A hidden ambiguity of the term "feedback" in its use as an explanatory mechanism for psychophysical visual phenomena

Talis Bachmann *

Institute of Public Law, University of Tartu, Tartu, Estonia
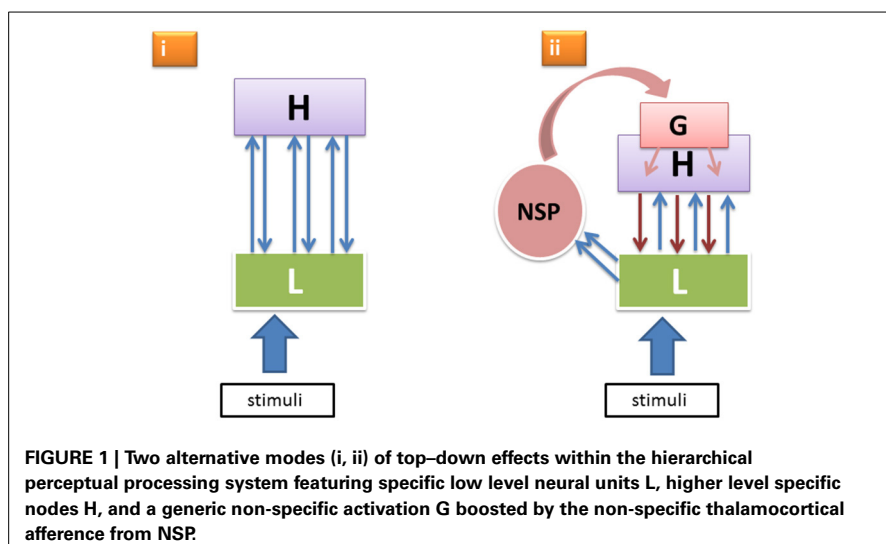**Correspondence: talis.bachmann@ut.ee*

Over the last decades many researchers have used concepts like "feedback," "reentrance," "backpropagation," "top–down (modulation)," or "reverse hierarchy" to specify the mechanisms that underlie various visual phenomena (e.g., Di Lollo et al., 2000; Lamme and Roelfsema, 2000; Pascual-Leone and Walsh, 2001; Supèr et al., 2001; Ro et al., 2003; Ahissar and Hochstein, 2004; Bar et al., 2006; Fahrenfort et al., 2007; Koivisto, 2012). An incomplete list of these phenomena includes visual (object substitution) masking, shape discrimination, illusory contours, illusory motion, priming effects, etc. Empirical evidence or theoretical argumentation in favor of the suggested mechanismic explanations mainly consists in finding or postulating an association between a temporally delayed, secondary activation of lower level neural units with correct reports of target stimuli, even though the higher level neural units in the processing hierarchy were already activated earlier. On that basis, feedforward processing has been argued to be insufficient for target perception. However, in most of the studies the relative *temporal order* of activity at different levels alone is taken as proof of reentrant modulation without precisely measuring the *neural sources* of this top–down effect. In principle, it is equally possible that the source of the higher level activity from which the top–down signals are sent back to earlier feature-encoding neural units (i) is specifically linked to those features by virtue of constituting the higher level nodes associated with specific attributes

of the target stimulus (thus mediating feature-binding for object integration) or (ii) is not specifically linked in this manner. In the latter case, the *source* of top–down modulation may be the result of the arousal or alerting boost triggered by the target stimulus via feedforward collateral activation of subcortical reticulo-thalamic units, which in turn is followed by the cortical spread of the thalamocortical activation, including the downpropagation of the non-specific wave of modulation to the early cortical areas. The non-specific system functions include arousal, attentional modulation, intercortical synchronization of neural activity, bringing the preconsciously processed specific content to awareness, "event-holding" the content in working memory, and alerting subjects to newly appearing objects and changes (Magoun, 1958; Purpura, 1970; Purpura and Schiff, 1997; Jones, 2001; Llinás and Ribary, 2001; Van der Werf et al., 2002; Ribary, 2005; Schiff et al., 2013; Saalmann, 2014). This non-specific system (NSP) targets layer-1 apical dendrites of the layer-5 and -6 pyramidal neurons. But since NSP-modulation is directed at the cortical neurons with specific representational functions, its function may go unacknowledged because the cortical units, when activated by NSP-modulation, can produce content-specific subjective effects misleading us to believe that the entire process has been specific throughout.

