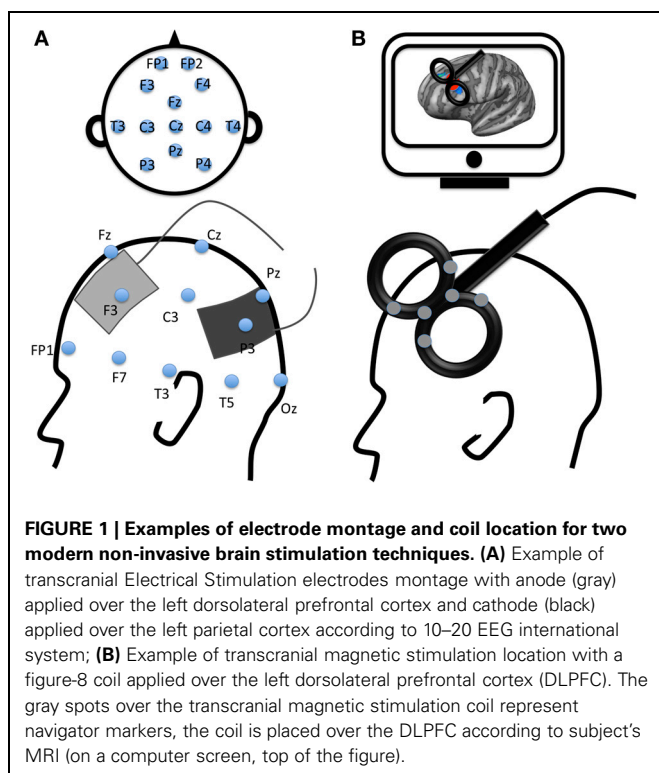
Specific mechanisms of how NIBS induces paradoxical facilitation in healthy individuals are not completely understood yet. Most researchers agree that from a neurophysiologic perspective, NIBS enhances behavioral performances by modulating a dynamic distributed brain network. From a conceptual perspective, three non-mutually exclusive frameworks have been proposed: the entrainment theory, the stochastic resonance model, and the zero-sum theory (for a review see Pascual-Leone et al., 2012). The entrainment theory posits that the brain can be brought into an oscillatory natural state that is known to be associated with a particular function. According to the entrainment model, NIBS mimics brain oscillations and has an effect by entraining the brain's natural state. For instance, applying slow oscillatory tDCS during sleep induced an increase in slow wave sleep and promoted memory in a frequency-specific manner (Marshall et al., 2004). The stochastic resonance model supposes that small amounts of noise injected into a system promote low-level signals leading to enhanced functions within this system. For instance, TMS at low intensity applied over the visual cortex (V5/MT) facilitated detection of weak motion signals, whereas higher intensities impaired detection of stronger motion signals (Schwarzkopf et al., 2011). Finally, the zero-sum theory posits that the brain has a finite power processing. According to this model, if NIBS induces a paradoxical facilitation, the opposite effect will also be observed that is a detrimental behavioral impact. For example, low frequency rTMS applied over the parietal cortex enhanced target detection in the ipsilateral visual hemi-field and worsened detection in the contralateral visual hemi-field (Hilgetag et al., 2001).

STUDIES USING NIBS TO INDUCE PARADOXICAL FACILITATIONS

We will here describe studies indicating that NIBS can enhance performance of healthy subjects on motor and cognitive tasks (attention, impulsivity, risk-taking, working memory, planning, and deceptive capacities).

EFFECTS OF NIBS ON MOTOR FUNCTIONS

The first application of TMS was on the human motor cortex (Barker et al., 1985); and the use of NIBS to promote motor functions in healthy subjects likely represents the richest literature on facilitations induced by rTMS or tES. We will here present studies reporting that NIBS can induce paradoxical facilitation of



motor functions in terms of precision, speed, strength, acceleration endurance, and execution of daily motor task. The majority of these NIBS studies targeted the primary motor cortex (M1), a region known to be involved in motor control (for a review see Schieber, 2001) and motor sequence learning (Penhune and Steele, 2012).

Effects of NIBS on motor precision

Studies tested the ability of NIBS to enhance precision of motor functions in healthy subjects. Buetefisch and colleagues tested the effects of low frequency rTMS applied over the left M1 on precision at motor pointing tasks. They used tasks requiring lower and higher demand of precision for both hands (i.e., ipsilateral and contralateral to the stimulated left M1). Participants receiving active rTMS were more accurate in the task demanding higher level of precision for both hands (with greater accuracy for the ipsilateral than the contralateral one), compared to when they received sham stimulation (Buetefisch et al., 2011). For the lower demand level of precision, no difference in precision was observed between active and sham stimulation conditions. Moreover, Matsuo and colleagues tested the precision in a circle-drawing task before and after healthy volunteers received either active or sham tDCS over the right M1. They found that participants receiving anodal tDCS displayed greater precision of the non-dominant-hand movement. No change in precision was observed when subjects received sham stimulation (Matsuo et al., 2011). Also, these enhanced motor abilities (i.e., deviation area and path length of the task) were observed up to 30 min after the end of the stimulation session (Matsuo et al., 2011).

Effects of NIBS on motor learning

Nitsche and colleagues investigated the effects of tDCS on implicit motor learning using a modified version of the *Serial Reaction Time Task* (SRTT). In this task, participants are instructed to respond as fast as possible on a response pad with four buttons to the apparition of a dot on a computer screen in one of the four positions (each button have to be pushed with a different finger of the right hand). Anodal tDCS was applied in separate groups of participants to different regions contralaterally to the performing hand: M1, premotor, and prefrontal cortices. Participants receiving anodal tDCS over M1 were faster at executing implicitly learned sequences compared to participants receiving tDCS over the premotor or prefrontal areas (Nitsche et al., 2003). This effect was replicated with rTMS. Healthy subjects who received low frequency rTMS over M1 were faster at executing a learned sequence movement with the hand ipsilaterally to the stimulated M1 without affecting performance with the contralateral hand as compared to rTMS applied to the contralateral M1, ipsilateral premotor area, or vertex (Kobayashi et al., 2004). This effect was reported for both M1, with a greater effect for the right M1. The authors reported no effect on accuracy as measured by error rate. The improvement of ipsilateral motor accuracy following 1 Hz rTMS over M1 can outlast the stimulation period up to 30 min (Avanzino et al., 2008). Similar findings were reported using high frequency rTMS applied over the right M1 in right-handed subjects. Subjects were faster and more accurate to execute a learned complex motor task with their left (non-dominant) hand when

they received active rTMS as compared to when they received sham stimulation (Kim et al., 2004). Vines et al. (2008) investigated the effects of tDCS in right-handed healthy participants in a finger sequence performance task. They studied four stimulation conditions: anodal tDCS over the non-dominant M1 coupled with cathodal tDCS over the dominant M1; anodal tDCS over the dominant M1 coupled with cathode over contralateral supraorbital region; anodal tDCS over the non-dominant M1 coupled with cathode over the contralateral supraorbital region, and sham tDCS. The anode applied over the non-dominant M1 coupled with the cathode over the dominant M1 enhanced motor performance in the contralateral (left) hand. Performance was measured by the total number of correct responses calculated as the mean percentage of change in the total number of correct sequential keystrokes at the finger-sequence performance task. The three other stimulation conditions did not lead to significant changes.

Effects of NIBS on muscle might

So far we discussed studies indicating that NIBS can improve motor accuracy, learning, and speed. Other studies suggested that NIBS can also promote motor strength, acceleration, and endurance. This has been shown in upper and lower body parts. Tanaka et al. (2009) investigated the impact of tDCS on leg motor strength at a *Pinch Force Test* in healthy subjects. They found that participants receiving anodal tDCS over the right M1 coupled with cathodal tDCS over the left supraorbital area displayed greater strength compared to those receiving cathodal tDCS over the right M1 coupled with anodal tDCS over the left supraorbital area and sham stimulation. Moreover, these effects outlasted the stimulation period by 60 min. Teo et al. (2011) studied the effects of intermittent theta burst stimulation (iTBS; known to increase excitability) over M1 on movement acceleration. They found that iTBS significantly increased peak acceleration of the thumb abduction movement compared to baseline performance. Cogiamanian et al. (2007) explored the effects of anodal tDCS over the right M1 coupled with cathodal tDCS over the right shoulder on muscular endurance in healthy subjects using a paradigm requiring submaximal isometric contraction of the left elbow flexor. They found that, compared to opposite electrode arrangement or sham conditions, anodal tDCS significantly increased endurance of participants (maximum voluntary contraction).

Effects of NIBS on execution of daily motor task

Boggio and colleagues investigated the effect of tDCS on motor performance at the *Jebsen Taylor Hand Function Test* (JTHF). The JTHF is a widely used task assessing motor activities often performed in daily life (e.g., picking up small objects and placing them in a can, stacking chequers, moving large light or heavy cans). Right-handed volunteers were faster at completing the JTHF with the left (non-dominant) hand when they received anodal tDCS over the right (non-dominant) M1. There was however no change between active and sham tDCS when performed with the right (dominant) hand (Boggio et al., 2006). In another study, participants who received active tDCS (anode over the right non-dominant M1 coupled with cathode over the dominant M1) combined with unilateral motor training and contralateral hand

restraint were faster at the JTHF than those who received sham tDCS combined with unilateral motor training and contralateral hand restraint (Williams et al., 2010). Also, Hummel et al. (2010) tested the effects of tDCS over the left M1 on motor performance measured by the JTHF in healthy subjects. They observed increased overall performance in the time to execute the task in participants receiving active tDCS as compared to when they received sham tDCS. Of note, this study included only elderly participants (mean age of 69 years). Also, these effects were sustained up to 30 min after the end of a single stimulation session. Thus, NIBS appears to decrease speed of motor movement execution.

In sum, NIBS applied over M1 can induce facilitations on various motor aspects such as precision, learning, strength, acceleration, endurance, and execution of daily motor task; and some of these enhancements included hand movements of daily life activities. It also has been proposed that NIBS can be used to enhance motor functions in the context of sportive performance (for a review see Banissy and Muggleton, 2013).

EFFECTS OF NIBS ON ATTENTIONAL SKILLS

Attention is a central cognitive process that is considered as a precursor of a large majority of other cognitive functions. Attention can be described as the capacity of sustainably focus cognitive resources on information while filtering or ignoring non-salient endogenous or extraneous information. Attention processes range from the ability to respond to specific visual, auditory, or tactile stimuli to higher cognitive processes of mental flexibility allowing simultaneous responses to multiple tasks. At the brain level, the attention network is a complex set of interactions implying numerous brain regions, especially the frontal and parietal cortices (Petersen and Posner, 2012) and numerous studies have investigated the effects of NIBS on these regions (Fecteau et al., 2006). We will here present studies reporting NIBS-induced paradoxical facilitation of various attentional processes: sustained attention, focused attention, selective attention, attentional switch, and inhibition.

Effects of NIBS on sustained attention

Sustained attention is the ability to maintain attention (vigilance) for sporadic critical events during long periods of time (Warm et al., 2008). It elicits a large cerebral network including right and left frontal regions. Nelson et al. (2013) measured the effects of tDCS on vigilance performance in military personnel with an air traffic controller simulator. As compared to sham, active tDCS (anodal over the left DLPFC coupled with cathodal over the right DLPFC, as well as the opposite electrode montage) resulted in enhanced accuracy that is an increased number of correct identified targets and a decreased number of false alarms. However, tDCS also resulted in slower RT. Thus, tDCS can improve sustained attention in setting mimicking work environments such as radar operators.

Effects of NIBS on focused attention

Focused attention represents the ability to concentrate the attentional locus toward a specific stimulus. The posterior part of the parietal cortex (PPC) is one of the areas often involved in focused attention, such as detecting a visual target presented in a specific

location (for a review see Corbetta and Shulman, 2002). In order to improve focused attention, numerous studies have applied NIBS over the PPC. For instance, a single session of low frequency rTMS applied over the right or left PPC improved detection of stimuli presented ipsilaterally to the stimulated site. The same rTMS protocol also impaired detection of stimuli presented in the contralateral visual field (Hilgetag et al., 2001). These findings were replicated using a single session of low frequency rTMS applied over the right dorsal PPC (Thut et al., 2005). The authors reported enhanced target detection in the right visual field (i.e., shorter RT) and impaired target detection in the left visual field (i.e., decreased accuracy) after rightward cueing in a visual attention detection task. NIBS appears to promote focused attention using stimuli other than visual as well. Anodal tDCS applied over the right PPC coupled with cathodal tDCS applied over the contralateral deltoid muscle improved attention to auditory stimuli presented contralaterally to the stimulation site, the left auditory field (Bolognini et al., 2011). Thus, NIBS can enhance attention in detecting some auditory and visual targets in healthy subjects.

