COMMENTS

Comment on "Competition for consciousness among visual events: The psychophysics of reentrant visual processes" (Di Lollo, Enns and Rensink, 2000)

Gregory Francis
Purdue University, Department of Psychological Sciences

Frouke Hermens
Universität Oldenburg

Di Lollo, Enns & Rensink (2000) reported properties of masking that they claimed were inconsistent with all current models. We show, through computer simulation, that many current models can account for their data. Although Di Lollo *et al.* argued that their data could only be accounted for with models that incorporate reentrant processing, we show that reentrant processing is not necessary.

This formatted manuscript was created by the author and does not correspond exactly to the manuscript published in the journal. References to page numbers should refer to the published article.

Introduction

Visual masking has been used both to study characteristics of visual perception and as a tool to explore other aspects of cognition (see Breitmeyer & Ögmen (2000) and Enns & Di Lollo (2000) for recent reviews). Given the long history and strong interest in masking, it is significant when fundamentally new properties of masking are discovered. In a series of publications, Enns and Di Lollo (1997, 2000) and their colleagues (Bischof & Di Lollo, 1995; Di Lollo, Bischof & Dixon, 1993; Di Lollo, Enns & Rensink, 2000) reported on three properties of masking. Experimental data that reveal all of these effects can be seen in Figure 1a.

The first property is referred to as common-onset masking because the target and mask appear at the same time. Many experiments on masking have focused on situations where the stimulus onset asynchrony (SOA) between the target and the mask are varied. Often the strongest masking occurs for a positive SOA, when the mask follows the target. Di Lollo *et al.* (1993) and Bischof and Di Lollo (1995) reported that strong masking occurs with SOA equal zero if the mask stimulus continues to be presented after the target stimulus turns off. The longer the duration that the mask continues after target offset, the stronger the masking effect. These effects are visible in Figure 1a, which plots percentage correct identifications of a feature of the target as a function of the mask alone duration. The data are reproduced from the third experiment of Di Lollo *et al.* (2000). For at least some of the curves, the percentage correct drops substantially with longer mask alone durations. While previous studies had reported masking for SOA equal zero (e.g., Spencer & Shuntich, 1970) and effects of mask duration (e.g., Breitmeyer, 1978), the experimental conditions were quite different from these new experiments.

A second property is the existence of masking effects with sparse masks. In many masking studies the target and mask are roughly equivalent in size, intensity, and duration. Enns and Di Lollo (1997) and Di Lollo *et al.* (2000) reported strong masking effects when the mask consisted of only four small dots on the corners of an imaginary square around the target. This kind of mask was used to produce the data in Figure 1a. This was not an entirely new finding as some masking from sparse masks occurred in studies by Werner (1935), Sherrick and Dember (1970) and Gilden, MacDonald and Lasaga (1988); but in combination with other properties it suggested to Enns and Di Lollo (1997) a new type of masking.

The third property is the role that attention seems to play in revealing the effect of a sparse mask. The separate lines in Figure 1a indicate conditions with varying numbers of possible

*Gregory Francis, Department of Psychological Sciences, Purdue University; Frouke Hermens, Institut für Kognitionsforschung, Universität Oldenburg, Germany.
Frouke Hermens is now at the Department of Medical Physics and Biophysics, University of Nijmegen, The Netherlands.
This material is based on work supported by the National Science Foundation under Grant 0108905. Gregory Francis was supported by a fellowship at the Hanse Wissenschaftskolleg, Delmenhorst, Germany.
Correspondence concerning this article should be addressed to Gregory Francis, Department of Psychological Sciences, 1364 Psychological Sciences Building, Purdue University, West Lafayette, Indiana 47907-1364. E-mail: gfrancis@psych.purdue.edu.