The focus of the present paper will be on the experimental-behavioral and neurobiological evidence in comparing the two processing modes, (i) and (ii), with

arguments from computational modeling left for some other occasion.

It is known that reticulo-thalamic, intralaminar and other matrix cells of the NSP project more heavily to lateral and frontal cortical areas and less so to the primary visual areas. (Even when rare examples of direct intralaminar-thalamic input to V1 were documented, these afferents were found to be much sparser than the more frontal ones—Miller and Benevento, 1979.) Moreover, this more rostrally directed thalamo-cortical flow can cause cortical responses as fast as or even faster than the afferent volleys through the specific geniculo-cortical pathways ignite primary visual cortical responses strongly enough (Kennedy and Baleydier, 1977; Kaufman and Rosenquist, 1985; Herkenham, 1986; Cruikshank et al., 2012; Liang et al., 2013; Saalmann, 2014). Thus, the primary cortical areas receive NSP-modulation not directly, but via the higher level cortical neurons that project onto apical parts of the layer-5 pyramidal neurons in the lower cortical areas. Consequently, as illustrated in **Figure 1**, we have two principal modes through which lower level neural units L responsible for encoding sensory features of perceptual objects receive top–down input from higher levels H: (i) from the specific nodes in H that were previously activated by L in a cortical feedforward manner and that now send reentrant signals back to L (here the feedforward-reentrant loop pertains to the specific sensory-perceptual attributes constituting a perceptual object LH); and (ii) from the

FIGURE 1 | Two alternative modes (i, ii) of top–down effects within the hierarchical perceptual processing system featuring specific low level neural units L, higher level specific nodes H, and a generic non-specific activation G boosted by the non-specific thalamocortical afference from NSP.

generic nodes G that were activated by the boost of the NSP directed at the more frontal and mid-level cortical neurons that now send their downpropagating wave to the lower level visual areas, including L.

When analyzing the experimental data from most of the studies that propose *specific* top–down linkages (i), there is no direct evidence that would invalidate the alternative, *non-specific* theory of downpropagation (ii). The specificity of visual experiences is due to the fact that the NSP-modulation arrives at specific early units L and may not be due to the specificity of the higher level from where this modulation arrives. Although the direct input from NSP to L may be weak, the top–down input from higher levels H/G driven by NSP may be strong enough to emphasize the specificity of the visual experience encoded in L. The pending task should be to try disentangle these two explanations experimentally. The experiments should ascertain whether the two modes of top–down modulation are incompatible or mutually complementary. In the latter case—how the two types of downpropagation are specifically combined and what relative roles each of them has? It is also possible that the standard views of reentrance (e.g., Di Lollo et al., 2000; Lamme and Roelfsema, 2000) may be valid in some empirical instances, difficult to ascertain in some other cases, and incompatible with the neurophysiological realities of processing in different experiments. Let me comment on some examples of

typical experiments aimed at supporting the standard views of reentrance listed below and see whether version (i) should be exclusively preferred or whether versions (i) and (ii) both are compatible with the experimental results.

1. In typical object substitution masking (OSM) a target stimulus (e.g., a Landolt C) is presented together with four dots that surround the target. When after a very brief delay the target is switched off, the four dots either are also switched off or remain displayed for varing duration acting as a post-mask (the simultaneous onset, asynchronous offset condition.) The delayed-offset condition leads to strong masking but in the simultaneous-offset condition masking is weak. The classic theory of OSM (Di Lollo et al., 2000, but see Põder, 2013) explains this by a reentrant model (a variety of model i) according to which target-activated units at level H activated by the target send reentrant signals back to level L in order to test whether levels H and L are consistent in representing the target. If mismatch is registered (e.g., when target signals do not arrive anymore and mask signals arrive instead), the iterative feedforward-reentrant cycles are interrupted and new iterative "hypothesis testing" begins for the new object— the mask. Because cycles of reentrance are necessary for registration of the stimulus in awareness, the target is not consciously perceived when reentrant