Effects of NIBS on selective attention

Selective attention is the ability to focus attentional resources oriented toward a given stimulus despite the presence of distracting or competing stimuli. Amongst the regions presumably involved in selective attention (Petersen and Posner, 2012), the right inferior frontal cortex (IFC) and the PPC have been targeted with NIBS to study selective attention. Selective attention can be studied using the *DARWARS Ambush! Threat Detection Task*. This task was initially designed to train US soldiers bound for Iraq. Subjects are presented with threatening and unthreatening targets that are concealed in realistic virtual situations. They are required to detect threatening targets, such as a bomb under a pile of rocks. Clark et al. (2012) evaluated the effects of tDCS on performance at the *DARWARS Ambush! Threat Detection Task*. Subjects who received active tDCS (anodal over either the right IFC or the right PPC coupled with cathodal over the contralateral upper arm) were significantly better than subjects who received sham tDCS. More specifically, they identified a greater number of correct threatening targets and reported a smaller number of false alarms (i.e., identifying unthreatening targets as threatening ones) at the detection task during and after the stimulation session. They were also increasingly faster to complete the task throughout the four training blocks. The same research team conducted another experiment using the *DARWARS Ambush!* (Falcone et al., 2012). First, they replicated their previous findings: anodal tDCS over the right IFC lead to better identification of threatening concealed objects, lesser number of false alarms, and faster learning curve, as compared to sham tDCS. In addition, they observed that this enhanced performance sustained 24 h after the end of the stimulation period. They conducted a third study with a similar design (Coffman et al., 2012). Here, they replicated their initial findings: subjects who received anodal tDCS over the right IFC were better than those who received sham stimulation (i.e., greater identification of threatening concealed objects, lesser number of false alarms, and faster learning curve), as compared to sham tDCS. The observed enhancements of threat detection with tDCS were associated with increased attention (i.e., alerting

attention, decreased RT to detect a cue). Overall, these studies indicate that selective attention can be enhanced by NIBS as shown by improved detection of threats.

Effects of NIBS on attentional switch

Attentional switch is the ability to change attentional resources from a given stimulus to another stimulus. It elicits activity in a large cerebral network including some frontal regions (e.g., the medial frontal cortex and the dorsolateral prefrontal cortex; DLPFC) and the pre-supplementary motor area (pre-SMA; Rushworth et al., 2002). Vanderhasselt and colleagues tested the effects of high frequency rTMS over the right DLPFC on attentional switch using a *Task-Switching Paradigm* (Vanderhasselt et al., 2006). In this task, participants had to respond with their hand to a visual stimulus presented on 8 different locations (pressing one of the 8 buttons) and with their foot to an auditory stimulus (pushing a pedal). Participants had to focus their attention to visual stimuli and then to switch their attention when the auditory stimuli occurred. They were faster at switching their attention when they received active rTMS than when they received sham rTMS.

Effects of NIBS on inhibition

Inhibition is defined here as the ability to refrain from initiating a response to a stimulus. The right IFC, DLPFCs, pre-SMA, M1, and PPC have been targeted with NIBS to diminish RT and improve accuracy of inhibitory control in healthy subjects. For instance, a single session of high frequency rTMS applied over the left DLPFC significantly decreased RT on incongruent trials at the *Stroop word and color task* as compared to sham stimulation (Vanderhasselt et al., 2007). The Stroop task requires participants to name the font color of the visually presented words. Subjects are usually faster at the congruent than the incongruent condition. The congruent condition consists of presenting the word blue written in blue. The incongruent condition consists for example of presenting the word blue written in red. Anodal tDCS over the left DLPFC coupled with cathodal over the contralateral supraorbital area diminished RT at the incongruent condition, compared to sham stimulation (Jeon and Han, 2012).

Inhibition can also be assessed with the *Stop Signal Task* (SST). In this task, an external stimulus signals participants to interrupt an already-initiated motor response. The SST involves a distributed cerebral network including the IFC, the pre-SMA, and the DLPFC of both hemispheres (Sharp et al., 2010). Studies reported that applying anodal tDCS over the right IFC (Jacobson et al., 2011; Ditye et al., 2012) or the right M1 (Kwon et al., 2013) reduced RT at the SST paradigm as compared to sham stimulation. Accuracy at the SST can also be improved with NIBS. The number of correct inhibited responses at the SST was greater in healthy subjects who received anodal tDCS over the pre-SMA as compared to subjects who received active stimulation over the left M1 (Hsu et al., 2011). NIBS can also enhance these inhibitory skills in healthy subjects in a similar task, the *Conners' Continuous Performance task* (Hwang et al., 2010). Here, participants must press a button each time any letter is presented except the "x" letter. The number of commission errors was reduced when subjects received high frequency rTMS

over the left DLPFC as compared to when they received sham stimulation.

The *Flanker Task* is a cognitive paradigm measuring inhibition. Specifically, it characterizes the ability to detect targets in the presence of distracting information. Subjects thus have to inhibit their attention toward distracting stimuli in order to focus their attention on relevant stimuli. Participants who received cathodal tDCS over the right PPC coupled with anodal tDCS over the contralateral supraorbital area were better at detecting targets at this task as compared to subjects who received anodal stimulation over the right PPC coupled with cathodal tDCS over the contralateral supraorbital area and subjects who received sham stimulation (Weiss and Lavidor, 2012). Of note, this NIBS-induced enhancement was not only found in low attentional load, but also in conditions requiring a high level of cognitive process (high-load scenes) when a stimulus is presented along with a great number of distractors.

In sum, NIBS can enhance attentional skills, such as decreasing RT and increasing accuracy at processing visual and/or auditory stimuli in healthy individuals. More specifically, NIBS can improve sustained attention, focused attention, selective attention, attentional switch, and inhibition.

EFFECTS OF NIBS ON IMPULSIVE BEHAVIOR

Some studies suggest that NIBS can modulate impulsive behavior. A rich literature in neuroimaging indicate that the DLPFC is critically involved in impulsive behavior (Rorie and Newsome, 2005). Based on this, the DLPFC has been the main targeted region with NIBS. The effects of NIBS on impulsive behavior have been tested using the *Delay Discounting Task*. This task assesses subject's tendencies to prefer smaller, more immediate rewards or larger, delayed rewards. Healthy subjects who received continuous Theta Burst Stimulation (cTBS; known to decrease excitability) over the right DLPFC choose more often larger, delayed rewards than smaller, immediate rewards, as compared to when they received sham stimulation or iTBS over the right DLPFC (Cho et al., 2012). Finally, in an ecological effort, Beeli and colleagues investigated the effect of anodal and cathodal tDCS over the left and right DLPFC on driving behavior (Beeli et al., 2008). They recorded several behaviors in a driving simulator such as distance from driver ahead and speed. They found that participants receiving anodal tDCS, applied either over the left or the right hemisphere, displayed more careful (less impulsive) driving behavior compared to baseline. As seen in attentional inhibition studies, these results suggest that NIBS can also lead to reduced impulsive behaviors.

EFFECTS OF NIBS ON RISK-TAKING

Risk-taking is known to elicit activity in several regions, critically including the DLPFC according to neuroimaging studies (Rao et al., 2008). The effects of NIBS applied over the DLPFC in healthy subjects were explored on risk-taking using the *Balloon Analog Risk Task* (BART). In this task, subjects are required to accumulate money by inflating a computerized balloon, whereby they increasingly face the risk of the balloon to explode and lose the accumulated gain. A single session of tDCS with both electrodes over the DLPFCs (i.e., anode placed over either the right

or left DLPFC coupled with the cathode over the contralateral DLPFC) led to a more conservative, risk-averse response style (i.e., decreased number of pumps) as compared to sham stimulation and to unilateral active stimulation (i.e., anodal placed over either the right or left DLPFC coupled with cathodal over the contralateral supraorbital area; Fecteau et al., 2007b). NIBS can also induce the opposite behavioral effect at the BART that is increasing risk-taking in healthy subjects. Participants receiving anodal tACS (6.5 Hz) over the left DLPFC coupled with cathodal over the right temporal cortex displayed greater risk-taking (i.e., increased number of pumps from balloons that did not explode) compared to participants receiving sham stimulation and participants receiving anodal right DLPFC tACS coupled with cathodal over the left temporal cortex (Sela et al., 2012).

The effects of NIBS on risk-taking were also investigated with the *Risk Task*. In the *Risk Task* participants have to choose between two options representing different levels of risk and balances of reward. Subjects receiving low frequency rTMS applied over the right DLPFC displayed riskier decision-making style compared to those receiving rTMS over the left DLPFC or sham rTMS (Knoch et al., 2006). NIBS can also decrease risk-taking using the same task. Subjects receiving tDCS (anodal over the right DLPFC coupled with cathodal over the left DLPFC) displayed suppressed risk-taking and decreased sensitivity to reward, as compared to subjects receiving sham tDCS (Fecteau et al., 2007a). Participants receiving active stimulation were also faster at making their choices compared to participants receiving sham stimulation. These studies converge to the suggestion that NIBS can modulate impulsive behaviors and risk-taking.

EFFECTS OF NIBS ON WORKING MEMORY

Working memory is a widely investigated cognitive function. Working memory allows to transiently maintain information. It encompasses a large brain network, especially the fronto-temporal network including the DLPFC. Working memory capacities can be assessed by the *Sternberg Task*. This task requires participants to recognize a previously presented item (verbal or non-verbal material) amongst distractors. It has been reported that healthy subjects were faster at the *Sternberg Task* when they received anodal tDCS over the left DLPFC coupled with cathodal tDCS over the right DLPFC as compared to when they received sham tDCS (Gladwin et al., 2012). This effect on the *Sternberg Task* was replicated in a study using high frequency rTMS applied over the left and right DLPFC. Participants were faster (but not more accurate) to perform the task after active (left and right DLPFC) rTMS compared to sham rTMS (Preston et al., 2010). It has also been reported that tDCS over the left DLPFC enhanced working memory as measured by the *backward digit span* (Jeon and Han, 2012). In this task, random sequences of numbers (range 0–9) are verbally presented to participants. The subjects have then to repeat the sequence of numbers in the reverse order. In an ecological effort, working memory can also be studied using an adapted version of the *Object-location learning* paradigm. In this task, subjects had to learn the accurate positions of buildings on a street map by looking at a series of correct and incorrect pairings of buildings (objects) and street map positions (locations). It has been reported that

accuracy (i.e., percentage of correct object-location recalls) was improved when subjects received active tDCS (anodal over the right temporoparietal junction coupled with cathodal over the contralateral supraorbital area) as compared to sham tDCS at this task in healthy elderly subjects (mean age of 62 years old; Floel et al., 2012). Interestingly, these effects were found after 1 week (i.e., delayed free recall). It thus appears that NIBS can enhance short term working memory performance in healthy subjects.

EFFECTS OF NIBS ON PLANNING

Planning represents the ability to divide behaviors step by step, in a particular order, to reach a specific goal (Unterrainer and Owen, 2006). It involves a large cerebral network including the DLPFC (Unterrainer and Owen, 2006). One well-known paradigm to measure planning is the *Tower of London Task*. In this task subjects are presented with three rods and a number of disks of different sizes which can slide onto any rod. They are invited to preplan mentally a sequence of moves from an initial state to match a goal state (initial thinking phase) and then to execute the moves one by one (execution phase). Some studies indicate that NIBS can improve the overall planning skills at the *Tower of London Task*. Dockery et al. (2009) investigated the impact of tDCS applied over the left DLPFC on this task in a crossover design. Participants were faster (when they received cathodal tDCS) and more accurate (when they received anodal tDCS) to complete the puzzle (preplan and execute) as compared to sham tDCS. Accuracy was calculated as the number of correct solutions divided by the total number of trials. A more recent study reported that cTBS applied over the left DLPFC can diminish the preplan time (initial thinking period) without changing performance at the *Tower of London Task*, compared to when participants received sham stimulation. Of note, iTBS applied over the same brain area lengthened speed of execution at this task (Kaller et al., 2013). Thus, NIBS applied over the DLPFC seems to enhance planning in healthy subjects.