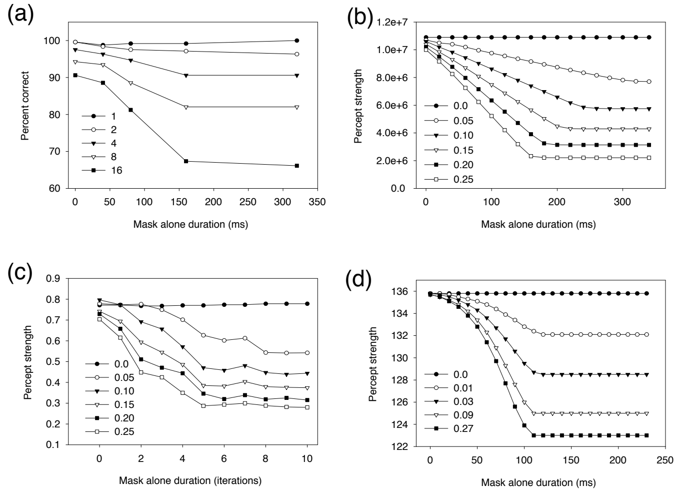

Figure 1: Masking effects for common-onset of the target and mask, a sparse mask, and varying target set sizes. (a) Experimental data from Di Lollo *et al.* (2000). The legend indicates the target set size. Masking occurs for a combination of large set size and long mask alone duration. (b)–(d) Simulation results from models of masking. The legend in each graph indicates the ratio of the mask signal intensity to the (fixed) target signal intensity; this value is hypothesized to be related to the effect of attentional focus that varies with set size. (b). Simulation results from the model of Weisstein (1968, 1972). (c) Simulation results from the model of Bridgeman (1971, 1978). (d) Simulation results from the model of Francis (1997).

targets (set size). In this experiment the mask also functioned as a cue to indicate which possible target item needed to be used for the observer's report. As Figure 1a shows, with a sparse mask, strong masking effects only seem to appear for a combination of long mask duration and large target set size. Attentional effects resulting from variation in target set size had been studied in masking before (e.g. Weisstein, 1966), but not in combination with common-onset masking and a sparse mask.

On the basis of these findings, and other related experimental results, Enns and Di Lollo (1997, 2000) and Di Lollo *et al.* (2000) concluded that their set of results is incompatible with current models and theories of masking and that new models based on reentrant processing were required. Contrary to this conclusion, we show that almost all current mathematical theories of backward visual masking can account for the properties of common-onset masking and the existence of masking with sparse masks. In addition, at least for the data in Figure 1a, it is fairly trivial to extend these models to account for the effects of target set size. Moreover, we show that reentrant processing is not necessary to account for the data.

Model simulations

Three out of four quantitative models of masking can reproduce the basic properties of masking found in Figure 1a. These include the models proposed by Weisstein (1968, 1972), Bridgeman (1971, 1978), and Francis (1997). The only quantitative model that cannot account for these effects is one proposed by Anbar and Anbar (1982), which, in its current form, is insensitive to variations in mask duration. For all of our simulations, we made no changes to the models or model parameters whatsoever. Details of the model equations and parameters can be found in Francis (2000).

Simulating common-onset masking is directly modeled by the relative timing of target and mask signals in the models. Simulating the presence of a sparse mask depends on the model. Both the Weisstein (1968, 1972) and Bridgeman (1971, 1978) models have insufficient representation of spatial extent and layout of the target and mask stimuli to directly portray a sparse mask. As a result, for these models the representation of a sparse mask is coded with a mask that has a weak intensity. The Francis (1997) model does include an explicit representation of extent and layout of the target and mask stimuli. To emulate a four-dot mask, the simulations used a sparse mask of the same type used to produce the data in Figure 10 of Francis (1997), which accounted for data by Sherrick and Dember (1970) on effects of mask completeness. The sparse mask contained only 28% of the pixels of a full mask.

None of the models make particular claims about the computational effect of attentional focus, and we expect that there are several possibilities that could be integrated into existing models. We considered a hypothesis that fits naturally into the framework of current models. We supposed that attentional focus on the cued target acts to prevent the mask from having a strong impact on the target signal. Decreases in the target set size thus would lead to a weaker mask signal and weaker overall masking. Attentional focus is thus modeled as a weakening of the mask intensity.

Figures 1b–d show simulation results under the common-onset masking paradigm for the Weisstein (1968, 1972), Bridgeman (1971, 1978) and Francis (1997) models, respectively. Here the separate curves are for different ratios of the intensity of the mask signal relative to the target signal (the actual intensities of the signals depended on the model). To map these results to the experimental data in Figure 1a we are hypothesizing that smaller target set sizes correspond to weaker mask signals (smaller ratios). With this hypothesis, the pattern of results is basically the same as for the experimental data in Figure 1a. No doubt better quantitative fits could be produced by varying model parameters. In particular, a change of parameters could rescale the x-axis so that it covers the same time range as the experimental data.