testing is prematurely interrupted by the stronger top–down mask signal. However, when mask's offset is synchronous with that of the target, the target-plus-mask is a composite object that provides both level L and H contents; hence, the target can be extracted from the composite representation that is maintained through the feedword-reentrant cycles. Let us see how the model (ii) works for OSM. Presentation of target evokes specific signaling along L-H vertical axis and also a collaterally ignited boost of NSP modulation. (NSP is necessary for awareness of the specific contents represented by L and H.) When asynchroneous-offset mask remains in view and target signals do not arrive anymore, the top–down activation G that was initiated fast at higher levels, but takes time to become active at lower levels "finds" mask related activity in L, but the target related activity has decayed already realtive to the mask activity, because the target was switched off earlier. Although the level G activity is non-specific, when its downpropagating generic influence reaches L it helps emphasize mask features because level L units themselves *are* specific. The mask-object representation becomes consciously perceived instead of target. Thus, models (i) and (ii) both are usable. At this point one may ask why not follow Ockham's rule and take the simpler one (i), i.e., the one with fewer hypotheses? However, the G units are important because neurobiological evidence has overwhelmingly shown that NSP is necessary for awareness of the specific contents represented by L and H.

2. In Lamme et al. (2000) monkeys were trained to discriminate visual targets. V1 responses began to differentiate the "seen" from the "unseen" trials after 125 ms. In subsequent studies occipital ERPs in humans differentiated visibility of masked targets after 109–141 ms or peaked at about 160 ms (Fahrenfort et al., 2007, 2008). Again, a variety of model (i) was used for explaining the results because specifically the temporally late target related activity at level L (which followed earlier time epochs sufficient for level H to have become

active in target processing) were associated with correct discrimination. And again, model (ii) can explain these empirical results: the late part of neural activity at L which is enhanced in trials where target is successfully discriminated may be modulated by the top–down process G passed down through levels H (or even bypassing stimulus-specific level H units either via direct fibers or level H units different from the stimulus-related ones).

3. Temporal precedence of high-level MEG activity which discriminated correct and incorrect target processing over low level activity in the study by Bar et al. (2006) was also interpreted as a variety of model (i). However, if activity of G at frontal sites fluctuates (fluctuation of the thalamocortical NSP activity is a norm rather than an exception) and dictates whether the top–down modulation is stronger or weaker, these experimental results can be interpreted also according to model (ii).

4. Ro et al. (2003) utilized transcranial magnetic stimulation (TMS) in a metacontrast masking paradigm and showed that TMS of visual cortex, when timed to produce visual suppression of an annulus (a metacontrast mask), induced recovery of a target disc which was imperceptible when TMS was not used. Moreover, TMS suppression of an annulus was more pronounced when a disk preceded it than when an annulus was presented alone. The authors assume that when the later activity, supposedly reflecting the reentrant effects is suppressed then target perceptibility can be reinstated. They argue that a prior visual stimulus can influence subsequent perception at early stages of visual encoding via feedback projections, supporting model (i). Alternatively, model (ii) can also be applied. It is known that a preceding brief stimulus (e.g., target) speeds up perception of the following stimulus (e.g., mask) (Bachmann, 1989; Scharlau, 2007). When target disc was presented before mask it may have speeded up masking annulus processing by presetting NSP modulation for its signals. This in turn may have optimized the effective processing delay

so as to coincide with the maximal TMS effect.

5. Up to now, both models appear to be equally applicable, but model (ii) provides an explanation of the results of an elegant experiment carried out by Wu et al. (2009) that model (i) cannot as readily provide. Capitalizing on the motion-induced blindness (MIB) phenomenon (Bonneh et al., 2001), where a static visual target-object continuously presented on a rotating background periodically disappears from awareness, they showed that a flashed stimulus that *caused* reappearance in awareness of the target was perceived *after* the reappearance of the target in consciousness. (The temporal value of reversal was about 100 ms, which is the value assumed to characterize the full cycle of reentrance based visual processing for awareness.) The temporal advantage of updating the conscious representation from the preexisting unconscious representation of the invisible static target was explained by a version of model (i), invoking reentry of neural signals after the first feed-forward sweep for a stimulus to be consciously perceived. Thus, MIB, by blocking reentry signals, prevents awareness. In Bachmann and Aru (2009) we pointed out some inconsistencies of this explanation and offered an explanation in terms of model (ii). When an object fades from awareness by MIB, its L and H level activity will be sustained because cortical specific signals are constantly present, but now it is dissociated from NSP-activity. When the flashed object is presented, the L/H process for representation of the flash occurs in parallel with a boost of the NSP-process igniting G. G leads to binding of the already present preconscious L/H-activity of the target with global consciousness-level representation. This process takes little time, because there is no need for build-up of the content-specific L/H representation of the target; consequently, its rapid reappearance in consciousness. The flashed object appears in consciousness not as fast, because its corresponding coherent L/H-representation must be built up, which takes time. The G that services *target* awareness has L/H