NIBS can also reduce reaction time to solve a problem in an *Analogic Reasoning Task*. This task requires participants to identify analogies between two sets of pictures of colored geometric shapes presented at the same time. Participants were faster at detecting analogies without affecting error rates when they received rTMS over the left DLPFC as compared to when they received rTMS over the right DLPFC and sham stimulation (Boroojerdi et al., 2001).

EFFECTS OF NIBS ON DECEPTIVE CAPACITIES

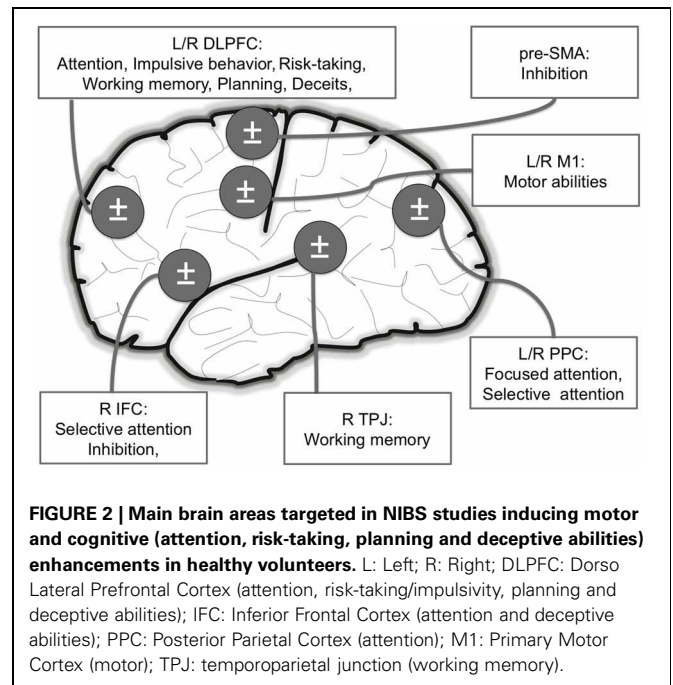
Deceptive capacities are commonly defined as the abilities to intentionally mislead another individual by falsifying truthful information in a credible way (Vrij et al., 2001). One of the most robust measures to identify deceitful from truthful answers is that deceitful answers are associated with longer onset (Walczyk et al., 2003). Another measure of deceit is the level of guilt as assessed with questions regarding the emotional state (e.g., “Did you feel guilty when lying?”; Caso et al., 2005). Lying elicited activity in several regions, including the DLPFC (Nunez et al., 2005) and the anterior prefrontal cortex (aPFC; Abe et al., 2007). First, it seems that production of lies can be improved (as well as impaired) by NIBS (Karton and Bachmann, 2011). This ability

was assessed in a task where subjects have to overtly name the color of a disc (blue or red) presented on a computer screen or lie. The authors investigated the effect of 1 Hz rTMS applied over the right and left DLPFC as compared to the same pattern of stimulation applied over the ipsilateral parietal cortex. The authors reported that participants produced less truthful answers after they received rTMS over the left DLPFC compared to when they received stimulation over the parietal cortex. Karim et al. (2010) evaluated the effects of tDCS on deceptive abilities using the *Guilty Knowledge Test*. In this task, subjects participate in a thief role-play in which they are supposed to steal money and then to attend to an interrogation. During the interrogation they have to respond to multi-choice questions, usually consisting of six possible answers; one of which that would only be known by a guilty person, the other five answers being equally plausible to an innocent person. Subjects who received active tDCS (anodal over the left parietal cortex coupled with cathodal over the right aPFC) were better at deceiving than when they received sham tDCS. More specifically, they were faster at lying and they reported lesser guilt. The opposite electrode montage (i.e., anodal over the left aPFC coupled with cathodal over the right parietal cortex) did not modulate deceptive behaviors (Karim et al., 2010). The effects of tDCS on other deceptive abilities were also investigated (Pecteau et al., 2013). Three kinds of stimulation parameters were compared: the anode over the right DLPFC coupled with the cathode over the left DLPFC, the opposite electrode arrangement (anodal over the left DLPFC coupled with cathodal over the right DLPFC) and sham tDCS. Main findings include that compared to subjects who received sham stimulation, those who received active tDCS (anodal over the right or left DLPFC coupled with cathodal over the contralateral region) were faster at recalling memorized untruthful answers. No change in RT was found in these subjects for providing truthful responses. In sum, although data are still limited, they suggest that NIBS may improve some deceptive abilities.

DISCUSSION

We reviewed here studies indicating that NIBS can improve normal performance in healthy subjects (see Figure 2). Specifically, these improvements were observed for motor abilities (e.g., greater muscular endurance), attentional processes (e.g., faster threat detection), impulsive behavior (e.g., choosing more often larger, delayed rewards than smaller, immediate rewards), risk-taking (e.g., displaying more careful behaviors, diminished or increased risk-taking), memory (e.g., increased working memory load), planning (e.g., enhanced fluid reasoning), and deceptive capacities (e.g., decreased RT in providing deceitful answers).

Interestingly, some of these motor and cognitive processes that can be enhanced using NIBS are already targeted in specific training programs for security purposes. Indeed, some approaches already exist to develop soldiers' motor abilities to emphasize combat readiness. Amongst them, the *Army Physical Fitness Test* is a common program to train physical performance in military. This program trains multimodal aspects of motor performance such as endurance, mobility, strength, and flexibility (Heinrich et al., 2012).



There are also several training programs to enhance cognitive functions for security purposes. Training attention to detect threatening stimuli constitutes one of the highest priorities for security services (see report from the Committee on Opportunities in Neuroscience for Future Army Applications and Council, 2009). Airport security screening staff are trained with *computer-based training* programs to improve their attentional skills in order to enhance their abilities to detect threatening objects in X-ray images (Schwaninger, 2004). As previously discussed, the *DAWARS Ambush!* program was developed to train soldiers to accurately detect threatening objects in realistic environment. Similarly, soldiers are trained to enhance their attentional skills in shooting using the *pop-up target friend or foe* programs (Kelley et al., 2011). In this training program, soldiers have to shoot or refrain from shooting targets representing either friends or foes. Accuracy and RT are trained during specific shooting training. Another training consists of developing automatic behavior to reduce aversive effect of stress on performance for which cognitive control is needed (Leach, 2004). In this way, soldiers are trained to create and follow cognitive automatisations so-called drills (e.g., *if you are under fire, you find cover*; Delahaij et al., 2006). There is also *The Reid training program* (Jayne and Buckley, 1999), which provides interrogation and interviewing techniques seminars. The goal of this training program is to develop adaptive attentional skills, planning abilities, memory abilities, and appropriate risk-taking. In sum, several of these motor and cognitive skills, as mentioned earlier, can be enhanced with NIBS in healthy subjects. Thus, one might speculate that NIBS may be a promising neuroenhancement tool for security purposes. However, transferability and meaningfulness of these NIBS-induced paradoxical facilitations into real life situations are not clear yet.

ARE NIBS-INDUCED PARADOXICAL FACILITATIONS TRANSFERABLE INTO REAL-LIFE SITUATIONS?

Before proposing NIBS as a neuroenhancement tool for security purposes, we have to discuss whether these enhancements may be transferable into real-life situations. Indeed, most of the NIBS-induced facilitation data reviewed here have been collected in laboratory settings. This particular environment using rigorous scientific methods is needed to identify as much as possible the exact changes that are induced by NIBS, not only the improvements, but also potential impairments with controls conditions for instance. This represents an important step toward the development of a new neuroenhancement technique. However, if we want to use NIBS to improve functions relevant in real-life situations, we need to explore whether they can be transferred into real-life.

One avenue to further transferability is to promote the ecological validity of the experimental tasks. Several factors can be promoted to boost the ecological validity of experimental testing. A first factor is how the function is measured. Most functions are measured with computer programs. For instance, target detection can be assessed in laboratory settings using the *Flanker task* (Lavie and Cox, 1997). More recently, target detection has been tested in a more ecological task: the *DAWARS Ambush!* As mentioned earlier, this computer-based program simulates foreign countries environments to train threat detection in war situations (e.g., detect land mines or the safe hidden path used by the enemy to avoid these mines into realistic environment). The effects of NIBS on target detection using the *Flanker* and the *DAWARS Ambush!* paradigms have also been tested. Target detection was improved with active as compared to sham stimulation in healthy subjects at the *Flanker task* (Weiss and Lavidor, 2012) and the *DARWARS Ambush!* (Clark et al., 2012; Falcone et al., 2012). Another example is impulsivity. A common way to test impulsivity level in laboratory settings is with a computer-based task, the SST (O'Brien and Gormley, 2013). Efforts have been made to test impulsivity in more ecological paradigms, such as using a driving simulator (Pearson et al., 2013). The effects of NIBS have been tested on impulsivity on these tasks. Active stimulation as compared to sham stimulation can lead to lower impulsivity level at the SST (Hsu et al., 2011) and at the driving simulator (Beeli et al., 2008). Another example is working memory. A widely used task to characterize working memory and learning is the *Sternberg Task*. In order to assess spatial working memory in a more ecological context, performance of subjects can be assessed using map-learning procedure based on existing maps (Bosco et al., 2004). In such *Street Map Task*, objects are placed on a map and participants have to remember the positions of the objects. The effects of NIBS have been tested on both the *Sternberg Task* and a *Street Map Task*. Results revealed that NIBS improved working memory at the *Sternberg Task* (Gladwin et al., 2012) and at the *Street Map Task* (Floel et al., 2012). These examples are good models to follow to promote the ecological value of laboratory setting without compromising scientific methodological rigor.

In order to promote the effects of NIBS in this population, we need to test the effects of NIBS on ecological tasks and mimic as much as possible external factors that might have an impact, such as performing under stressful situations. Technological advances

such as the development of immersive 3D scenarios will certainly optimize smooth translation from laboratory programs into real-life situations. A last point to discuss concerns the generalization of these NIBS-induced improvement at specific task to the whole functioning (global intelligence) as it can be the case with cognitive training (Jaeggi et al., 2010). Now, let's say that in the best-case scenario, NIBS can be transferred into real-life situations. The next question is: *Are these NIBS-induced paradoxical facilitations meaningful for real-life situations?*

ARE NIBS-INDUCED PARADOXICAL FACILITATIONS MEANINGFUL FOR REAL-LIFE SITUATIONS?

Throughout this paper we presented studies showing paradoxical facilitation induced by NIBS on various motor and cognitive functions. If these NIBS-induced motor and cognitive enhancements are transferable in real-life situations, another question that remains is whether they are meaningful for security purposes. Meaningfulness is defined here as the magnitude and the duration of the effects, in other words *Are they big enough to have a real impact?*

Magnitude of these NIBS-induced facilitations is widely variable. Although statistically significant, whether the magnitude of these enhancements is meaningful for daily-life situations is not clear yet. For instance, Pascual-Leone et al. (2012) estimated a mean reduction of 32 milliseconds from studies using NIBS to improve motor RT. In the specific context of speed shooting performances, ~13 milliseconds would be the difference between elite and rookie police officers (Vickers and Lewinski, 2012). Therefore, an improvement of 32 ms may make a vital difference in the context of a one-on-one gunfight or during aircraft combat (dogfight). This suggested that the magnitude of NIBS-induced enhancements might have a real interest for soldiers and police officers. On the other hand, the magnitude of the enhancement typically observed using NIBS are rather the same as those observed using pharmacological enhancers such as caffeine (Husain and Mehta, 2011). Duration of these NIBS-induced paradoxical facilitations is widely variable across studies, from several minutes to several months (Dockery et al., 2009; Reis et al., 2009). Duration of these effects obviously plays an important role in determining whether these enhancements are meaningful for real-life situations or determining the best timing to stimulate or re-stimulate. Even when tested in laboratory settings in which testing is rigorously controlled, the real duration of these enhancements remains uncertain.

