Commentary: Principles, Approaches and Challenges of Applying Big Data in Safety Psychology Research

Davide Giusino [1,2], Federico Fraboni [1], Marco De Angelis [1] and Luca Pietrantoni [1,2]*

[1] *Department of Psychology, University of Bologna, Bologna, Italy,* [2] *Interdepartmental Center for Industrial Research in Advanced Mechanical Engineering Applications and Materials Technology, University of Bologna, Bologna, Italy*

A Commentary on

Principles, Approaches and Challenges of Applying Big Data in Safety Psychology Research
by Kang, L., Wu, C., and Wang, B. (2019). Front. Psychol. 10:1596. doi: 10.3389/fpsyg.2019.01596

INTRODUCTION

This commentary builds upon the recent theoretical paper by Kang et al. (2019) to advance the debate currently going on about psychology and big data. The aim is to discuss the feasibility of extending the conceptualization proposed by the authors—i.e., Big Data of Safety Psychology (BDSP)—to other branches of psychology going beyond the only safety domain, ultimately pointing out a *big data of whatever psychology* scenario. This will lead to suggest a perspective enrichment from a solely *big data applied to psychology* paradigm toward a much less advocated *psychology applied to big data*.

BIG DATA OF WHATEVER PSYCHOLOGY

Big Data are generally defined as data being high in volume, velocity, and variety. That is, a huge amount of data is produced at an inedited fast pace coming from broadly diverse sources. Big data are characterized by varying quality (*veracity* issues) and are usually unstructured (raw digital information like texts or images) vs. structured (data traditionally representable on statistical software spreadsheets like numbers or Likert-type measurements). Big Data result from a global *datafication* phenomenon that is impacting contemporary human everyday life (Chen and Wojcik, 2016). Indeed, nowadays people are constantly using various types of online and offline digital ICTs—especially social media—which generate great quantities of data.

Although not free from pitfalls, Big Data are offering benefits to many academic and industrial sectors. This is because information can be inferred, and predictions made, based on the inductive observation of real-world behavioral patterns displayed by unprecedently large datasets. In the history of psychology, the official turning point in big data uptake can be traced to the 2016 *Psychological Methods* special issue entirely devoted to such topic (Harlow and Oswald, 2016). From there on, big data approaches to psychology have become increasingly popular (Jones, 2017), and several scholars have initiated both theoretical and empirical investigation entailing the use of big data within different sub-disciplines of psychology.

Among these, Kang et al. (2019) introduced the BDSP concept to hold how big data adoption and utilization may positively affect workplace health and safety. They define BDSP as "structured, semi-structured, and unstructured datasets formed by psychological index parameters

and behavior, which provide potential and valuable psychological knowledge and rules to solve the psychological issues related to safety with the help of big data technology" (p. 3). Of course, safety is an extremely interesting and practically relevant field for big data implementation as it is aimed at managing risks and preventing accidents. Nevertheless, the paper provides input to wonder why we should limit our scope to the safety domain and not including other ones into the *big data of psychology* discourse. Examples of big data applications are identifiable in numerous areas of psychology, such as organizational (Guzzo et al., 2015; Tonidandel et al., 2016), educational (aka learning analytics; Watson and Christensen, 2017; Maldonado-Mahauad et al., 2018; Viberg et al., 2018; Elia et al., 2019; Shorfuzzaman et al., 2019), marketing (Hopp and Vargo, 2017; Matz and Netzer, 2017; Erceg et al., 2018; Ibrahim and Wang, 2019), personality (Bleidorn et al., 2017; Boyd and Pennebaker, 2017; Gerlach et al., 2018; Hinds and Joinson, 2019), emotion (aka affective computing; D'Mello et al., 2018; Chatterjee et al., 2019; Gruda and Hasan, 2019), psycholinguistics (Ridgeway et al., 2017; Johns, 2019; Luo et al., 2019), clinical (Anestis et al., 2016; Russ et al., 2018), cognitive (Medina and Fischer-Baum, 2017; Bhatia and Walasek, 2019), community (O'Brien, 2016), group (Guadagno et al., 2018), music (Greenberg and Rentfrow, 2017), political (Ma-Kellams et al., 2018), and positive psychology (Luhmann, 2017; Yaden et al., 2018). Additionally, the representation of BDSP as being the intersection of safety science, data science, and psychology, seems to equally fit other psychology branches. For instance, we might say that *big data of organizational psychology* constitute the link between organization science, data science, and psychology. As well, one may state that *big data of political psychology* connect political science, data science, and psychology. Basically, it will suffice to replace the first term of the equation. Finally, the seven listed characteristics, along with the 15 types of BDSP, are also easily transferable to different psychological sub-domains. This line of reasoning leads us to conclude that we are actually in front of a *big data of whatever psychology* (BDWP) landscape.