content of the target ready on the "waiting list" but the G process has to wait as a "dummy process" until the L/H contents of the *flashed object* are ready to be modulated.

It appears that experiments have difficulty in distinguishing between the two models. This raises the question whether a computational/mathematical argument could be developed that allows to test different predictions about experimental data on the basis of the two models. Sadly, space does not allow me to dwell into this important perspective which must be dealt with in future research.

CONCLUSION

In this opinion paper I argued for the view that in the majority of the standard experimental studies set to support the model of top–down processing featuring exclusively the specific system components also the combined non-specific/specific model seems equally valid.

REFERENCES

Ahissar, M., and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends Cogn. Sci.* 8, 457–464. doi: 10.1016/j.tics.2004.08.011

Bachmann, T. (1989). Microgenesis as traced by the transient paired-forms paradigm. *Acta Psychol.* 70, 3–17. doi: 10.1016/0001-6918(89)90056-5

Bachmann, T., and Aru, J. (2009). Mechanisms of consciousness: why the effect can precede its cause in awareness – Comentary on Wu et al. *Curr. Biol.* Available online at: http://www.cell.com/current-biology/comments/S0960-9822%2809%2901836-3

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, M., et al. (2006). Top-down facilitation of visual recognition. *Proc. Natl. Acad. Sci. U.S.A.* 103, 449–454. doi: 10.1073/pnas.0507062103

Bonneh, Y. S., Cooperman, A., and Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature* 411, 798–801. doi: 10.1038/35081073

Cruikshank, S. J., Ahmed, O. J., Stevens, T. R., Patrick, S. L., Gonzalez, A. N., Elmaleh, M., et al. (2012). Thalamic control of Layer 1 circuits in prefrontal cortex. *J. Neurosci.* 32, 17813–17823. doi: 10.1523/JNEUROSCI.3231-12.2012

Di Lollo, V., Enns, J. T., and Rensink, R. A. (2000). Competition for consciousness among visual events: the psychophysics of reentrant visual processes. *J. Exp. Psychol. Gen.* 129, 481–507. doi: 10.1037/0096-3445.129.4.481

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. *J. Cogn. Neurosci.* 19, 1488–1497. doi: 10.1162/jocn.2007.19.9.1488

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2008). The spatiotemporal profile of cortical processing leading up to visual perception. *J. Vis.* 8, 12.1–12.12. doi: 10.1167/8.1.12

Herkenham, M. (1986). "New perspectives on the organization and evolution of nonspecific thalamocortical projections," in *Sensory-Motor Areas and Aspects of Cortical Connectivity,* Chapter 11, eds E. G. Jones and A. Peters (New York, NY: Plenum Press), 403–423. doi: 10.1007/978-1-4613-2149-1_11

Jones, E. G. (2001). The thalamic matrix and thalamocortical synchrony. *Trends Neurosci.* 24, 595–601. doi: 10.1016/S0166-2236(00)01922-6

Kaufman, E. F. S., and Rosenquist, A. C. (1985). Efferent projections from thalamic intralaminar nuclei in the cat. *Brain Res.* 335, 257–279. doi: 10.1016/0006-8993(85)90478-0

Kennedy, H., and Baleydier, C. (1977). Direct projections from thalamic intralaminar nuclei to extra-striate visual cortex in the cat traced with horseradish peroxidase. *Exp. Brain Res.* 28, 133–139. doi: 10.1007/BF00237091

