Commentary

Is the Musical Stroop Effect Able to Keep Its Promises?

A Reply to Akiva-Kabiri and Henik (2014), Gast (2014), Moeller and Frings (2014), and Zakay (2014)

Laurent Grégoire, Pierre Perruchet, and Bénédicte Poulin-Charronnat

CNRS-UMR 5022, University of Bourgogne, Dijon, France

Abstract. Grégoire, Perruchet, and Poulin-Charronnat (2013) claimed that the Musical Stroop task, which reveals the automaticity of note naming in musician experts, provides a new tool for studying the development of automatisms through extensive training in natural settings. Many of the criticisms presented in the four commentaries published in this issue appear to be based on a misunderstanding of our procedure, or questionable postulates. We maintain that the Musical Stroop Effect offers promising possibilities for further research on automaticity, with the main proviso that the current procedure makes it difficult to tease apart facilitation and interference.

Keywords: automatism, Stroop effect, interference, musical expertise, note naming

The basic arrangement of the experiments reported in Grégoire, Perruchet, and Poulin-Charronnat (2013) comprised a staff with a note in various positions. A name of a note was printed inside the note. In the congruent condition, the note name was congruent with the note position, whereas in the incongruent condition, name and position were incongruent. The main result was that musicians asked to read aloud the written name of the notes showed impaired processing in the incongruent condition with regard to the congruent condition, a result coined as the *Musical Stroop Effect* (MSE). In this reply, we focus on the points that have been presented as potential limitations of the procedure in the four previous commentaries. We thank the commentators for other helpful and constructive comments, which space limitation does not allow discussing.

The Musical Stroop Task Is Just Another Stroop-Like Test

Zakay (2014) writes that the musical Stroop task "is no more than another Stroop-like test." The material is indeed similar to a standard picture-word interference task, as repeatedly noted by Moeller and Frings (2014), but, crucially, the MSE is the *reverse* of the effect explored in most Stroop-like tests. Reading is involved, but as the object, rather than the source, of interference. The procedure was designed to investigate the automaticity of note naming in musicians, in the same way as the classical Stroop task investigates the automaticity of word reading. We agree with Zakay (2014) that the musical Stroop task does not replace the standard color-word version, but this was not our objective. More modestly, we intended to provide a tool that would be better suited than the other Stroop tasks to address a specific question in future research: How does Stroop interference evolve with practice? The key point is that the emergence of automaticity in musicians provides the opportunity of better control on the conditions of training than reading acquisition, notably because the level of musical expertise can be easily decoupled from age and academic level. As Moeller and Frings (2014) rightly point out, there are several conflicting hypotheses about how interference may evolve with increasing expertise, and exploring this evolution looks as promising.

There Is No Note-Picture Note-Name Association

Zakay (2014), and to a lesser extent Moeller and Frings (2014), contend that the note name would not be activated

Experimental Psychology 2014; Vol. 61(1):80–83
DOI: 10.1027/1618-3169/a000222

by a note picture, thus negating the very existence of the automatism we were tracking in our experiments. A major problem with this contention is that, if right, it would leave unexplained why an MSE occurred in our experiments.

Moeller and Frings (2014) invoke as a definitive argument for justifying that the name of the note cannot be activated: ''Just think of a trumpeter!'' It must be understood that assuming the automaticity of note naming in musicians does not amount to claim that a note picture generates an irrepressible need to name aloud the note, which would be obviously incompatible with the practice of a wind instrument, among other activities. If an overt response were required, then the classical Stroop effect would not exist in the first place, given that reading aloud is certainly infrequent in adults. As suggested by Gast (2014), it is even possible that the note name does not really evoke a *response*, even subvocal. The MSE would come from the learned association of note and note name, conceived as a stimulus-stimulus relationship, which would generate interference during the encoding phase when the note/note name contingency is broken. This is a sensible hypothesis, which warrants further investigations.

Of course, we are not asserting that the mastery of note naming is required for any form of musical performance. As Akiva-Kabiri and Henik (2014) observe, note labeling may not be needed for playing by ear, or still playing from memory or improvising. Our contention is, however, that note naming is a key component of music reading. As Hodges and Nolker (2011) wrote: ''Although there are many oral musical traditions and practices, music reading hold a special place in contemporary music education curricula. [...] Virtually all beginning instrumental method books and private instructional books [...] have sections on music reading, as do general music basal series'' (p. 61).