Several factors can influence the magnitude and duration of these paradoxical facilitations, thus ultimately transferability of laboratory findings into real-life situations. These factors can be related to (1) the NIBS device, (2) the brain state, and (3) the behaviors.

- (1) Factors related to the NIBS device that can influence facilitation include the stimulation parameters. These parameters such as frequency, intensity, number of pulses, and number of sessions can influence the magnitude and duration of paradoxical facilitations. For instance, Iyer et al. (2005) found greater effects with 2 mA than 1 mA on verbal fluency in healthy subjects.

- (2) Brain state can also influence the effects of NIBS on paradoxical facilitation. State dependency can be defined as the baseline state of brain dealing with many factors such as fatigue, sleep, experience, and personality traits (e.g., Silvanto and Pascual-Leone, 2008; Silvanto et al., 2008). The effects of NIBS can be state dependent. One example comes from tDCS. Slow oscillatory-tDCS (i.e., 0.75 Hz) applied bilaterally over the DLPFC during slow wave sleep increased retention of declarative memory capacities (word pairs previously learned), as compared to sham stimulation. This improvement of memory capacities was associated with an increased sleep depth and slow oscillatory activity (<3 Hz), whereas the power in the faster frequency EEG bands (theta, alpha, and beta) was reduced. In contrast, with the same protocol of slow oscillatory-tDCS, but applied during the wake retention interval, there were no effects on declarative memory (Marshall et al., 2004). Thus, administering NIBS during a specific sleep phase facilitated sleep-dependent consolidation of declarative memories. This kind of result highlights the importance of state-dependency in NIBS-induced paradoxical facilitation. In other words, in order to optimize NIBS efficacy, we have to determine the best state the subject needs to be before, during and after stimulation. For instance, Kwon et al. (2013) reported that NIBS improved inhibition when it was applied while subjects performed the SST, but it had no effect when it was applied before the SST.
- (3) Behavioral level at baseline can also influence the effects of NIBS on paradoxical facilitation. Even when performance of a group of subjects is considered normal, within the normal range, some subjects displayed better performance than others (e.g., normal distribution). This baseline level of performance may influence the effects of NIBS. For example, NIBS improved visual working memory skills in low performing subjects, but not in higher performing ones (Tseng et al., 2012).

Age and gender can influence behavioral performance as well as the effects of NIBS. Indeed, baseline performance can vary according to subject's age and gender. Throughout life, our skills naturally change. For example older individuals present slower RT to motion onset than younger ones (Porciatti et al., 1999). Attentional capacities also change with aging (McDowd and Craik, 1988). The same observation has been reported on planning abilities with older adults displaying worse performance at the *Tower of London task* than younger adults (Phillips et al., 2006). Normal aging also affects working memory. For example, it has been reported that older participants displayed both reduced accuracy and slower RT at working memory tasks compared to younger participants (Gazzaley et al., 2005). In sum, it is well-accepted that motor and cognitive performance change through aging (see review from Glisky, 2007). The influence of age on NIBS-induced paradoxical facilitation has not been however extensively investigated yet (for a review see Freitas et al., 2013). One study reported that rTMS induced greater facilitation of inhibition at the *Go/NoGo task* in younger than older adults (age range 28–37 years; Huang et al., 2004), whereas another study reported that NIBS led to greater improvement of motor

skills in older than younger participants (age range 56–87 years; Hummel et al., 2010). On one hand, it is possible that NIBS induces larger facilitation in younger than older adults. Indeed, age was reported to correlate negatively with the duration of NIBS-induced neurophysiological effects: longer-lasting effects were found in younger than older healthy subjects. It is speculated that this change in cortical plasticity through aging is linked to normal motor and cognitive decline (Freitas et al., 2013). On the other hand, it is possible that normal performance in older individuals might be easier to improve with NIBS than in younger ones. We could call this motor or cognitive *rejuvenation* that is making older individuals performing as when they were younger.

Gender may also be a considerable factor when using NIBS to induce facilitation in healthy subjects. At the behavioral level, baseline performance can differ according to gender. For example, men are more accurate at a throwing task than women (Moreno-Briseno et al., 2010). Cognitive performance has also been reported different according to gender in numerous functions (for a review, see Zaidi, 2010), such as attentional inhibition (Halari et al., 2005), visual-spatial attention (Rubia et al., 2010), and spatial working memory (Duff and Hampson, 2001). The influence of gender on NIBS-induced effects has not been rigorously studied and remains to be further characterized (Ridding and Ziemann, 2010). Most NIBS studies are not specifically designed to test for gender differences. In sum, further studies are needed to characterize the real influence of several factors, including those related to the device, brain state, behavioral level at baseline, age, and gender on NIBS-induced paradoxical facilitation. Better knowledge of these factors will certainly help to smooth transferability and increase meaningfulness of laboratory setting protocols into real-life contexts.

Another way to improve transferability and meaningfulness of the NIBS induced effects might be to use NIBS as an *add-on* to existing training programs. NIBS may promote capacities that are critical for security purposes. Some studies reported that the combination of motor training and NIBS lead to greater motor improvements than to a single method approach (e.g., physical exercise alone; Bolognini et al., 2009; Williams et al., 2010). This has also been reported in cognition. Combining cognitive training with NIBS resulted in greater effects than single method approach (e.g., stimulation alone). For instance, the combination of a *n-back* training and active tDCS resulted in greater performance at the *digit span task* than tDCS used as a single method approach and the combination of the *n-back* training and sham tDCS (Andrews et al., 2011). Thus, existing programs developed for security personnel might benefit from combining them with NIBS.

ETHICAL CONCERNS OF USING NIBS-INDUCED PARADOXICAL FACILITATION IN HEALTHY SUBJECTS

Although this is out of the scope of this review paper, it is important to mention that this field—inducing paradoxical facilitations with NIBS in healthy subjects—calls for fair and well-balanced discussions on ethics. This discussion should be to some extent in accordance with lines of conduct from the use of other neuroenhancers, such as smart pills (for review Illes and Bird, 2006; Forlini et al., 2013). At this point, whether or not it is ethical to use

NIBS as a neuroenhancement tool for security purposes remains an open debate. If it is, another question remains: *Is it safe?*

SAFETY CONCERNS OF USING NIBS-INDUCED PARADOXICAL FACILITATION IN HEALTHY SUBJECTS

There are known risks and hypothetical risks associated with the use of NIBS. These risks are reviewed by different groups on the use of NIBS (Wassermann, 1998; Iyer et al., 2005; Rossi et al., 2009). The classic protocol that is considered safe to reduce depressive symptoms in patients with major depression refractory to medications consists of delivering daily session (a session a day, from Monday through Friday) of high frequency rTMS during 3–6 weeks (O'Reardon et al., 2007). Repeated sessions are delivered in order to induce longer lasting clinical benefits. Common side-effects related to this protocol include headaches or cutaneous discomfort.

In healthy subjects, the use of tDCS has been reported to be safe with a single session in 103 subjects (Iyer et al., 2005). However, there are no safety guidelines for the administration of repeated NIBS sessions over a long period of time in healthy individuals. We cannot solely and directly derive them from safety guidelines established for clinical populations. One reason is that the effects of a given NIBS protocol known to be safe (and even salutary) in a clinical population may not be safe in healthy volunteers. For instance, delivering high frequency rTMS over the left DLPFC can alleviate depressive symptoms in patients with depression (i.e., clinical benefit), but can hinder mood in healthy subjects (i.e., would be considered as a side-effect). Hence, we must consider the possibility that a same NIBS protocol might lead to opposite behavioral effects depending on the studied populations.

Regarding the NIBS-induced enhancement studies, another important related aspect that must be taken into consideration is the possibility of incidentally eliciting other effects. For instance, in line with the zero-sum theory principle, rTMS resulted in

improved detection of targets in the ipsi- or contra-lateral visual-field and in impaired detection in the opposite visual field (Thut et al., 2005; Buettefisch et al., 2011). NIBS-induced facilitation of motor function can also shift the speed/accuracy trade-off function (Reis et al., 2009; Nelson et al., 2013). This dual effect is not new, nor restricted to the use of NIBS. This speed/accuracy trade-off is commonly observed in cognitive programs (Van Veen et al., 2008). Novice inspectors of aircrafts are trained to detect defects with immersive virtual scenarios. This training leads to increased attentional accuracy, but also to increased RT to detect threatening defects (Sadasivan et al., 2005). We might not be able to prevent or minimize some of these trade-offs yet, but the benefit/risk ratio should be carefully addressed. With regards to hypothetical risks, it is also important to keep in mind some results from the animal literature. It is well-known that animal can develop an addiction to auto-electrical stimulation. This represent an hypothetical risk for humans to develop an addiction to neuroenhancers (Heinz et al., 2012).

CONCLUSION

In this article we reviewed experimental data supporting that NIBS can enhance motor (precision, speed, strength, acceleration endurance, and execution of daily motor task) and cognitive functions (attention, impulsivity, risk-taking, working memory, planning, and deceptive capacities) in healthy individuals. Some of these functions are already trained with existing programs for security services. It is thus tempting to speculate that NIBS may serve as a neuroenhancer tool for security purposes. However, numerous questions remain to be answered to do so. We believe that two important questions are (1) *Are these paradoxical facilitations induced in laboratory settings transferable into real-life situations?* and (2) *If they are transferable, are they meaningful for real-life events?* Furthermore, ethical and safety concerns should be carefully addressed.

REFERENCES

- Abe, N., Suzuki, M., Mori, E., Itoh, M., and Fujii, T. (2007). Deceiving others: distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *J. Cogn. Neurosci.* 19, 287–295. doi: 10.1162/jocn.2007.19.2.287
- Andrews, S. C., Hoy, K. E., Enticott, P. G., Daskalakis, Z. J., and Fitzgerald, P. B. (2011). Improving working memory: the effect of combining cognitive activity and anodal transcranial direct current stimulation to the left dorsolateral prefrontal cortex. *Brain Stimul.* 4, 84–89. doi: 10.1016/j.brs.2010.06.004
- Avanzino, L., Bove, M., Trompetto, C., Tacchino, A., Ogliastro, C., and Abbruzzese, G. (2008). 1-Hz repetitive TMS over ipsilateral motor cortex influences the performance of sequential finger movements of different complexity. *Eur. J. Neurosci.* 27, 1285–1291. doi: 10.1111/j.1460-9568.2008.06086.x
- Banissy, M. J., and Muggleton, N. G. (2013). Transcranial direct current stimulation in sports training: potential approaches. *Front. Hum. Neurosci.* 7:129. doi: 10.3389/fnhum.2013.00129
- Barker, A. T., Jalinous, R., and Freeston, I. L. (1985). Non-invasive magnetic stimulation of human motor cortex. *Lancet* 1, 1106–1107. doi: 10.1016/S0140-6736(85)92413-4
- Beeli, G., Koenke, S., Gasser, K., and Jancke, L. (2008). Brain stimulation modulates driving behavior. *Behav. Brain Funct.* 4:34. doi: 10.1186/1744-9081-4-34
- Boggio, P. S., Castro, L. O., Savagim, E. A., Brait, R., Cruz, V. C., Rocha, R. R., et al. (2006). Enhancement of non-dominant hand motor function by anodal transcranial direct current stimulation. *Neurosci. Lett.* 404, 232–236. doi: 10.1016/j.neulet.2006.05.051
- Bolognini, N., Pascual-Leone, A., and Fregni, F. (2009). Using non-invasive brain stimulation to augment motor training-induced plasticity. *J. Neuroeng. Rehabil.* 17, 6–8. doi: 10.1186/1743-0003-6-8
- Bolognini, N., Rossetti, A., Casati, C., Mancini, F., and Vallar, G. (2011). Neuromodulation of multisensory perception: a tDCS study of the sound-induced flash illusion. *Neuropsychologia* 49, 231–237. doi: 10.1016/j.neuropsychologia.2010.11.015
- Boroojerdi, B., Phipps, M., Kopylev, L., Wharton, C. M., Cohen, L. G., and Grafman, J. (2001). Enhancing analogic reasoning with rTMS over the left prefrontal cortex. *Neurology* 56, 526–528. doi: 10.1212/WNL.56.4.526
- Bosco, A., Longoni, A. M., and Vecchi, T. (2004). Gender effects in spatial orientation: Cognitive profiles and mental strategies. *Appl. Cogn. Psychol.* 18, 519–532. doi: 10.1002/acp.1000
- Buettefisch, C. M., Hines, B., Shuster, L., Pergami, P., and Mathes, A. (2011). Motor demand-dependent improvement in accuracy following low-frequency transcranial magnetic stimulation of left motor cortex. *J. Neurophysiol.* 106, 1614–1621. doi: 10.1152/jn.00048.2011
- Caso, L., Gnisci, A., Vrij, A., and Mann, S. (2005). Processes underlying deception: an empirical analysis of truth and lies when manipulating the stakes. *J. Invest. Psychol. Offender Profiling* 2, 195–202. doi: 10.1002/jip.32
- Cho, S. S., Pellicchia, G., Ko, J. H., Ray, N., Obeso, I., Houle, S., et al.