FROM "BIG DATA APPLIED TO PSYCHOLOGY" TOWARD "PSYCHOLOGY APPLIED TO BIG DATA"

All BDWP literature outputs show one common feature. They all encompass big data as a tool or instrument at the disposal of psychological disciplines. So, no more doubts are around big data usefulness in psychology—and for sure we need further methodological developments and researchers' upskilling to exploit it (Liem et al., 2018; Bleidorn and Hopwood, 2019). Today, the question is no longer whether big data should be used in psychological applications, or what the role of big data in psychology could be.

Therefore, new conceptualization needs regarding the relationship between psychology and big data are stimulated. As a call for discipline self-reflection, we suggest that the question is now becoming what the role of psychology might become in a big data era. This would imply a paradigm shift from *big data applied to psychology* toward *psychology applied to big data*. For instance, it is recognized that big data are associated with individual, cultural, organizational, structural, and technological barriers hindering their acceptance. These entail concerns about privacy and personal information confidentiality (Fast and Jago, 2019), lack of IT infrastructure readiness, and poor analytic skills (Alharthi et al., 2017). These factors can determine resistance to their adoption (Raguseo, 2018), increase people's skepticism toward this technology (Van Rijmenam, 2014), and thus undermine its potential benefits.

How are psychological disciplines going to address similar issues, for example by deploying human-technology interaction frameworks (e.g., Technology Acceptance Model; Davis, 1989) or organizational change management strategies (e.g., Calvard, 2016)? Are psychologists going to only use already available tools or also develop new tailored, targeted ones? Future research can use this kind of questions to start stepping forward within the theoretical, empirical and practical debate about psychology and big data.

AUTHOR CONTRIBUTIONS

DG wrote the first draft of the manuscript. FF, MD, and LP contributed to manuscript revision, read, and approved the submitted version.

ACKNOWLEDGMENTS

REFERENCES

Alharthi, A., Krotov, V., and Bowman, M. (2017). Addressing barriers to big data. *Bus. Horizons.* 60, 285–292. doi: 10.1016/j.bushor.2017.01.002

Anestis, M. D., Law, K. C., Jin, H., Houtsma, C., Khazem, L. R., and Assavedo, B. L. (2016). Treating the capability for suicide: a vital and understudied frontier in suicide prevention. *Suicide Life Threat. Behav.* 47, 523–537. doi: 10.1111/sltb.12311

Bhatia, S., and Walasek, L. (2019). Association and response accuracy in the wild. *Mem. Cogn.* 47, 292–298. doi: 10.3758/s13421-018-0869-6

Bleidorn, W., and Hopwood, C. J. (2019). Using machine learning to advance personality assessment and theory. *Pers. Soc. Psychol. Rev.* 23, 190–203. doi: 10.1177/1088868318772990

Bleidorn, W., Hopwood, C. J., and Wright, A. G. C. (2017). Using big data to advance personality theory. *Curr. Opin. Behav. Sci.* 18, 79–82. doi: 10.1016/j.cobeha.2017.08.004

Boyd, R. L., and Pennebaker, J. W. (2017). Language-based personality: a new approach to personality in a digital world. *Curr. Opin. Behav. Sci.* 18, 63–68. doi: 10.1016/j.cobeha.2017.07.017

Calvard, T. S. (2016). Big data, organizational learning, and sensemaking: theorizing interpretive challenges under conditions of dynamic complexity. *Manage. Learn.* 47, 65–82. doi: 10.1177/1350507615592113

