

The Stroop effect and mental imagery

Bence Nanay 
University of Antwerp, Belgium

Perception

1–7

© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/03010066231212152

journals.sagepub.com/home/pec

Abstract

The classic Stroop task is very simple: you have to name the color of words printed on a page. If these words are color words (like “red” or “blue”), where the color named and the color it is printed in are different (say, “red” printed in blue), the reaction time increases significantly. My aim is to argue that the existing psychological explanations of the Stroop effect need to be supplemented. The Stroop effect is not exclusively about access to motor control. It is also, to a large extent, about interferences in perceptual processing. To put it briefly, reading the color word triggers—laterally and automatically—visual imagery of the color and this interferes with the processing of the perceived color of the word. In other words, the Stroop effect is to a large extent a sensory phenomenon, and it has less to do with attention, conflict monitoring, or other higher-level phenomena.

Keywords

imagery, Stroop effect, attention, conflict monitoring

Date Received: 30 January 2023; accepted: 13 October 2023

Introduction

One of the most widely researched psychological phenomena of all times is the Stroop effect (see Stroop, 1935). The classic Stroop task is very simple: you have to name the color of words printed on a page. If these words are color words (like “red” or “blue”), where the color named and the color it is printed in are different (say, “red” printed in blue), the reaction time increases significantly.

What explains this odd difference? There are two major explanations, the first one dominant in the second half of the 20th century, the second dominant in the last 20 years. According the first one, the Stroop effect is about attention capture. The linguistic stimulus captures our attention, and as a consequence, less attention remains for the processing of the color stimulus (see

Corresponding author:

Bence Nanay, Centre for Philosophical Psychology, University of Antwerp, 15 Prinsstraat, Antwerpen, 2000, Belgium.

Email: bence.nanay@uantwerpen.be

MacLeod, 1991 for a summary). According to the second one, the Stroop effect is about conflict monitoring and control: there are control mechanisms that detect the conflict between the linguistic and the color stimulus and they prioritize the processing of the language stimulus (Botvinick et al., 2001).¹

The attention account and the conflict monitoring account of the Stroop effect are very different inasmuch as the former gives a fully bottom-up explanation, whereas the latter a top-down one of the effect the semantic meaning of the word has on the processing of color. But they share an important premise, namely, that the Stroop effect is about access to motor control. Depending on whether the word “red” is printed in red or blue, our access to the motor control (of reading the word) is different and this explains the difference in our reaction time. This is clear enough in the attention account, but it is also what is behind the conflict monitoring account, where “conflict may be operationally defined as the simultaneous activation of incompatible representations [...] e.g., representations of alternative responses” (Botvinick et al., 2001, p. 630).

My aim in this paper is to argue that the Stroop effect is not exclusively about access to motor control. It is also, to a large extent, about interferences in perceptual processing. To put it briefly, reading the color word triggers—laterally and automatically—visual imagery of the color and this interferes with the processing of the perceived color of the word.

In section “Mental Imagery”, I outline the concept of mental imagery that is relevant in this discussion and in section “Language Processing and Mental Imagery”, I provide empirical evidence for the various ways in which language processing and mental imagery interact. In section “Back to the Stroop Effect”, I argue that these interactions provide a clear case for a very early perceptual interference from language processing to perceptual processing that explains some aspects of the Stroop effect in a much more straightforward manner than either the attention account or the conflict monitoring account could.

Mental Imagery

The term “mental imagery” was first consistently used in the early days of experimental psychology in the second half of the 19th century and while it has clearly made it to our ordinary language, the way psychologists and neuroscientists use the concept is not as an ordinary language category. Here is a representative definition from a review article on mental imagery in the journal *Trends in Cognitive Sciences*: “We use the term ‘mental imagery’ to refer to representations [...] of sensory information without a direct external stimulus” (Pearson et al., 2015, p. 590; see also Nanay, 2015, 2018).

This definition captures the pre-theoretical notion of mental imagery, which we tend to have in mind when, for example, thinking about the experience of closing our eyes and visualizing an apple. That experience is a representation of sensory information without direct external stimulus. But the concept of mental imagery has a much wider scope than just the experience of visualizing.