- (2012). Effect of continuous theta burst stimulation of the right dorsolateral prefrontal cortex on cerebral blood flow changes during decision making. *Brain Stimul.* 5, 116–123. doi: 10.1016/j.brs.2012.03.007
- Clark, V. P., Coffman, B. A., Mayer, A. R., Weisend, M. P., Lane, T. D., Calhoun, V. D., et al. (2012). TDCS guided using fMRI significantly accelerates learning to identify concealed objects. *Neuroimage* 59, 117–128. doi: 10.1016/j.neuroimage.2010.11.036
- Clark, V. P., Coffman, B. A., Trumbo, M. C., and Gasparovic, C. (2011). Transcranial direct current stimulation (tDCS) produces localized and specific alterations in neurochemistry: a ¹H magnetic resonance spectroscopy study. *Neurosci. Lett.* 500, 67–71. doi: 10.1016/j.neulet.2011.05.244
- Coffman, B. A., Trumbo, M. C., Flores, R. A., Garcia, C. M., Van Der Merwe, A. J., Wassermann, E. M., et al. (2012). Impact of tDCS on performance and learning of target detection: interaction with stimulus characteristics and experimental design. *Neuropsychologia* 50, 1594–1602. doi: 10.1016/j.neuropsychologia.2012.03.012
- Cogiamanian, E., Marceglia, S., Ardolino, G., Barbieri, S., and Priori, A. (2007). Improved isometric force endurance after transcranial direct current stimulation over the human motor cortical areas. *Eur. J. Neurosci.* 26, 242–249. doi: 10.1111/j.1460-9568.2007.05633.x
- Committee on Opportunities in Neuroscience for Future Army Applications, N. R. (2009). *Opportunities in Neuroscience for Future Army Applications*. Washington, DC: The National Academies Press.
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi: 10.1038/nrn755
- Delahaj, R., Gaillard, A. W. K., and Soeters, J. M. L. M. (2006). “Stress training and the new military environment,” in *Human Dimensions in Military Operations – Military Leaders’ Strategies for Addressing Stress and Psychological Support, Meeting Proceedings RTO-MP-HFM-134, Paper 17A*. (Neuilly-sur-Seine: RTO), 17A-1–17A-10. Available online at: <http://www.rto.nato.int/abstracts.asp>
- Ditely, T., Jacobson, L., Walsh, V., and Lavidor, M. (2012). Modulating behavioral inhibition by tDCS combined with cognitive training. *Exp. Brain Res.* 219, 363–368. doi: 10.1007/s00221-012-3098-4
- Dockery, C. A., Hueckel-Weng, R., Birbaumer, N., and Plewnia, C. (2009). Enhancement of planning ability by transcranial direct current stimulation. *J. Neurosci.* 29, 7271–7277. doi: 10.1523/JNEUROSCI.0065-09.2009
- Duff, S. J., and Hampson, E. (2001). A sex difference on a novel spatial working memory task in humans. *Brain Cogn.* 47, 470–493. doi: 10.1006/brcg.2001.1326
- Falcone, B., Coffman, B. A., Clark, V. P., and Parasuraman, R. (2012). Transcranial direct current stimulation augments perceptual sensitivity and 24-hour retention in a complex threat detection task. *PLoS ONE* 7:e34993. doi:10.1371/journal.pone.0034993
- Fecteau, S., Boggio, P., Fregni, F., and Pascual-Leone, A. (2013). Modulation of untruthful responses with non-invasive brain stimulation. *Front. Psychiatry* 3:97. doi: 10.3389/fpsy.2012.00097
- Fecteau, S., Knoch, D., Fregni, F., Sultani, N., Boggio, P., and Pascual-Leone, A. (2007a). Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. *J. Neurosci.* 27, 12500–12505. doi: 10.1523/jneurosci.3283-07.2007
- Fecteau, S., Pascual-Leone, A., Zald, D. H., Liguori, P., Theoret, H., Boggio, P. S., et al. (2007b). Activation of prefrontal cortex by transcranial direct current stimulation reduces appetite for risk during ambiguous decision making. *J. Neurosci.* 27, 6212–6218. doi: 10.1523/jneurosci.0314-07.2007
- Fecteau, S., Pascual-Leone, A., and Theoret, H. (2006). Paradoxical facilitation of attention in healthy humans. *Behav. Neurol.* 17, 159–162.
- Floel, A., Sutorp, W., Kohl, O., Kurten, J., Lohmann, H., Breitenstein, C., et al. (2012). Non-invasive brain stimulation improves object-location learning in the elderly. *Neurobiol. Aging* 33, 1682–1689. doi: 10.1016/j.neurobiolaging.2011.05.007
- Forlini, C., Hall, W., Maxwell, B., Outram, S. M., Reiner, P. B., Repantis, D., et al. (2013). Navigating the enhancement landscape. Ethical issues in research on cognitive enhancers for healthy individuals. *EMBO Rep.* 14, 123–128. doi: 10.1038/embor.2012.225
- Freitas, C., Farzan, F., and Pascual-Leone, A. (2013). Assessing brain plasticity across the lifespan with transcranial magnetic stimulation: why, how, and what is the ultimate goal? *Front. Neurosci.* 7:42. doi: 10.3389/fnins.2013.00042
- Gazzaley, A., Cooney, J. W., Rissman, J., and D’Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nat. Neurosci.* 8, 1298–1300. doi: 10.1038/nn1543
- Gladwin, T. E., Den Uyl, T. E., and Wiers, R. W. (2012). Anodal tDCS of dorsolateral prefrontal cortex during an Implicit Association Test. *Neurosci. Lett.* 517, 82–86. doi: 10.1016/j.neulet.2012.04.025
- Glisky, E. L. (2007). “Changes in cognitive function in human aging,” in *Brain Aging: Models, Methods, and Mechanisms*, ed D. R. Riddle (Boca Raton, FL: CRC Press), 3–20.
- Halari, R., Hines, M., Kumari, V., Mehrotra, R., Wheeler, M., Ng, V., et al. (2005). Sex differences and individual differences in cognitive performance and their relationship to endogenous gonadal hormones and gonadotropins. *Behav. Neurosci.* 119, 104–117. doi: 10.1037/0735-7044.119.1.104
- Heinrich, K. M., Spencer, V., Fehl, N., and Poston, W. S. (2012). Mission essential fitness: comparison of functional circuit training to traditional Army physical training for active duty military. *Mil. Med.* 177, 1125–1130.
- Heinz, A., Kipke, R., Heimann, H., and Wiesing, U. (2012). Cognitive neuroenhancement: false assumptions in the ethical debate. *J. Med. Ethics* 38, 372–375. doi: 10.1136/medethics-2011-100041
- Hilgetag, C. C., Theoret, H., and Pascual-Leone, A. (2001). Enhanced visual spatial attention ipsilateral to rTMS-induced “virtual lesions” of human parietal cortex. *Nat. Neurosci.* 4, 953–957. doi: 10.1038/nn0901-953
- Hoogendam, J. M., Ramakers, G. M., and Di Lazzaro, V. (2010). Physiology of repetitive transcranial magnetic stimulation of the human brain. *Brain Stimul.* 3, 95–118. doi: 10.1016/j.brs.2009.10.005
- Hsu, T. Y., Tseng, L. Y., Yu, J. X., Kuo, W. J., Hung, D. L., Tzeng, O. J., et al. (2011). Modulating inhibitory control with direct current stimulation of the superior medial frontal cortex. *Neuroimage* 56, 2249–2257. doi: 10.1016/j.neuroimage.2011.03.059
- Huang, C. C., Su, T. P., Shan, I. K., and Wei, I. H. (2004). Effect of 5 Hz repetitive transcranial magnetic stimulation on cognition during a Go/NoGo task. *J. Psychiatr. Res.* 38, 513–520. doi: 10.1016/j.jpsychires.2004.01.006
- Hummel, F. C., Heise, K., Celnik, P., Floel, A., Gerloff, C., and Cohen, L. G. (2010). Facilitating skilled right hand motor function in older subjects by anodal polarization over the left primary motor cortex. *Neurobiol. Aging* 31, 2160–2168. doi: 10.1016/j.neurobiolaging.2008.12.008
- Husain, M., and Mehta, M. A. (2011). Cognitive enhancement by drugs in health and disease. *Trends Cogn. Sci.* 15, 28–36. doi: 10.1016/j.tics.2010.11.002
- Hwang, J. H., Kim, S. H., Park, C. S., Bang, S. A., and Kim, S. E. (2010). Acute high-frequency rTMS of the left dorsolateral prefrontal cortex and attentional control in healthy young men. *Brain Res.* 1329, 152–158. doi: 10.1016/j.brainres.2010.03.013
- Illes, J., and Bird, S. J. (2006). Neuroethics: a modern context for ethics in neuroscience. *Trends Neurosci.* 29, 511–517. doi: 10.1016/j.tins.2006.07.002
- Iyer, M. B., Mattu, U., Grafman, J., Lomarev, M., Sato, S., and Wassermann, E. M. (2005). Safety and cognitive effect of frontal DC brain polarization in healthy individuals. *Neurology* 64, 872–875. doi: 10.1212/01.WNL.0000152986.07469.E9
- Jacobson, L., Ezra, A., Berger, U., and Lavidor, M. (2012). Modulating oscillatory brain activity correlates of behavioral inhibition using transcranial direct current stimulation. *Clin. Neurophysiol.* 123, 979–984. doi: 10.1016/j.clinph.2011.09.016
- Jacobson, L., Javitt, D. C., and Lavidor, M. (2011). Activation of inhibition: diminishing impulsive behavior by direct current stimulation over the inferior frontal gyrus. *J. Cogn. Neurosci.* 23, 3380–3387. doi: 10.1162/jocn_a_00020
- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y. F., Jonides, J., and Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning – implications for training and transfer. *Intelligence* 38, 625–635. doi: 10.1016/j.intell.2010.09.001
- Jayne, B. C., and Buckley, J. T. (1999). *The Investigator Anthology: A Compilation of Articles and Essays about the Reid Technique of Interviewing and Interrogation*. Chicago, IL: John E. Reid and Associates.