Trading Verbal Responses Against Manual Responses Would Be Better

Starting from the postulate that musical training is primarily directed toward the automatization of the motor programs involved in the practice of a musical instrument, several commentators suggest requiring the production of the note on a musical instrument, as a complement (Gast, 2014) or as a better alternative (Moeller & Frings, 2014; Zakay, 2014) to verbal responses. Although appealing at first glance, this suggestion overlooks the fact that the stimulus to which participants are asked to respond in the musical Stroop task is *not* a note on a staff, but a written note name. To illustrate what happens in these conditions, let us consider Zakay and Glicksohn's (1985) study. Among the many conditions of their experiment with a sample of pianists, the authors introduced verbal (reading aloud) and motor (pressing the appropriate key on the piano) modes of responding to written note names. It turned out that motor responses were, on average, considerably longer than verbal responses (650 ms vs. 430 ms, respectively). The reason is straightforward: A written note name does not provide sufficient information to trigger a unique motor response. In particular, a note name does not specify the octave, hence obliging pianists to an arbitrary choice between seven or eight possible piano keys for each note. It looks unlikely that response production processes that imply time-consuming intentional decisions could be impacted by motor automatisms.

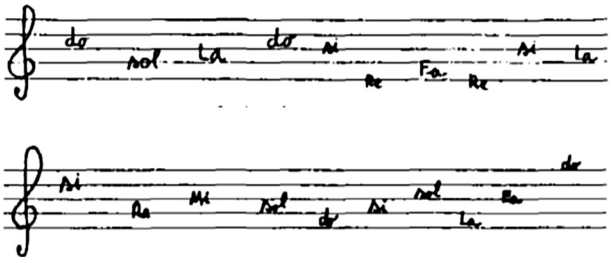

Figure 1. Stimuli used by Zakay and Glicksohn (1985). In the conditions of interest for our concern, pianists had to press the appropriate piano keys corresponding to the written names of the notes, irrespective of their location. Note names and note locations were either congruent (upper staff) or incongruent (lower staff).

In apparent contradiction with our analysis, however, the data reported with motor responses in Zakay and Glicksohn's Table 1 showed a difference between incongruent and congruent items, which even largely exceeded in amplitude the standard Stroop effect (287 ms). Although the authors made neither descriptive nor inferential analyses of this effect, Zakay (2014) retrospectively describes it as an earlier demonstration of an MSE with a motor response. This claim is questionable. An alternative, and much more plausible explanation for the 287 ms difference appears when looking at the materials used by Zakay and Glicksohn, which is partially reproduced in Figure 1 (see also our discussion of this study in Grégoire et al., 2013). The crucial point is that the time-consuming selection of a specific key is no longer necessary in the congruent condition because, by contrast with the incongruent condition, the note location on the staff now designates a correct and unique key for the response. May be another procedure would be successful in revealing an MSE with motor responses, but the design remains to be invented. Our current feeling is that the lack of a one-to-one mapping between a note name and a motor response (whatever the musical instrument) raises an insurmountable obstacle to simply trade verbal mode against motor mode of responding in the musical Stroop task.

The Two Competing Dimensions Are Integrated (or Separated)

According to Moeller and Frings (2014), the note picture and the written note name that is written inside belong to the same object. Moeller and Frings (2014) worry: ''Thus, it cannot be ruled out that the interference of note naming

on note name reading would diminish if written note name and note would not belong to the same object. This seems to be the case for other variants of the Stroop effect and would argue against a strong claim of automatic processing.'' Zakay (2014) describes our arrangement in a diametrically opposite way on the separated/integrated dimension, thus asserting that our dimensions are separated. The point is that, ironically, even though the diagnoses of Moeller and Frings (2014) and Zakay (2014) are opposite, their conclusions are identical: Our design is defective.