Chatterjee, A., Gupta, U., Chinnakotla, M. K., Srikanth, R., Galley, M., and Agrawal, P. (2019). Understanding emotions in text using deep learning and big data. *Comput. Hum. Behav.* 93, 309–317. doi: 10.1016/j.chb.2018.12.09

Chen, E. E., and Wojcik, S. P. (2016). A practical guide to big data research in psychology. *Psychol. Methods* 21, 458–474. doi: 10.1037/met0000111

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 13, 319–340. doi: 10.2307/249008

D'Mello, S., Kappas, A., and Gratch, J. (2018). The affective computing approach to affect measurement. *Methods Emot. Res.* 10, 174–183. doi: 10.1177/1754073917696583

Elia, G., Solazzo, G., Lorenzo, G., and Passiante, G. (2019). Assessing learners' satisfaction in collaborative online courses through a big data approach. *Comput. Hum. Behav.* 92, 589–599. doi: 10.1016/j.chb.2018.04.033

Erceg, N., Burghart, M., Cottone, A., Lorimer, J., Manku, K., Pütz, H., et al. (2018). The effect of moral congruence of calls to action and salient social norms on online charitable donations: a protocol study. *Front. Psychol.* 9:1913. doi: 10.3389/fpsyg.2018.01913

Fast, N. J., and Jago, A. S. (2019). Privacy matters... Or does it? Algorithms, rationalization, and the erosion of concern for privacy. *Curr. Opin. Psychol.* 31, 44–48. doi: 10.1016/j.copsyc.2019.07.011

Gerlach, M., Farb, B., Revelle, W., and Nunes Amaral, L. A. (2018). A robust data-driven approach identifies four personality types across four large data sets. *Nat. Hum. Behav.* 2, 735–742. doi: 10.1038/s41562-018-0419-z

Greenberg, D. M., and Rentfrow, P. J. (2017). Music and big data: a new frontier. *Curr. Opin. Behav. Sci.* 18, 50–56. doi: 10.1016/j.cobeha.2017.07.007

Gruda, D., and Hasan, S. (2019). Feeling anxious? Perceiving anxiety in tweets using machine learning. *Comput. Hum. Behav.* 98, 245–255. doi: 10.1016/j.chb.2019.04.020

Guadagno, R. E., Nelson, M., and Lee, L. L. (2018). Peace data standard: a practical and theoretical framework for using technology to examine intergroup interactions. *Front. Psychol.* 9:734. doi: 10.3389/fpsyg.2018.00734

Guzzo, R. A., Fink, A. A., King, E. B., Tonidandel, S., and Landis, R. S. (2015). Big data recommendations for industrial–organizational psychology. *Ind. Org. Psychol.* 8, 491–508. doi: 10.1017/iop.2015.40

Harlow, L. L., and Oswald, F. L. (2016). Big data in psychology: introduction to special issue. *Psychol. Methods* 21, 447–457. doi: 10.1037/met0000120

Hinds, J., and Joinson, A. (2019). Human and computer personality prediction from digital footprints. *Curr. Opin. Psychol. Sci.* 28, 204–211. doi: 10.1177/0963721419827849

Hopp, T., and Vargo, C. J. (2017). Does negative campaign advertising stimulate uncivil communication on social media? Measuring audience response using big data. *Comput. Hum. Behav.* 68, 368–377. doi: 10.1016/j.chb.2016.11.034

Ibrahim, N. F., and Wang, X. (2019). Decoding the sentiment dynamics of online retailing customers: time series analysis of social media. *Comput. Hum. Behav.* 96, 32–45. doi: 10.1016/j.chb.2019.02.004

Johns, B. T. (2019). Mining a crowdsourced dictionary to understand consistency and preference in word meanings. *Front. Psychol.* 10:268. doi: 10.3389/fpsyg.2019.00268

Jones, M. N. (2017). "Developing cognitive theory by mining large-scale naturalistic data", in *Big Data in Cognitive Science*, ed M. N. Jones (New York, NY: Taylor and Francis). p. 1–12.