- Jeon, S. Y., and Han, S. J. (2012). Improvement of the working memory and naming by transcranial direct current stimulation. *Ann. Rehabil. Med.* 36, 585–595. doi: 10.5535/arm.2012.36.5.585
- Kaller, C. P., Heinze, K., Frenkel, A., Lappchen, C. H., Unterrainer, J. M., Weiller, C., et al. (2013). Differential impact of continuous theta-burst stimulation over left and right DLPFC on planning. *Hum. Brain Mapp.* 34, 36–51. doi: 10.1002/hbm.21423
- Kapur, N. (1996). Paradoxical functional facilitation in brain-behaviour research. A critical review. *Brain* 119(Pt 5), 1775–1790. doi: 10.1093/brain/119.5.1775
- Karim, A. A., Schneider, M., Lotze, M., Veit, R., Sauseng, P., Braun, C., et al. (2010). The truth about lying: inhibition of the anterior prefrontal cortex improves deceptive behavior. *Cereb. Cortex* 20, 205–213. doi: 10.1093/cercor/bhp090
- Karton, I., and Bachmann, T. (2011). Effect of prefrontal transcranial magnetic stimulation on spontaneous truth-telling. *Behav. Brain Res.* 225, 209–214. doi: 10.1016/j.bbr.2011.07.028
- Keeser, D., Meindl, T., Bor, J., Palm, U., Pogarell, O., Mulert, C., et al. (2011). Prefrontal transcranial direct current stimulation changes connectivity of resting-state networks during fMRI. *J. Neurosci.* 31, 15284–15293. doi: 10.1523/JNEUROSCI.0542-11.2011
- Kelley, A. M., Athy, J. R., King, M., Erickson, B., Charamonte, J., Vasbinder, M., et al. (2011). *Think Before You Shoot: The Relationship between Cognition and Marksmanship*. Fort Rucker, Alabama: USAARL. Report No. 2011-23.
- Kim, Y. H., Park, J. W., Ko, M. H., Jang, S. H., and Lee, P. K. W. (2004). Facilitative effect of high frequency subthreshold repetitive transcranial magnetic stimulation on complex sequential motor learning in humans. *Neurosci. Lett.* 367, 181–185. doi: 10.1016/j.neulet.2004.05.113
- Knoch, D., Gianotti, L. R., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., et al. (2006). Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *J. Neurosci.* 26, 6469–6472. doi: 10.1523/JNEUROSCI.0804-06.2006
- Kobayashi, M., Hutchinson, S., Theoret, H., Schlaug, G., and Pascual-Leone, A. (2004). Repetitive TMS of the motor cortex improves ipsilateral sequential simple finger movements. *Neurology* 62, 91–98. doi: 10.1212/WNL.62.1.91
- Kwon, J. W., Nam, S. H., Lee, N. K., Son, S. M., Choi, Y. W., and Kim, C. S. (2013). The effect of transcranial direct current stimulation on the motor suppression in stop-signal task. *Neurorehabilitation* 32, 191–196. doi: 10.3233/NRE-130836
- Ladavas, E., Petronio, A., and Umiltà, C. (1990). The deployment of visual attention in the intact field of hemineglect patients. *Cortex* 26, 307–317. doi: 10.1016/S0010-9452(13)80083-4
- Lavie, N., and Cox, S. (1997). On the efficiency of visual selective attention: Efficient visual search leads to inefficient distractor rejection. *Psychol. Sci.* 8, 395–398. doi: 10.1111/j.1467-9280.1997.tb00432.x
- Leach, J. (2004). Why people “freeze” in an emergency: temporal and cognitive constraints on survival responses. *Aviat. Space Environ. Med.* 75, 539–542.
- Marshall, L., Molle, M., Hallschmid, M., and Born, J. (2004). Transcranial direct current stimulation during sleep improves declarative memory. *J. Neurosci.* 24, 9985–9992. doi: 10.1523/JNEUROSCI.2725-04.2004
- Matsuo, A., Maeoka, H., Hiyamizu, M., Shomoto, K., Morioka, S., and Seki, K. (2011). Enhancement of precise hand movement by transcranial direct current stimulation. *Neuroreport* 22, 78–82. doi: 10.1097/WNR.0b013e32834298b3
- McDowd, J. M., and Craik, F. I. M. (1988). Effects of aging and task-difficulty on divided attention performance. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 267–280. doi: 10.1037/0096-1523.14.2.267
- Moreno-Briseno, P., Diaz, R., Campos-Romo, A., and Fernandez-Ruiz, J. (2010). Sex-related differences in motor learning and performance. *Behav. Brain Funct.* 6:74. doi: 10.1186/1744-9081-6-74
- Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., and Parasuraman, R. (2013). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *Neuroimage*. doi: 10.1016/j.neuroimage.2012.11.061. [Epub ahead of print].
- Nitsche, M. A., Doemkes, S., Karakose, T., Antal, A., Liebetanz, D., Lang, N., et al. (2007). Shaping the effects of transcranial direct current stimulation of the human motor cortex. *J. Neurophysiol.* 97, 3109–3117. doi: 10.1152/jn.01312.2006
- Nitsche, M. A., Schauenburg, A., Lang, N., Liebetanz, D., Exner, C., Paulus, W., et al. (2003). Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human. *J. Cogn. Neurosci.* 15, 619–626. doi: 10.1162/089892903321662994
- Nunez, J. M., Casey, B. J., Egner, T., Hare, T., and Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *Neuroimage* 25, 267–277. doi: 10.1016/j.neuroimage.2004.10.041
- O'Brien, E., and Gormley, M. (2013). The contribution of inhibitory deficits to dangerous driving among young people. *Accid. Anal. Prev.* 51, 238–242. doi: 10.1016/j.aap.2012.11.024
- O'Reardon, J. P., Solvason, H. B., Janicak, P. G., Sampson, S., Isenberg, K. E., Nahas, Z., et al. (2007). Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol. Psychiatry* 62, 1208–1216. doi: 10.1016/j.biopsych.2007.07.007
- Pascual-Leone, A., Horvath, J. C., and Robertson, E. M. (2012). “Enhancement of normal cognitive abilities through noninvasive brain stimulation,” in *Cortical Connectivity: Brain Stimulation for Assessing and Modulating Cortical Connectivity and Function*, eds R. Chen, and J. C. Rothwell (Berlin; Heidelberg: Springer), 207–249.
- Pearson, M. R., Murphy, E. M., and Doanec, A. N. (2013). Impulsivity-like traits and risky driving behaviors among college students. *Accid. Anal. Prev.* 53, 142–148. doi: 10.1016/j.aap.2013.01.009
- Penhune, V. B., and Steele, C. J. (2012). Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behav. Brain Res.* 226, 579–591. doi: 10.1016/j.bbr.2011.09.044
- Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* 35, 73–89. doi: 10.1146/annurev-neuro-062111-150525
- Phillips, L. H., Kliegel, M., and Martin, M. (2006). Age and planning tasks: the influence of ecological validity. *Int. J. Aging Hum. Dev.* 62, 175–184. doi: 10.2190/EM1W-HAYC-TMLM-WW8X
- Porciatti, V., Fiorentini, A., Morrone, M. C., and Burr, D. C. (1999). The effects of ageing on reaction times to motion onset. *Vision Res.* 39, 2157–2164. doi: 10.1016/S0042-6989(98)00288-0
- Preston, G., Anderson, E., Silva, C., Goldberg, T., and Wassermann, E. M. (2010). Effects of 10 Hz rTMS on the neural efficiency of working memory. *J. Cogn. Neurosci.* 22, 447–456. doi: 10.1162/jocn.2009.21209
- Rao, H., Korczykowski, M., Pluta, J., Hoang, A., and Detre, J. A. (2008). Neural correlates of voluntary and involuntary risk taking in the human brain: an fMRI Study of the Balloon Analog Risk Task (BART). *Neuroimage* 42, 902–910. doi: 10.1016/j.neuroimage.2008.05.046
- Reis, J., Schambra, H. M., Cohen, L. G., Buch, E. R., Fritsch, B., Zarahn, E., et al. (2009). Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1590–1595. doi: 10.1073/pnas.0805413106
- Ridding, M. C., and Ziemann, U. (2010). Determinants of the induction of cortical plasticity by non-invasive brain stimulation in healthy subjects. *J. Physiol.* 588, 2291–2304. doi: 10.1113/jphysiol.2010.190314
- Rorie, A. E., and Newsome, W. T. (2005). A general mechanism for decision-making in the human brain. *Trends Cogn. Sci.* 9, 41–43. doi: 10.1016/j.tics.2004.12.007
- Rossi, S., Hallett, M., Rossini, P. M., and Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin. Neurophysiol.* 120, 2008–2039. doi: 10.1016/j.clinph.2009.08.016
- Rubia, K., Hyde, Z., Halari, R., Giampietro, V., and Smith, A. (2010). Effects of age and sex on developmental neural networks of visual-spatial attention allocation. *Neuroimage* 51, 817–827. doi: 10.1016/j.neuroimage.2010.02.058
- Rushworth, M. F., Hadland, K. A., Paus, T., and Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *J. Neurophysiol.* 87, 2577–2592.
- Sadasivan, S., Greenstein, J. S., Gramopadhye, A. K., and Duchowski, A. T. (2005). “Use of eye movement as feedforward training for a synthetic aircraft inspection task,” in: *Proceedings of CHI'05*, (New York, NY), 141–149.
- Sandrini, M., Umiltà, C., and Rusconi, E. (2011). The use of transcranial magnetic stimulation in cognitive

- neuroscience: a new synthesis of methodological issues. *Neurosci. Biobehav. Rev.* 35, 516–536. doi: 10.1016/j.neubiorev.2010.06.005
- Schieber, M. H. (2001). Constraints on somatotopic organization in the primary motor cortex. *J. Neurophysiol.* 86, 2125–2143.
- Schwabinger, A. (2004). Computer based training: a powerful tool to the enhancement of human factors. *Aviat. Secur. Int. FEB*, 31–36.
- Schwarzkopf, D. S., Silvano, J., and Rees, G. (2011). Stochastic resonance effects reveal the neural mechanisms of transcranial magnetic stimulation. *J. Neurosci.* 31, 3143–3147. doi: 10.1523/JNEUROSCI.4863-10.2011
- Sela, T., Kilim, A., and Lavidor, M. (2012). Transcranial alternating current stimulation increases risk-taking behavior in the balloon analog risk task. *Front. Neurosci.* 6:22. doi: 10.3389/fnins.2012.00022
- Sharp, D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., et al. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6106–6111. doi: 10.1073/pnas.1000175107
- Silvano, J., Muggleton, N. G., and Walsh, V. (2008). State dependency in brain stimulation studies of perception and cognition. *Trends Cogn. Sci.* 12, 447–454. doi: 10.1016/j.tics.2008.09.004
- Silvano, J., and Pascual-Leone, A. (2008). State-dependency of transcranial magnetic stimulation. *Brain Topogr.* 21, 1–10. doi: 10.1007/s10548-008-0067-0
- Stagg, C. J., and Nitsche, M. A. (2011). Physiological basis of transcranial direct current stimulation. *Neuroscientist* 17, 37–53. doi: 10.1177/1073858410386614
- Tanaka, S., Hanakawa, T., Honda, M., and Watanabe, K. (2009). Enhancement of pinch force in the lower leg by anodal transcranial direct current stimulation. *Exp. Brain Res.* 196, 459–465. doi: 10.1007/s00221-009-1863-9
- Teo, J. T., Swayne, O. B., Cheeran, B., Greenwood, R. J., and Rothwell, J. C. (2011). Human theta burst stimulation enhances subsequent motor learning and increases performance variability. *Cereb. Cortex* 21, 1627–1638. doi: 10.1093/cercor/bhq231
- Thut, G., Nietzel, A., and Pascual-Leone, A. (2005). Dorsal posterior parietal rTMS affects voluntary orienting of visuospatial attention. *Cereb. Cortex* 15, 628–638. doi: 10.1093/cercor/bhh164
- Tseng, P., Hsu, T. Y., Chang, C. F., Tzeng, O. J., Hung, D. L., Muggleton, N. G., et al. (2012). Unleashing potential: transcranial direct current stimulation over the right posterior parietal cortex improves change detection in low-performing individuals. *J. Neurosci.* 32, 10554–10561. doi: 10.1523/JNEUROSCI.0362-12.2012
- Unterrainer, J. M., and Owen, M. J. (2006). Planning and problem solving: from neuropsychology to functional neuroimaging. *J. Physiol.* 99, 308–317. doi: 10.1016/j.jphysparis.2006.03.014
- Utz, K. S., Dimova, V., Oppenlander, K., and Kerkhoff, G. (2010). Electrified minds: Transcranial direct current stimulation (tDCS) and Galvanic Vestibular Stimulation (GVS) as methods of non-invasive brain stimulation in neuropsychology. A review of current data and future implications. *Neuropsychologia* 48, 2789–2810. doi: 10.1016/j.neuropsychologia.2010.06.002
- Vanderhasselt, M. A., De Raedt, R., Baeken, C., Leyman, L., Clerinx, P., and D'haenen, H. (2007). The influence of rTMS over the right dorsolateral prefrontal cortex on top-down attentional processes. *Brain Res.* 1137, 111–116. doi: 10.1016/j.brainres.2006.12.050
- Vanderhasselt, M. A., De Raedt, R., Baeken, C., Leyman, L., and D'haenen, H. (2006). The influence of rTMS over the right dorsolateral prefrontal cortex on intentional set switching. *Exp. Brain Res.* 172, 561–565. doi: 10.1007/s00221-006-0540-5
- Van Veen, V., Krug, M. K., and Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *J. Cogn. Neurosci.* 20, 1952–1965. doi: 10.1162/jocn.2008.20146
- Vickers, J. N., and Lewinski, W. (2012). Performing under pressure: gaze control, decision making and shooting performance of elite and rookie police officers. *Hum. Mov. Sci.* 31, 101–117. doi: 10.1016/j.humov.2011.04.004
- Vines, B. W., Nair, D., and Schlaug, G. (2008). Modulating activity in the motor cortex affects performance for the two hands differently depending upon which hemisphere is stimulated. *Eur. J. Neurosci.* 28, 1667–1673. doi: 10.1111/j.1460-9568.2008.06459.x
- Vrij, A., Edward, K., and Bull, R. (2001). People's insight into their own behaviour and speech content while lying. *Br. J. Psychol.* 92, 373–389. doi: 10.1348/000712601162248
- Walczyk, J. J., Roper, K. S., Seemann, E., and Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: response time as a cue to deception. *Appl. Cogn. Psychol.* 17, 755–774. doi: 10.1002/acp.914
- Warm, J. S., Parasuraman, R., and Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Hum. Factors* 50, 433–441. doi: 10.1518/001872008X312152
- Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the International Workshop on the Safety of Repetitive Transcranial Magnetic Stimulation, June 5–7, 1996. *Electroencephalogr. Clin. Neurophysiol.* 108, 1–16. doi: 10.1016/S0168-5597(97)00096-8
- Weiss, M., and Lavidor, M. (2012). When less is more: evidence for a facilitative cathodal tDCS effect in attentional abilities. *J. Cogn. Neurosci.* 24, 1826–1833. doi: 10.1162/jocn_a_00248
- Williams, J. A., Pascual-Leone, A., and Fregni, F. (2010). Interhemispheric modulation induced by cortical stimulation and motor training. *Phys. Ther.* 90, 398–410. doi: 10.2522/ptj.20090075
- Zaidi, Z. F. (2010). Gender differences in human brain: a review. *Open Anat. J.* 2, 37–55. doi: 10.2174/1877609401002010037