These criticisms are puzzling because, to our knowledge, no one has ever questioned the validity of the conclusions stemming from a Stroop study on the ground that the level of integration of the two competing dimensions was too strong (Moeller & Frings, 2014) or too weak (Zakay, 2014). It has been well documented that integrated dimensions produce the largest amount of Stroop interference, and that interference decreases when the competing stimuli move farther one from each other (MacLeod, 1991). Therefore, it is highly predictable that if the picture shown in the right panel of Moeller and Frings's (2014) Figure 1 were used instead of the currently used picture, the MSE would be reduced. This indeed could be put forth against a strong claim about the automaticity of note naming, as commonly acknowledged for reading since the 1980's, but we fail to see how this would impact our own work. To be sure, we have never claimed that the automaticity of note naming in musicians would be stronger than the automaticity of reading.

Teasing Apart Facilitation and Interference

Akiva-Kabiri and Henik (2014) point out the difficulty to tease apart interfering and facilitating components of the MSE. The question is: Where a note name should be printed to give a baseline reading time to which performance in congruent and incongruent conditions could be compared? A seemingly obvious response is ''out of the staff,'' but our results revealed that reading a printed word out of the staff is much shorter than reading the same word inside a staff, even for nonmusicians, presumably due to the perceptual complexity induced by the staff.

Akiva-Kabiri and Henik (2014) recommend printing the note name on an inverted staff. This would indeed control for some physical components of perceptual complexity, and we are relatively confident that this condition would equalize the reading times between the neutral, congruent, and incongruent conditions for nonmusicians. We are more skeptical about the neutrality of an inverted staff for musicians, because we suspect that musicians may either neglect the reversal (hence coding the note as in the congruent and incongruent trials) or take the reversal into account, hence processing the stimuli as if the whole score was displayed upside down.

As a result, we agree that the difficulty of dissociating facilitation and interference in the MSE is an important limitation of the paradigm, and we acknowledge that our almost exclusive focus on the interference component is a questionable shortcut. This being so, in the same way as we see no reason to deny that some facilitation could account for a part of the MSE as in other Stroop paradigms, we see no reason to believe that this part would be larger than in other paradigms. MacLeod (1991) concludes from his review on this issue that ''facilitation [in the congruent condition] is much less than the corresponding interference in the incongruent condition'' (p. 175). He adds: ''and the choice of control condition may be crucial.'' The last sentence usefully recalls that even if the choice of a control condition may be especially problematic for the MSE, it would be wrong to believe that other Stroop paradigms are totally free from similar intricacies. Some investigators have even questioned the use of a baseline to measure facilitation and interference effects in the standard color-word task (e.g., Lindsay & Jacoby, 1994).

Does Music Practice Offer the Best Opportunity to Track Automatism Formation?

As mentioned above, the main objective of replacing reading by note naming in the Stroop task is the possibility of better control on the conditions of training. Several commentators note various problems that could arise when trying to control the level of practice in music, nevertheless. For instance, Akiva-Kabiri and Henik (2014) note the variability of the musical notation systems across cultures and instruments. Along the same line, Gast (2014) points out that the learning of musical notation is not a linear and incremental process, which may complicate the interpretation of learning curves across years of musical school. These are useful caveats. Using the number of years of practice as the single criterion for selecting participants is certainly insufficient, and further studies should refine the selection criteria. Measuring the speed and accuracy of note naming with a standard test could be a useful complement.

Akiva-Kabiri and Henik (2014) suggest that investigating the automaticity of reading during second language acquisition could provide a better control over training than the MSE. Our feeling is that the practical feasibility of one or another paradigm may depend on conditions specific to each country. With regard to the current landscape in France, the acquisition of second language is hard to decouple from age and academic level, whereas the organization of musical teaching ensures nice conditions of investigation. Indeed, musical teaching is mainly provided by music schools that recruit people of any age, and which place great emphasis on music reading. In addition, Grégoire et al. (2013) noted that many conceptual issues on bilingualism are currently unsolved, and that the complexity of the involved processes could be damaging for drawing clear conclusions on automatisms. That being said, we fully share the view that further research following this approach would be worthwhile.