First, mental imagery, like perception, can happen in all sense modalities. Mental imagery can be visual, but it can also be auditory, olfactory, gustatory, and tactile. Second, while visualizing an apple amounts to a voluntary use of mental imagery, there is also involuntary mental imagery, like flashbacks or earworms—annoying tunes that go through our head in spite of the fact that we really don’t want them to. Third, while in the case of visualizing, mental imagery is not accompanied by the feeling of presence—you’re not actually taking the apple to be in front of you—, some other forms of mental imagery may be accompanied by the feeling of presence, for example, in the case of lucid dreaming and in some forms of hallucinations (which are widely taken to be forms of mental imagery in psychiatry).

The definition I have been using is a negative definition. It defines mental imagery as (to rephrase a bit) sensory representation not triggered directly by sensory input. But it leaves open the question

about what this sensory representation is triggered by (directly). In some cases, it is triggered by top-down processes, as in the case of closing your eyes and visualizing an apple. But in other cases, it is triggered laterally, by, for example, input in another sense modality. When you watch the TV muted, for example, your auditory representation (and often your salient auditory experience) is not directly triggered by the auditory input—there is no auditory input as the TV is muted. It is directly triggered by the visual input of the images on TV (Calvert et al., 1997; Hertrich et al., 2011; Nanay, 2018; Pekkola et al., 2005; Spence & Deroy, 2013).

It should be clear that while the definition of mental imagery I have been using does seem to capture the ordinary usage of the term, it also carves up mental phenomena somewhat differently. As we have seen, it allows for involuntary imagery. But it also allows for unconscious mental imagery as nothing in the definition says that the perceptual representation that is not triggered directly by sensory input must be a conscious representation.

We have an overwhelming amount of evidence that perception may be conscious or unconscious (e.g., Kouider & Dehaene, 2007). But if perceptual representations that are directly triggered by sensory input (i.e., perception) may be unconscious, then it would be arbitrary to posit that perceptual representations that are not directly triggered by sensory input (i.e., mental imagery) may not be. Further, some people report having no conscious mental imagery—these people are called aphantasics and in the last two decades or so a lot of experimental studies were conducted to find out about the causes and nature of aphantasia (see, e.g., Zeman et al., 2007). And while aphantasia seems to be a non-monolithic phenomenon, where many different things can lead to the lack of conscious mental imagery, there is clear evidence that at least a subset of aphantasics, while reporting to have no conscious mental imagery at all, do have mental imagery in the sense of perceptual representation that is not directly triggered by sensory input. They have unconscious mental imagery (Nanay, 2021).

In short, mental imagery may be voluntary or involuntary and it may be conscious or unconscious. It is a scientifically respectable (and even publicly observable) category that is well suited to play a role in the explanation of psychological phenomena.

Language Processing and Mental Imagery

We now know that language processing is not completely detachable from mental imagery. Both generating linguistic utterances and hearing/reading them utilizes mental imagery. Some of the empirical findings supporting these claims come from neuroimaging. Describing a scene relies on our ability to generate mental imagery—early cortical representations not directly triggered by sensory input (Mar, 2004; Zadbood et al., 2017). Even more importantly, hearing a description invariably triggers mental imagery—again, not necessarily conscious mental imagery, but early cortical representations not directly triggered by sensory input and it is this imagistic representation that is remembered, not the words we heard (Zwaan, 2016; Zwaan & Radvansky, 1998).

We understand a fair amount of how this happens and, crucially, we know a lot about the ways in which linguistic labels change (and speed up) perceptual processes and we also know a fair amount about the time scale of this influence. The most important piece of finding both from EEG and from eye tracking studies is that linguistic labels influence shape recognition in less than 100 ms (Boutonnet & Lupyan, 2015; de Groot et al., 2016; Noorman et al., 2018—it should be acknowledged that in these experiments, the onset of the linguistic label preceded the onset of the shape to be recognized). This is a very similar time-frame as how long it takes for the stimulus to reach V4 (Zamarashkina et al., 2020)—that is, extremely fast (note that word recognition does take significantly longer, see Hauk et al., 2012).