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 April 2013; accepted: 20 July 2013; published online: 14 August 2013.
Citation: Levasseur-Moreau J, Brunelin J and Fecteau S (2013) Non-invasive brain stimulation can induce paradoxical facilitation. Are these neuroenhancements transferable and meaningful to security services? *Front. Hum. Neurosci.* 7:449. doi: 10.3389/fnhum.2013.00449
Copyright © 2013 Levasseur-Moreau, Brunelin and Fecteau. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Why non-invasive brain stimulation should not be used in military and security services

Bernhard Sehm* and Patrick Ragert*

Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences Leipzig, Leipzig, Germany

*Correspondence: sehm@cbs.mpg.de; ragert@cbs.mpg.de

Edited by:

Elena Rusconi, University College London, UK

Keywords: non-invasive brain stimulation, tDCS, TMS, ethics, military, security

In a recent review article by Levasseur-Moreau et al. (2013), the authors discussed the effects of non-invasive brain stimulation (NIBS) on cognitive functions and proposed a potential application of NIBS in security or military personnel. We believe that this research endeavor is questionable since it might disclose several scientific as well as ethical concerns. In the following, we highlight our reservations about the potential use of NIBS in army and/or security services.

Over the past decades, non-invasive brain stimulation (NIBS) techniques such as transcranial magnetic (TMS) or transcranial direct current stimulation (tDCS) have been extensively used to investigate brain function and brain plasticity in the living human brain. Early studies provided evidence that NIBS is capable of evoking short-lasting modulatory effects on brain functions. Based on this finding, subsequent proof-of-principle studies quickly progressed to also affect motor and cognitive functions. Originally, NIBS techniques were primarily used in basic research to unravel physiological brain processes and/or to establish brain-behavior relationships. The underlying motivation for many researchers is to extend the boundaries of knowledge and to translate findings of basic research into clinical science, that is, to develop new adjuvant therapeutic tools.

In fact, NIBS might be a promising tool for the treatment of neurological and psychiatric diseases (Floel, 2013). For example, Hummel and colleagues showed a beneficial effect of a short period of tDCS in chronic stroke patients on paretic hand function (Hummel et al., 2005). Based on this finding, the authors and other subsequent studies (Lindenber-

et al., 2010) suggested that such interventional strategies in combination with customary rehabilitative treatments may play an adjuvant role in neurorehabilitation. Nevertheless, clinical trials on larger patient samples are still needed to confirm the promising results that have been achieved so far by smaller clinical studies.

Apart from a translation to the clinical settings, it has been suggested to use these techniques for “neuroenhancement” in cognitive abilities or sports, fueling a vivid discussion concerning ethical issues of the use of NIBS in healthy human subjects (Hamilton et al., 2011; Brukamp and Gross, 2012; Cohen Kadosh et al., 2012; Banissy and Muggleton, 2013). However, to our mind, the use of these techniques in military or security personnel goes even a step further and accentuates concerns as compared to the use in “civilians.” First, the use of NIBS in military or security services is problematic with respect to the autonomy of individuals receiving NIBS: In the military context, the risk of coercion is much more pronounced and autonomous decisions cannot always be warranted (Tennison and Moreno, 2012). Second, safety issues might be aggravated in this context and might not only apply to the person receiving NIBS but also to third persons. Both safety and autonomy represent principles that may help to identify ethical problems and guide related decisions (Beauchamp and Childress, 1994; Walker, 2009; Brukamp and Gross, 2012).

WHAT ARE THE LONG-TERM EFFECTS OR SIDE EFFECTS OF NIBS?

The long-term behavioral effects of NIBS are yet unknown. Single NIBS applications typically result in transient effects

on behavior and brain physiology. A few studies, however, indicated, that repeated NIBS applications over several consecutive days during motor or cognitive learning might induce longer-lasting behavioral improvements (Levasseur-Moreau et al., 2013). These findings are certainly of great interest for the application of NIBS in neurorehabilitation, where long-lasting brain changes and associated functional improvements are a desired goal of any treatment. However, we still do not know how specific such changes are and whether improvements in one function may be associated with deterioration in others, as raised by a recent article (Brem et al., 2013). In a clinical setting, patients are under close medical supervision and individually elected for specific treatments, based on a careful assessment of individual risks and benefits. In addition, due to a longitudinal medical monitoring, potential long-term changes may possibly be identified. This, however, does not hold true in military/security context. Therefore, to our mind, it raises ethical questions whether the induction of long-lasting brain changes in healthy individuals, and in particular in military and/or security personnel, should be an aim or even just a tolerated “side effect” of neuroscientific research. Even though hypothetical, the question that comes up is: Do we want to take the risk of changing the brain processing in people who (i) potentially cannot make autonomous decisions concerning the application of NIBS and (ii) are responsible for their own lives as well as the lives of others?

Medical side effects of NIBS described so far in the literature are seldom and usually not severe (with the exception that specific NIBS protocols increase the risk of epileptic seizures). In analogy

to unknown long-term effects discussed above, the risk–benefit ratio of NIBS should be carefully evaluated since potential medical risks especially related to repeated brain stimulation are still not well-known (e.g., do repeated applications of NIBS increase the risk of epileptic seizures?). Therefore, it certainly cannot be excluded that a repeated exposure to NIBS might result in unforeseeable health issues for the “treated” individual. While this concern is not specific to the application of NIBS in military settings, it again might be especially severe, since the individual might not be able to weigh the risks and benefits and make an autonomous decision (Bruckamp and Gross, 2012). On this notion Tennison and Moreno (2012) state that “if a warfighter is allowed no autonomous freedom to accept or decline an enhancement intervention [...] then the ethical implications are immense.”

ARE THE EFFECTS OF NIBS TRANSFERABLE TO THE “REAL WORLD”?

What do we know about the generalization of NIBS-induced effects on everyday life situations? Until now, scientific evidence for NIBS effects has been limited to relatively simplified experimental settings which might not necessarily be valid outside controlled laboratory settings. In order to argue that stimulation of specific brain areas is related to a “meaningful” behavioral effect, researchers usually try to isolate a cognitive process of interest (the dependent variable, e.g., spatial attention) while minimizing or controlling for potential “confounding variables” such as mood changes and so forth. However, there is still limited evidence that NIBS effects can at all be beneficial in real-life situations—where we are subject to complex perceptual, cognitive, and emotional interactions.

Some recent studies investigated the effects of tDCS on visual detection abilities in a task that is specifically designed for military training programs to “familiarize military personnel with the Middle Eastern environment before deployment” (Clark et al., 2012; “*DARWARS Ambush! Threat Detection Task*”). Here, in a so-called “threat detection task” concealed bombs and “enemy combatants” have to

be detected in a virtual reality setting that simulates a Middle Eastern environment. While this might be somewhat more realistic, it still remains a computer simulation and surely cannot mimic real-life situations of soldiers and/or security personnel who may need to make fast decisions under extreme and life-threatening conditions with potentially enormous attentional and/or emotional load. We do not know how NIBS techniques affect human behavior in such complex real-life situations. For example, an unwanted and unexpected modulation of the attentional state, decision-making or emotional factors might negatively affect behavioral outcome. Therefore, the use of these techniques must be considered unsafe in particular for third persons that might be harmed by the actions of “dysregulated” individuals receiving NIBS.

HOW SPECIFIC IS THE MODULATION OF BRAIN FUNCTION USING NIBS?

Despite recent progress, it still remains elusive how specific NIBS protocols act on behavior and/or neural processing. Focal brain stimulation might potentially be suitable for enhancing some abilities in a laboratory setting, but we do not know yet at which costs. As mentioned above, it has been proposed that NIBS performed to enhance a specific ability of interest may be deleterious to another (Hilgetag et al., 2001; Hamilton et al., 2011; Brem et al., 2013). Obviously, due to the limited spatial accuracy of NIBS we do not modulate one segregated brain area that is responsible for one specific function. Instead, recent studies combining NIBS and neuroimaging demonstrate that whole brain functional networks are affected by “focal” stimulation, and increases in functional activity or connectivity of certain brain regions are often accompanied by a decrease in others (Bestmann et al., 2004; Polania et al., 2011; Sehm et al., 2012).

A recent study investigated the effects of tDCS applied over the frontal cortex during a 40-minute vigilance task that was designed to simulate the work of an air traffic controller (Nelson et al., 2012). TDCS over the prefrontal cortex caused a sustained target detection performance thus counteracting a physiological

decrease of vigilance in the volunteering military personnel. However, in the same study, tDCS did not only modulate perceptual sensitivity—in the framework of signal detection theory (McMillan, 2005)—but also induced a liberalization in the decision criterion, that is, the internal criterion to differentiate signal from noise. In a similar way, a study by Pavlidou et al. (2012) reported improved visual discrimination of human and animal motion induced by tDCS over premotor cortex but at the costs of an increase in the false alarm rate. However, another study did find a specific effect of tDCS on perceptual sensitivity and no effect on the decision criterion (Falcone et al., 2012). Thus, the results across studies are inconsistent which might depend on differences in NIBS parameters and/or task design. Nevertheless, they question whether only “basic” perceptual abilities are modulated by NIBS or whether additionally the perceptual decision criterion is affected by brain stimulation. This, however, might be an essential issue in military settings. For example, a liberalization of the decision criterion may result in more “hits” but at the costs of more “false alarms.” In the military setting, a “false alarm” that causes a military reaction might have disastrous consequences.

In this context it might be important to consider questions related to the responsibility of individuals undergoing NIBS whose actions harmed themselves or others. Is a soldier that is receiving NIBS responsible for erroneous decisions? Can “wrong” brain stimulation parameters be blamed? These questions still remain unanswered but have tremendous moral and legal implications.

CONCLUSION

We here critically discussed the potential application of NIBS in military or security services as proposed in a recent article (Levasseur-Moreau et al., 2013). In our opinion, relevant ethical and scientific concerns as outlined in this article question such implications. In this light, we hope that our arguments will contribute to and stimulate a constructive discussion about the potential use of NIBS in military and/or security services.