These simulation results refute the claim made by Di Lollo *et al.* (2000) that no current models can account for their experimental data. Indeed, the basic mechanisms to account for the experimental data were embedded in models that are nearly 30 years old. This discrepancy between the conclusions of Di Lollo *et al.* (2000) and the actual capabilities of existing models to account for the experimental data occurs because Di Lollo *et al.* present one theory of masking as representative of all theories of masking. They suggested that all current theories are based on the hypothesis that masking interactions result solely from inhibition generated by transient signals. Such transient inhibition would be insensitive to variations in mask duration.

Di Lollo *et al.* further argue that all such models are incapable of producing strong masking with common onset of the target and mask. If the models were truly constructed this way, Di Lollo *et al.* would be correct in claiming that the models could not account for their data. However, the models are not based solely on inhibition from transient signals, and, as Figure 1 shows, the models can generate substantial masking with common-onset. Moreover, even the most prominent theory that utilizes transient inhibition (Breitmeyer & Ganz, 1976; Breitmeyer, 1984) also hypothesizes additional mechanisms. Breitmeyer and Ganz (1976) proposed that Type A masking effects (where the strongest masking occurs for common onset of the target and mask) are due to interactions between sustained signals of the target and mask. A quantitative version of the full Breitmeyer and Ganz theory has never been developed, but it seems likely that the theory could account for properties of common-onset masking.

Di Lollo *et al.* (2000) argued that feed-forward theories of information processing could not account for their experimental data (most notably the properties of common-onset masking) and that models with reentrant (feed-back) processing were necessary. The data do not support such strong claims. In fact, two key model characteristics account for common-onset masking in all the models (including the model proposed by Di Lollo *et al.*). First, all of these models hypothesize that the target stimulus engenders some type of persisting trace in the visual system. The strength of the target percept (which leads to reports of visibility or percent detection) is based on the magnitude of this persisting trace. For the Di Lollo *et al.* model, the percept strength is related to the magnitude of this trace at the moment attention is focused on the target. For the other models, the percept strength is related to an integral of the persisting trace over time. Second, all the models hypothesize that a signal corresponding to the mask weakly interacts with the persisting trace to make it smaller and reduce the strength of the target percept. The masking effect plays out over time as the target's persisting trace fades away. The mask-to-target interaction at any given point in time must be weak to allow variation in the mask duration to produce measurable effects. If the mask-to-target interaction was so strong as to completely erase the target's persisting trace at the mask's onset, then increases in mask duration would never lead to stronger masking (there would be a floor effect).

These two model properties are a subset of model characteristics that were identified by Francis (2000) as necessary for a class of models (this class includes all the models discussed here) to account for u-shaped backward masking functions, which are generated by fixing target and mask durations and varying SOA. Thus, while common-onset masking may provide constraints on model parameters, it offers fewer constraints on fundamental model hypotheses than other types of masking.

Although reentrant processing can produce necessary model characteristics to account for the data, reentrant processing is not itself a necessary characteristic. Indeed, the model of Weisstein (1972) is strictly feed-forward, yet it accounts for the experimental data of Di Lollo *et al.* fairly well. The models of Bridgeman (1971, 1978) and Francis (1997) include feed-back (reentrant processing), and here the feed-back provides a mechanism for either producing the target's persisting trace or insuring that the mask-to-target interaction is weak (Francis, 2000).

Di Lollo *et al.* (2000) also suggested that current models could not account for their experimental results because the current models are contour based, and there are not enough contours in the sparse masks to produce strong masking. There is a legitimate question about whether masking is contour based or is object based (as is suggested by Di Lollo *et al.*). It is true that the current models of masking assume that the mask inhibits the target through some interaction of contours. It is also true that these models all suggest that with fewer or weaker contours, the strength of masking should weaken. However, it does not logically follow that a small amount of contour cannot exhibit *any* masking. For example, if the experimental task is particularly challenging, even weak inhibition from the mask to the target can give rise to substantial decrements in task performance. Moreover, it is not clear that four dots, as used by Di Lollo *et al.* (2000), is an appropriate definition of a "small" amount of contour. It may be that any visible stimulus has sufficient contour to produce masking in the context of distributed attention.

Conclusions

Contrary to the claims of Di Lollo *et al.* (2000), we have shown that most current quantitative models of backward masking can account for properties of common-onset masking and for masking with sparse masks. We have also shown that a fairly simple assumption regarding the role of attention can account for the effects of target set size reported by Di Lollo *et al.* (2000).