Crucially, this less than 100 ms it takes for linguistic labels to influence shape recognition is much shorter than the time that would be needed for perceptual processing to reach all the way

up to higher level representations and then trickle all the way down again to the primary visual cortex (see Lamme & Roelfsema, 2000; Thorpe et al., 1996 for the temporal unfolding of visual processing in unimodal cases and see Kringelbach et al., 2015 for a summary of the relative slowness of non-early cortical processing).

By means of comparison, amodal completion (the visual representation of occluded parts of perceived objects) is taken to be bottom-up or laterally influenced on the basis of timing studies although it happens slightly slower than 100 ms. Amodal completion in the early cortices happens within 100–200 ms of retinal stimulation (Rauschenberger et al., 2001; Sekuler & Palmer, 1992—this is true even of complex visual stimuli, like faces, see Chen et al., 2009; see also Lerner et al., 2004; Rauschenberger et al. 2006; Yun et al., 2018 for detailed studies that track the (very quick) temporal unfolding of amodal completion in different parts of the visual cortex). If the 100–200 ms of amodal completion is explained in terms of lateral influence, then the less than 100 ms of the influence of linguistic labeling can also be explained in terms of lateral influence.

This means that linguistic processing and mental imagery interact at an extremely early stage of perceptual processing—by any account in early cortical processing.

Back to the Stroop Effect

My aim is to argue that in the light of these results about the relation between language processing and mental imagery, we have good reasons to hold that reading the color word triggers—laterally and automatically—visual imagery of the color and this interferes with the processing of the perceived color of the word and this is what explains the Stroop effect. In other words, the conflict between the color and the meaning of the word starts much earlier than motor control.

Here is an experiment that supports this hypothesis directly (there may be some indirect support from findings about the Stroop effect for color-related words as well (like “sky” [for blue] and “fire” [for red]—see Dairyple-Alford, 1972). A recent experiment shows that even if we control for all the attentional and other mechanisms that determine motor control, the activation patterns in V4—the part of the visual cortex that is responsible for color processing—would be difficult to explain unless we posit early sensory involvement in the Stroop effect (Purmann & Pollmann, 2015).

Given that V4 is devoted (mainly) to color processing, it is active throughout any color Stroop task. More generally, the involvement of V4 in the Stroop task is somewhat difficult to examine experimentally given that without the functioning of these regions, the effect goes away. So some tricks are required to gain any insight about exactly how early cortical color processing is involved in the Stroop task. The experiments in Purmann & Pollmann, 2015 examined the ways in which the previous trial in a series of Stroop tasks influences the current trial. So the question they raised is how your early sensory cortices behave depending on the order of these trials. If you read the word “red” printed in blue, there is a conflict—it is an “incongruent trial.” If you read the word “blue” printed in blue, there is no conflict—it is referred to as a “congruent trial.”

The question is whether early sensory processing is different depending on whether an incongruent trial was preceded by another incongruent trial. And what the results show is that activities in V4 are very different depending on whether the previous trial was congruent or incongruent. Interestingly, the same effect was not observed in language processing regions of the brain, only in V4. If we take the Stroop task to be about motor control, these results make little sense. But if, as I am suggesting, it is at least partly about sensory processing, these results are exactly what we should expect.

The color of the word activates V4 bottom up (that’s perception). And the reading of the word activates V4 laterally and automatically (that’s mental imagery). And the processing of the

perceived color is slowed down because of the interference of the mental imagery. In short, the conflict between the color and the meaning of the word starts already in perceptual processing.

A word of caution about the scope of the claim I argued for in this paper. While the main findings of the Stroop effect can be explained in terms of the lateral and automatic activation of mental imagery, I don't want to pretend that all aspects of the Stroop effect can be explained with the help of this explanatory scheme. For example, we know that subjects show greater interference on the first few trials in each block of testing than on subsequent trials in the series (Henik et al., 1997). Also, there is less interference on incongruent trials if they are frequent in comparison with congruent trials (Lindsey & Jacoby, 1994). I don't think the appeal to the laterally and automatically triggered mental imagery will help us explain these findings.