Kang, L., Wu, C., and Wang, B. (2019). Principles, approaches and challenges of applying big data in safety psychology research. *Front. Psychol.* 10:1596. doi: 10.3389/fpsyg.2019.01596

Liem, C. C. S., Langer, M., Demetriou, A., Hiemstra, A. M. F., Wicaksana, A. S., Born, M. P., et al. (2018). "Psychology meets machine learning: Interdisciplinary perspectives on algorithmic job candidate screening," in

Explainable and Interpretable Models in Computer Vision and Machine Learning, eds H. J. Escalante, S. Escalera, I. Guyon, X. Baró, Y. Güçlütürk, U. Güçlü, and M. van Gerven (Cham: Springer), 197–253.

Luhmann, M. (2017). Using big data to study subjective well-being. *Curr. Opin. Behav. Sci.* 18, 28–33. doi: 10.1016/j.cobeha.2017.07.006

Luo, M., Robbins, M. L., Martin, M., and Demiray, B. (2019). Real-life language use across different interlocutors: a naturalistic observation study of adults varying in age. *Front. Psychol.* 10:1412. doi: 10.3389/fpsyg.2019.01412

Ma-Kellams, C., Bishop, B., Zhang, M. F., and Villagrana, B. (2018). Using "big data" versus alternative measures of aggregate data to predict the U.S. 2016 presidential election. *Psychol. Rep.* 121, 726–735. doi: 10.1177/0033294117736318

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., and Munoz-Gama, J. (2018). Mining theory-based patterns from big data: identifying self-regulated learning strategies in Massive Open Online Courses. *Comput. Hum. Behav.* 80, 179–196. doi: 10.1016/j.chb.2017.11.011

Matz, S. C., and Netzer, O. (2017). Using big data as a window into consumers' psychology. *Curr. Opin. Behav. Sci.* 18, 7–12. doi: 10.1016/j.cobeha.2017.05.009

Medina, J., and Fischer-Baum, S. (2017). Single-case cognitive neuropsychology in the age of big data. *Cogn. Neuropsychol.* 34, 440–448. doi: 10.1080/02643294.2017.1321537

O'Brien, D. T. (2016). Lamp lighters and sidewalk smoothers: how individual residents contribute to the maintenance of the urban commons. *Am. J. Commun. Psychol.* 58, 391–409. doi: 10.1002/ajcp.1209e

Raguseo, E. (2018). Big data technologies: an empirical investigation on their adoption, benefits and risks for companies. *Int. J. Inform. Manage.* 38, 187–195. doi: 10.1016/j.ijnfomgt.2017.07.008

Ridgeway, K., Mozer, M. C., and Bowles, A. R. (2017). Forgetting of foreign-language skills: a corpus-based analysis of online tutoring software. *Cogn. Sci.* 41, 924–949. doi: 10.1111/cogs.12385

Russ, T. C., Woelbert, E., Davis, K. A. S., Hafferty, J. D., Ibrahim, Z., Inkster, B., et al. (2018). How data science can advance mental health research. *Nat. Hum. Behav.* 3, 24–32. doi: 10.1038/s41562-018-0470-9

Shorfuzzaman, M., Hossain, M. S., Nazir, A., Muhammad, G., and Alamri, A. (2019). Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment. *Comput. Hum. Behav.* 92, 578–588. doi: 10.1016/j.chb.2018.07.002

Tonidandel, S., King, E. B., and Cortina, J. M. (2016). *Big Data at Work: The Data Science Revolution and Organizational Psychology*. New York, NY: Routledge.

Van Rijmenam, M. (2014). *Think Bigger: Developing a Successful Big Data Strategy for Your Business*. New York, NY: Amacom.

Viberg, O., Hatakka, M., Bälter, O., and Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Comput. Hum. Behav.* 89, 98–110. doi: 10.1016/j.chb.2018.07.027

Watson, R. J., and Christensen, J. L. (2017). Big data and student engagement among vulnerable youth: a review. *Curr. Opin. Behav. Sci.* 18, 23–27. doi: 10.1016/j.cobeha.2017.07.004

Yaden, D. B., Eichstaedt, J. C., and Medaglia, J. D. (2018). The future of technology in positive psychology: methodological advances in the science of well-being. *Front. Psychol.* 9:962. doi: 10.3389/fpsyg.2018.00962