This is not a minor point of contention. On the basis of their experimental findings, Enns and Di Lollo (1997) and Di Lollo *et al.* (2000) argued that they had found a new form of masking and that this new form of masking required a new explanation. These conclusions have been repeated in additional studies (e.g., Jiang & Chun, 2001a,b; Enns, in press) that attempt to investigate the new type of masking and the new theory. It is important to properly characterize how the new data and new theory fit into current theories. Although current theories of masking account for the data very nicely, there is certainly nothing wrong with proposing a new explanation of masking. The theory put forth by Di Lollo *et al.* (2000) has several intriguing characteristics, and we hope it is developed further. However, the existing data do not allow us to clearly prefer this new model over other models that have been previously analyzed and applied to other data sets. Moreover, since existing models can account for both old and new forms of masking, it raises doubts about whether Enns and Di Lollo (1997) have found a truly new form of masking.

We do agree with Di Lollo *et al.* (2000) about the importance of integrating the role of spatial attention into theories of masking. We look forward to additional data on the role of spatial attention in masking and hope that such data will either guide the development of current models, or allow us to reject them.

References

Anbar, S. & Anbar, D. (1982). Visual masking: A unified approach. *Perception*, *11*, 427–439.

Bischof, W. F. & Di Lollo, V. (1995). Motion and metacontrast with simultaneous onset of stimuli. *Journal of the Optical Society of America A*, *12*, 1623–1636.

Breitmeyer, B. (1978). Metacontrast masking as a function of mask energy. *Bulletin of the Psychonomic Society*, *12*, 50–52.

Breitmeyer, B. (1984). *Visual masking: An integrative approach.* New York: Oxford University Press.

Breitmeyer, B. & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, *83*, 1–36.

Breitmeyer, B. & Ögmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*, 1572–1595.

Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychological Review*, *78*, 528–539.

Bridgeman, B. (1978). Distributed sensory coding applied to simulations of iconic storage and metacontrast. *Bulletin of Mathematical Biology*, *40*, 605–623.

Enns, J. T. (in press). Visual binding in the standing wave illusion. *Psychonomic Bulletin & Review.*

Enns, J. T. & Di Lollo, V. (1997). Object substitution: a new form of masking in unattended visual locations. *Psychological Science*, *8*, 135–139.

Enns, J. T. & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences*, *4*, 345–352.

Di Lollo, V., Bischof, W. F., & Dixon, P. (1993). Stimulus-onset asynchrony is not necessary for motion perception or metacontrast masking. *Psychological Science*, *4*, 260–263.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General*, *129*, 481–507.

Francis, G. (1997). Cortical dynamics of lateral inhibition: Metacontrast masking. *Psychological Review*, *104*, 572–594.

Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review*, *107*, 768–785.

Gilden, D., MacDonald, K., & Lasaga, M. (1988). Masking with minimal contours: Selective inhibition with low spatial frequencies. *Perception & Psychophysics*, *44*, 127–132.

Jiang, Y., & Chun, M. M. (2001a). Asymmetric object substitution masking. *Journal of Experimental Psychology : Human Perception and Performance*, *27*, 895-918.

Jiang, Y., & Chun, M. M. (2001b). The spatial gradient of visual masking by object substitution. *Vision Research*, *41*, 3121-3131.

Sherrick, M. & Dember, W. (1970). Completeness and spatial distribution of mask contours as factors in backward visual masking. *Journal of Experimental Psychology*, *84*, 179–180.

Spencer, T. J. & Shuntich, R. (1970). Evidence for an interruption theory of backward masking. *Journal of Experimental Psychology*, *85*, 198–203.

Weisstein, N. (1966). Backward masking and models of perceptual processing. *Journal of Experimental Psychology*, *72*, 232–240.

Weisstein, N. (1968). A Rashevsky-Landahl neural net: Simulation of metacontrast. *Psychological Review*, *75*, 494–521.

Weisstein, N. (1972). Metacontrast. In D. Jameson & L. Hurvich (Eds.) *Handbook of sensory physiology* (Vol. 7, No. 4, *Visual psychophysics*). Berlin: Springer-Verlag.

Werner, H. (1935). Studies on contour: I. Qualitative analysis. *American Journal of Psychology*, *47*, 40–64.

Received October 16, 2001
Revision received March 7, 2002
Accepted March 14, 2002