Nonetheless, we can conclude that the Stroop effect is, at least partially, a sensory phenomenon, and it has less to do with attention, conflict monitoring, or other higher-level phenomena than previously supposed. While it may give us insights into the nature of attention and automaticity or into the intricacies of conflict monitoring and cognitive control, its theoretical import may be even more significant. In fact, the way language processing and perceptual processing interact in the case of the Stroop effect can open up new research directions both about early cortical sensory processing and about language processing, besides touching on some of the deepest (and earliest) philosophical questions about the relation between perception and language.

Author Contribution(s)

Bence Nanay: Conceptualization; Formal analysis; Investigation; Project administration; Writing – original draft; Writing – review & editing.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Bence Nanay  <https://orcid.org/0000-0002-2835-6530>

Note

1. The Stroop effect is an important phenomenon in its own right, but it has also been used in many branches and paradigms of psychological research in order to determine the automaticity of our various psychological responses. Just one example: it is widely held that a necessary feature of synesthetic experiences is that they show the Stroop effect. When deciding whether more unusual forms of synesthesia (e.g., swimming-style synesthesia, where different styles of swimming—backstroke, breaststroke, etc.—bring about synesthetic color experiences, see Nikolić et al., 2011; Rothen et al., 2013) really counts as genuine synesthesia, the hallmark is to show that they show the Stroop effect.

References

- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652. <https://doi.org/10.1037/0033-295X.108.3.624>
- Boutonnet, B., & Lupyan, G. (2015). Words jump-start vision: A label advantage in object recognition. *Journal of Neuroscience*, *35*, 9329–9335. <https://doi.org/10.1523/JNEUROSCI.5111-14.2015>

- Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., Woodruff, P. W. R., Iversen, S. D., David, A. S., & Williams, S. C. R. (1997). Activation of auditory cortex during silent lipreading. *Science*, *276*, 593–596. <https://doi.org/10.1126/science.276.5312.593>
- Chen, J., Liu, B., Chen, B., & Fang, F. (2009). Time course of amodal completion in face perception. *Vision Research*, *49*, 752–758. <https://doi.org/10.1016/j.visres.2009.02.005>
- Dairymply-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception and Psychophysics*, *11*, 274–276. <https://doi.org/10.3758/BF03210377>
- de Groot, F., Huettig, F., & Olivers, C. N. L. (2016). When meaning matters: The temporal dynamics of semantic influences on visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 180–196. <https://doi.org/10.1037/xhp0000102>
- Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word Reading: Evidence from fast behavioral and brain responses. *NeuroImage*, *60*, 1462–1477. <https://doi.org/10.1016/j.neuroimage.2012.01.061>
- Henik, A., Bibi, U., Yanai, M., & Tzelgov, J. (1997). The stroop effect is largest during first trials. *Abstracts of the Psychonomic Society*, *2*, 57.
- Hertrich, I., Dietrich, S., & Ackermann, H. (2011). Cross-modal interactions during perception of audiovisual speech and nonspeech signals: An fMRI study. *Journal of Cognitive Neuroscience*, *23*, 221–237. <https://doi.org/10.1162/jocn.2010.21421>
- Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 857–875. <https://doi.org/10.1098/rstb.2007.2093>
- Kringelbach, M. L., McIntosh, A. R., Ritter, P., Jirsa, V. K., & Deco, G. (2015). The rediscovery of slowness: Exploring the timing of cognition. *Trends in Cognitive Sciences*, *19*, 616–628. <https://doi.org/10.1016/j.tics.2015.07.011>
- Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*, 571–579. [https://doi.org/10.1016/S0166-2236\(00\)01657-X](https://doi.org/10.1016/S0166-2236(00)01657-X)
- Lerner, Y., Harel, M., & Malach, R. (2004). Rapid completion effects in human high-order visual areas. *NeuroImage*, *21*, 516–526. <https://doi.org/10.1016/j.neuroimage.2003.08.046>
- Lindsey, D. S., & Jacoby, L. I. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 219–234. <https://doi.org/10.1037/0096-1523.20.2.219>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- Mar, R. A. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia*, *42*, 1414–1434. <https://doi.org/10.1016/j.neuropsychologia.2003.12.016>
- Nanay, B. (2015). Perceptual content and the content of mental imagery. *Philosophical Studies*, *172*, 1723–1736. <https://doi.org/10.1007/s11098-014-0392-y>
- Nanay, B. (2018). Multimodal mental imagery. *Cortex*, *105*, 125–134. <https://doi.org/10.1016/j.cortex.2017.07.006>
- Nanay, B. (2021). Unconscious mental imagery. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, 20190689. <https://doi.org/10.1098/rstb.2019.0689>
- Nikolić, D., Jürgens, U., Rothen, N., Meier, B., & Mroczko, A. (2011). Swimming-style synesthesia. *Cortex*, *47*, 874–879. <https://doi.org/10.1016/j.cortex.2011.02.008>
- Noorman, S., Neville, D. A., & Simanova, I. (2018). Words affect visual perception by activating object shape representations. *Scientific Reports*, *8*, 14156. <https://doi.org/10.1038/s41598-018-32483-2>
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015). Mental imagery: Functional mechanisms and clinical applications. *Trends in Cognitive Sciences*, *19*, 590–602. <https://doi.org/10.1016/j.tics.2015.08.003>
- Pekkola, J., Ojanen, V., Autti, T., Jaaskelainen, I. P., Mottonen, R., Tarkainen, A., & Sams, M. (2005). Primary auditory cortex activation by visual speech: An fMRI study at 3 T. *NeuroReport*, *16*, 125–128. <https://doi.org/10.1097/00001756-200502080-00010>