Koivisto, M. (2012). Is reentry critical for visual awareness of object presence? *Vis. Res.* 63, 43–49. doi: 10.1016/j.visres.2012.05.001

Lamme, V. A. F., Supèr, H., Landman, R., Roelfsema, P. R., and Spekreijse, H. (2000). The role of primary visual cortex (V1) in visual awareness. *Vis. Res.* 40, 1507–1521. doi: 10.1016/S0042-6989(99)00243-6

Lamme, V. A., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/S0166-2236(00)01657-X

Liang, M., Mouraux, A., and Iannetti, G. D. (2013). Bypassing primary sensory cortices–A direct thalamocortical pathway for transmitting salient sensory information. *Cereb. Cortex* 23, 1–11. doi: 10.1093/cercor/bhr363

Llinás, R., and Ribary, U. (2001). Consciousnes and the brain. The thalamocortical dialogue in health and disease. *Ann. N.Y. Acad. Sci.* 929, 166–175. doi: 10.1111/j.1749-6632.2001.tb05715.x

Magoun, H. W. (1958). *The Waking Brain.* Springfield, Ill: C.C.Thomas. doi: 10.1037/11149-000

Miller, J. W., and Benevento, L. A. (1979). Demonstration of a direct projection from the IL central lateral nucleus to the primary visual cortex. *Neurosci. Lett.* 14, 229–234. doi: 10.1016/0304-3940(79)96153-6

Pascual-Leone, A., and Walsh, V. (2001). Fast back-projections from the motion to the primary visual area necessary for visual awareness. *Science* 292, 510–512. doi: 10.1126/science.1057099

Põder, E. (2013). Attentional gating models of object substitution masking. *J. Exp. Psychol. Gen.* 142, 1130–1141. doi: 10.1037/a0030575

Purpura, D. (1970). "Operations and processes in thalamic and synaptically related neural subsystems," in *The Neurosciences. Second Study Program,* ed F. O. Schmitt (New York, NY: Rockefeller University Press), 458–470.

Purpura, K. P., and Schiff, N. D. (1997). The thalamic intralaminar nuclei: a role in visual awareness. *Neuroscientist* 3, 8–15.

Ribary, U. (2005). "Dynamics of thalamo-cortical network oscillations and human perception," in *Progress in Brain Research 150,* ed S. Laureys (Amsterdam: Elsevier), 127–142. doi: 10.1016/S0079-6123(05)50010-4

Ro, T., Breitmeyer, B., Burton, P., Singhal, N. S., and Lane, D. (2003). Feedback contributions to visual awareness in human occipital cortex. *Curr. Biol.* 11, 1038–1041. doi: 10.1016/S0960-9822(03)00337-3

Saalmann, Y. B. (2014). Intralaminar and medial thalamic influence on cortical synchrony, information transmission and cognition. *Front. Syst. Neurosci.* 8:83. doi: 10.3389/fnsys.2014.00083

Scharlau, I. (2007). Perceptual latency priming: a measure of attentional facilitation. *Psychol. Res.* 71, 678–686. doi: 10.1007/s00426-006-0056-4

Schiff, N. D., Shah, S. A., Hudson, A. E., Nauvel, T., Kalik, S. F., and Purpura, K. P. (2013). Gating of attentional effort through the central thalamus. *J. Neurophysiol.* 109, 1152–1163. doi: 10.1152/jn.00317.2011

Supèr, H., Spekreijse, H., and Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in figure-ground activity of monkey primary visual cortex (V1). *Nat. Neurosci.* 4, 304–310. doi: 10.1038/85170

Van der Werf, Y. D., Witter, M. P., and Groenewegen, H. J. (2002). The intralaminar and midline nuclei of the thalamus. Anatomical and functional evidence for participation in processes of arousal and awareness. *Brain Res.Brain Res. Rev.* 39, 107–140. doi: 10.1016/S0165-0173(02)00181-9

Wu, C.-T., Busch, N. A., Fabre-Thorpe, M., and VanRullen, R. (2009). The temporal interplay between conscious and unconscious perceptual streams. *Curr. Biol.* 19, 2003–2007. doi: 10.1016/j.cub.2009.10.017