Gast (2014) usefully recalls that creating an automatism does not necessarily require years of practice, and that some Stroop-like effects can be observed after a quite limited amount of training in laboratory conditions. As Gast (2014) notes, this approach has the obvious advantage of ensuring the best experimental control. It remains to be seen whether the very same phenomena are observed at a micro-level in a one-session laboratory task and at a macro-level after years of consistent practice. Some Stroop-like effects have been observed with arbitrary mapping in laboratory, as Gast (2014) mentions, but to our knowledge, newly acquired automatisms remain unable to interfere with reading (MacLeod, 1998), thus suggesting that laboratory practice cannot serve as a substitute for extensive practice in real-world conditions. For instance, Gast (2014) rightly asserts that "MacLeod and Dunbar (1988) showed that newly learned naming responses can lead to interference in a Stroop task after five one-hour training sessions," but it is important to add that the observed pattern of interference did *not* imply word reading. Again, we believe that conceiving the different paradigms in terms of complementarity is a better approach than reasoning in terms of competition and exclusive alternatives.

Conclusion

Zakay (2014) worries that "when names of notes are written within note pictures on a staff, a very particular and unusual non-ecological condition is created." We agree. But, except for experimental psychologists, is it more usual to be asked to name the color of incongruent color words? The grounding principle of a Stroop task is to involve an unusual situation, because this condition is necessary to mislead our automatisms. Of course, there are other paradigms for assessing automaticity, as Gast (2014) recalls. However, insofar as one includes Stroop paradigms among the worthwhile approaches of automaticity, we did not find in the commentaries any arguments that could invalidate the ability of our procedure to fulfill its primary objective, namely providing a better control on the conditions and the level of training than most classical versions.

Acknowledgments

This research was supported by Grant No. ANR-09-JCJC-0129 and the grant Conseil Régional de Bourgogne No. 2012-9201AAO05S01353.

References

Akiva-Kabiri, L., & Henik, A. (2014). Additional insights: Commentary on "The musical Stroop effect: Opening a new avenue to research on automatisms" by L. Grégoire, P. Perruchet, and B. Poulin-Charronnat (*Experimental Psychology,* 2013, Vol. 60, pp. 269–278). *Experimental Psychology, 61*, 75–77. doi: 10.1027/1618-3169/a000208

Gast, A. (2014). What is learned, and when? Commentary on "The musical Stroop effect: Opening a new avenue to research on automatisms" by L. Grégoire, P. Perruchet, and B. Poulin-Charronnat (*Experimental Psychology,* 2013, Vol. 60, pp. 269–278). *Experimental Psychology, 61*, 71–74. doi: 10.1027/1618-3169/a000206

Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2013). The musical Stroop effect: Opening a new avenue to research on automatisms. *Experimental Psychology, 60*, 269–278. doi: 10.1027/1618-3169/a000197

Hodges, D., & Nolker, B. (2011). The acquisition of music reading skills. In R. Colwell & P. Webster (Eds.), *MENC handbook of research on music learning, volume II: applications* (pp. 61–91). Oxford, UK: Oxford University Press.

Lindsay, D. S., & Jacoby, L. L. (1994). Stroop process dissociations: The relationship between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 219–234.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.

MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition, 26*, 201–211.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 126–135.

Moeller, B., & Frings, C. (2014). How automatic is the musical Stroop effect? Commentary on "The musical Stroop effect: Opening a new avenue to research on automatisms" by L. Grégoire, P. Perruchet, and B. Poulin-Charronnat (*Experimental Psychology,* 2013, Vol. 60, pp. 269–278). *Experimental Psychology, 61*, 68–70. doi: 10.1027/1618-3169/a000204

Zakay, D. (2014). Can the "musical Stroop" task replace the classical Stroop task? Commentary on "The musical Stroop effect: Opening a new avenue to research on automatisms" by L. Grégoire, P. Perruchet, and B. Poulin-Charronnat (*Experimental Psychology,* 2013, Vol. 60, pp. 269–278). *Experimental Psychology, 61*, 78–79. doi: 10.1027/1618-3169/a000211

Zakay, D., & Glicksohn, J. (1985). Stimulus congruity and S-R compatibility as determinants of interference in a Stroop-like task. *Canadian Journal of Psychology, 39*, 414–423.

Received April 22, 2013
Revision received May 16, 2013
Accepted May 21, 2013
Published online January 21, 2014

Laurent Grégoire

LEAD-CNRS UMR5022
Université de Bourgogne
Pôle AAFE
11 Esplanade Erasme
2100 Dijon
France
E-mail laurent.gregoire@u-bourgogne.fr