REFERENCES

- Banissy, M. J., and Muggleton, N. G. (2013). Transcranial direct current stimulation in sports training: potential approaches. *Front. Hum. Neurosci.* 7:129. doi: 10.3389/fnhum.2013.00129
- Beauchamp, T. L., and Childress, J. F. (1994). *Principles of Biomedical Ethics*. New York, NY: Oxford University Press.
- Bestmann, S., Baudewig, J., Siebner, H. R., Rothwell, J. C., and Frahm, J. (2004). Functional MRI of the immediate impact of transcranial magnetic stimulation on cortical and subcortical motor circuits. *Eur. J. Neurosci.* 19, 1950–1962. doi: 10.1111/j.1460-9568.2004.03277.x
- Brem, A. K., Fried, P. J., Horvath, J. C., Robertson, E. M., and Pascual-Leone, A. (2013). Is neuroenhancement by noninvasive brain stimulation a net zero-sum proposition? *Neuroimage*. doi: 10.1016/j.neuroimage.2013.07.038. [Epub ahead print].
- Bruckamp, K., and Gross, D. (2012). “Neuroenhancement - a controversial topic in contemporary medical ethics,” in *Contemporary Issues in Bioethics*, ed D. P. A. Clark (Rijeka: InTech), 39–51.
- Clark, V. P., Coffman, B. A., Mayer, A. R., Weisend, M. P., Lane, T. D., Calhoun, V. D., et al. (2012). TDCS guided using fMRI significantly accelerates learning to identify concealed objects. *Neuroimage* 59, 117–128. doi: 10.1016/j.neuroimage.2010.11.036
- Cohen Kadosh, R., Levy, N., O’shea, J., Shea, N., and Savulescu, J. (2012). The neuroethics of non-invasive brain stimulation. *Curr. Biol.* 22, R108–R111. doi: 10.1016/j.cub.2012.01.013
- Falcone, B., Coffman, B. A., Clark, V. P., and Parasuraman, R. (2012). Transcranial direct current stimulation augments perceptual sensitivity and 24-hour retention in a complex threat detection task. *PLoS ONE* 7:e34993. doi: 10.1371/journal.pone.0034993
- Floel, A. (2013). tDCS-enhanced motor and cognitive function in neurological diseases. *Neuroimage*. doi: 10.1016/j.neuroimage.2013.05.098. [Epub ahead print].
- Hamilton, R., Messing, S., and Chatterjee, A. (2011). Rethinking the thinking cap: ethics of neural enhancement using noninvasive brain stimulation. *Neurology* 76, 187–193. doi: 10.1212/WNL.0b013e318205d50d
- Hilgetag, C. C., Theoret, H., and Pascual-Leone, A. (2001). Enhanced visual spatial attention ipsilateral to rTMS-induced ‘virtual lesions’ of human parietal cortex. *Nat. Neurosci.* 4, 953–957. doi: 10.1038/nn0901-953
- Hummel, F., Celnik, P., Giraux, P., Floel, A., Wu, W. H., Gerloff, C., et al. (2005). Effects of non-invasive cortical stimulation on skilled motor function in chronic stroke. *Brain* 128, 490–499. doi: 10.1093/brain/awh369
- Levasseur-Moreau, J., Brunelin, J., and Fecteau, S. (2013). Noninvasive brain stimulation can induce paradoxical facilitation. Are these neuroenhancements transferable and meaningful to security services? *Front. Hum. Neurosci.* 7:449. doi: 10.3389/fnhum.2013.00449
- Lindenberg, R., Renga, V., Zhu, L. L., Nair, D., and Schlaug, G. (2010). Bihemispheric brain stimulation facilitates motor recovery in chronic stroke patients. *Neurology* 75, 2176–2184. doi: 10.1212/WNL.0b013e318202013a
- McMillan, N. A. (2005). *Detection Theory*. Mahwah, NJ: Erlbaum.
- Nelson, J. T., McKinley, R. A., Golob, E. J., Warm, J. S., and Parasuraman, R. (2012). Enhancing vigilance in operators with prefrontal cortex transcranial direct current stimulation (tDCS). *Neuroimage*. doi: 10.1016/j.neuroimage.2012.11.061. [Epub ahead print].
- Pavlidou, A., Edwards, M., Lange, J., Schnitzler, A., and Bell, J. (2012). “Anodal stimulation of premotor cortex facilitates the recognition of different forms of movements,” in *NeuroVisionen 8 Conference* (Düsseldorf).
- Polania, R., Paulus, W., Antal, A., and Nitsche, M. A. (2011). Introducing graph theory to track for neuroplastic alterations in the resting human brain: a transcranial direct current stimulation study. *Neuroimage* 54, 2287–2296. doi: 10.1016/j.neuroimage.2010.09.085
- Sehm, B., Schaefer, A., Kipping, J., Margulies, D., Conde, V., Taubert, M., et al. (2012). Dynamic modulation of intrinsic functional connectivity by transcranial direct current stimulation. *J. Neurophysiol.* 108, 3253–3263. doi: 10.1152/jn.00606.2012
- Tennison, M. N., and Moreno, J. D. (2012). Neuroscience, ethics, and national security: the state of the art. *PLoS Biol.* 10:e1001289. doi: 10.1371/journal.pbio.1001289
- Walker, T. (2009). What principlism misses. *J. Med. Ethics* 35, 229–231. discussion: 232–223. doi: 10.1136/jme.2008.027227

Received: 22 July 2013; accepted: 21 August 2013; published online: 09 September 2013.

Citation: Sehm B and Ragert P (2013) Why non-invasive brain stimulation should not be used in military and security services. *Front. Hum. Neurosci.* 7:553. doi: 10.3389/fnhum.2013.00553

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Sehm and Ragert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Is it ethical and safe to use non-invasive brain stimulation as a cognitive and motor enhancer device for military services? A reply to Sehm and Ragert (2013)

Jerome Brunelin^{1,2}, Jean Levasseur-Moreau¹ and Shirley Fecteau^{1,3*}

¹ Centre Interdisciplinaire de Recherche en Réadaptation et en Intégration Sociale de l'Université Laval, Centre de Recherche de l'Institut Universitaire en Santé Mentale de Québec, Faculté de Médecine, Université Laval, Québec, QC, Canada

² Centre Hospitalier le Vinatier, Université de Lyon, F-69003, Université Claude Bernard Lyon I, EA 4615, Bron, Lyon, France

³ Berenson-Allen Center for Noninvasive Brain Stimulation, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

*Correspondence: shirley.fecteau@fmed.ulaval.ca

Edited by:

Elena Rusconi, University College London, UK

Keywords: non-invasive brain stimulation, motor function, cognitive function, security, transcranial direct current stimulation (tDCS), neuroenhancement, transcranial magnetic stimulation (TMS)

We appreciate the comment from Sehm and Ragert (2013) on our review article published in the Frontiers Research Topic on *Neuroscience perspectives on Security* (Levasseur-Moreau et al., 2013) and we want to clearly and briefly reaffirm our position to avoid misinterpretations.

The goal of our article was to review data suggesting that non-invasive brain stimulation (NIBS) can enhance performance in healthy volunteers (in line with our expertise) and to focus on motor and cognitive functions that are relevant for security purposes (in line with the Research Topic). We did not take position on whether or not NIBS may eventually serve security because we believe that there are ethics and safety aspects to be studied before considering NIBS as a neuroenhancer device for healthy individuals.

The goal of our review paper was not to examine the ethics and safety of NIBS. We did call for an open and fair debate on the ethics and safety of using NIBS in healthy volunteers because we consider it is our responsibility to at least acknowledge these aspects although they were beyond the scope of our paper. We believe that review articles (as our) should not be considered as encouragement to an irresponsible use of NIBS and we therefore, thank Sehm and Ragert for taking part in this debate.

As neuroscientists we cannot ignore data suggesting to some extent that NIBS might eventually be used as a cognitive enhancer and it is our obligation to discuss their limitations in terms of safety, ethics,

transferability, and meaningfulness as we did in Levasseur-Moreau et al. (2013). We thus, raised potential safety and ethical concerns of using NIBS in healthy participants and we referred readers to articles specifically addressing these major questions (e.g., Illes and Bird, 2006; Forlini et al., 2013).

This debate on ethics and safety on the use of NIBS for cognitive enhancement should certainly be pursued among scientists (e.g., Bikson et al., 2013) but we should also seek participation of policy makers, ethicists and manufacturers, since, whether we like it or not, there is a fast-growing market promoting do-it-yourself brain stimulation devices proposing NIBS for a recreational use, especially transcranial Direct Current Stimulation (tDCS). We thus, welcome the effort of policy makers such as The California Department of Public Health (CDPH) that “warned consumers not to use the unapproved medical device sold on the Internet as a tDCS Home Device Kit” (see <http://www.cdph.ca.gov/Pages/NR13-029.aspx>). As already stated in Levasseur-Moreau et al. (2013), we believe that this discussion should also encompass all potential nonpharmacologic neuroenhancers and brain-boosting drugs that can improve performance of healthy individuals including military personnel.

If our position on the ethics or safety in the use of NIBS in our article has been misinterpreted or is unclear, we hereby want to reaffirm it because these topics strongly matter to us: in our opinion, benefits and

risks in terms of ethics and safety must be clearly weighed before any use of NIBS as a cognitive enhancer in healthy population. NIBS protocols must be reviewed by independent and competent institutional review boards. Stimulation sessions should be delivered by adequately trained staff in a secure environment (e.g., hospital setting) and with strict inclusion and exclusion criteria to ensure safety of participants in accordance with international guidelines (for instance, see Rossi et al., 2009).

REFERENCES

- Bikson, M., Bestmann, S., and Edwards, D. (2013). Transcranial devices are not playthings. *Nature* 501, 167. doi: 10.1038/501167b
- Forlini, C., Hall, W., Maxwell, B., Outram, S. M., Reiner, P. B., Repantis, D., et al. (2013). Navigating the enhancement landscape. Ethical issues in research on cognitive enhancers for healthy individuals. *EMBO Rep.* 14, 123–128. doi: 10.1038/embo.2012.225
- Illes, J., and Bird, S. J. (2006). Neuroethics: a modern context for ethics in neuroscience. *Trends Neurosci.* 29, 511–517. doi: 10.1016/j.tins.2006.07.002
- Levasseur-Moreau, J., Brunelin, J., and Fecteau, S. (2013). Non-invasive brain stimulation can induce paradoxical facilitation. Are these neuroenhancements transferable and meaningful to security services? *Front. Hum. Neurosci.* 7:449. doi: 10.3389/fnhum.2013.00449
- Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., and Safety of TMS Consensus Group. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin. Neurophysiol.* 120, 2008–2039. doi: 10.1016/j.clinph.2009.08.016
- Sehm, B., and Ragert, P. (2013). Why non-invasive brain stimulation should not be used in military and security services. *Front. Hum. Neurosci.* 7:553. doi: 10.3389/fnhum.2013.00553

Received: 12 November 2013; accepted: 29 November 2013; published online: 16 December 2013.

Citation: Brunelin J, Levasseur-Moreau J and Fecteau S (2013) Is it ethical and safe to use non-invasive brain stimulation as a cognitive and motor enhancer device for military services? A reply to Sehm and Ragert (2013).

Front. Hum. Neurosci. 7:874. doi: 10.3389/fnhum.2013.00874

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Brunelin, Levasseur-Moreau and Fecteau. This is an open-access article distributed under the terms of the Creative Commons Attribution License

(CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

ADVANTAGES OF PUBLISHING IN FRONTIERS



FAST PUBLICATION

Average 90 days
from submission
to publication



COLLABORATIVE PEER-REVIEW

Designed to be rigorous –
yet also collaborative, fair and
constructive



RESEARCH NETWORK

Our network
increases readership
for your article



OPEN ACCESS

Articles are free to read,
for greatest visibility



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



GLOBAL SPREAD

Six million monthly
page views worldwide



COPYRIGHT TO AUTHORS

No limit to
article distribution
and re-use



IMPACT METRICS

Advanced metrics
track your
article's impact



SUPPORT

By our Swiss-based
editorial team