- Purmann, S., & Pollmann, S. (2015). Adaptation to recent conflict in the classical color-word stroop-task mainly involves facilitation of processing of task-relevant information. *Frontiers in Human Neuroscience*, *9*, 88. <https://doi.org/10.3389/fnhum.2015.00088>
- Rauschenberger, R., Liu, T., Slotnick, S. D., & Yantis, S. (2006). Temporally unfolding neural representation of pictorial occlusion. *Psychological Science*, *17*, 358–364.
- Rauschenberger, R., & Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature*, *410*, 369–372. <https://doi.org/10.1038/35066577>
- Rothen, N., Nikolic, D., Jurgens, U. M., Mroczko-Wasowicz, A., Cock, J., & Meier, B. (2013). Psychophysiological evidence for the genuineness of swimming-style colour synaesthesia. *Consciousness and Cognition*, *22*, 35–46. <https://doi.org/10.1016/j.concog.2012.11.005>
- Sekuler, A. B., & Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General*, *121*, 95–111. <https://doi.org/10.1037/0096-3445.121.1.95>
- Spence, C., & Deroy, O. (2013). Crossmodal imagery. In S. Lacey & R. Lawson (Eds.), *Multisensory imagery* (pp. 157–183). Springer.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. <https://doi.org/10.1037/h0054651>
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522. <https://doi.org/10.1038/381520a0>
- Yun, Y., Hazenberg, S. J., & van Lier, R. (2018). Temporal properties of amodal completion: Influences of knowledge. *Vision Research*, *145*, 21–30. <https://doi.org/10.1016/j.visres.2018.02.011>
- Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How we transmit memories to other brains: Constructing shared neural representations via communication. *Cerebral Cortex*, *27*, 4988–5000. <https://doi.org/10.1093/cercor/bhx202>
- Zamarashkina, P., Popovkina, D. V., & Pasupathy, A. (2020). Timing of response onset and offset in macaque V4: Stimulus and task dependence. *Journal of Neurophysiology*, *123*, 2311–2325.
- Zeman, A., McGonigle, D., Gountouna, E., Torrens, L., Della Sala, S., & Logie, R. (2007). Blind imagination: Brain activation after loss of the mind's eye. *Journal of Neurology Neurosurgery & Psychiatry*, *78*, 209–209.
- Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*, 1028–1034. <https://doi.org/10.3758/s13423-015-0864-x>
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185. <https://doi.org/10.1037/0033-2909.123.2.162>