Psychological Research
DOI 10.1007/s00426-007-0108-4

ORIGINAL ARTICLE

Offline and online automatic number comparison

Filip Van Opstal · Agnes Moors · Wim Fias · Tom Verguts

Received: 31 March 2006 / Accepted: 16 January 2007
© Springer-Verlag 2007

Abstract This paper focuses on automatic number processing as instantiated in the size congruity effect. It was recently argued that long-term "associations between individual digits and the attributes 'small' and 'large' create the size congruity effect" (Choplin and Logan 2003, abstract, p. 17). Moreover, these authors proposed the additional assumption that the relevant connections are acquired over a lifetime of experience with numbers. We show that at least one of these assumptions is not true: either the size congruity effect derives from an (online) comparison effect between two numbers at the time of stimulus presentation (violating the first assumption) or the relevant connections flexibly change (offline) between trials during the course of one experimental session (violating the second assumption).

Offline and online automatic number comparison

Many authors have found numbers and numerical processing an interesting domain to (re-)evaluate extant issues in cognitive science (e.g., locus of negative priming, Lammertyn & Fias, 2005; unconscious semantic priming, Naccache & Dehaene, 2001; consciousness as an all-or-none phenomenon, Sergent & Dehaene, 2004). One of these issues is automaticity (Logan, 1988). It has been shown by many authors that a number's magnitude is processed even when it is irrelevant (Dehaene, Bossini, & Giraux, 1993) or disadvantageous (Henik & Tzelgov, 1982) to optimal execution of the experimental task. Number stimuli have been used to address questions such as how to explain automaticity (Choplin & Logan, 2003), whether there are side conditions to automaticity (Pansky & Algom, 1999), and how automatic and non-automatic processes interact (Gevers, Caessens, & Fias, 2005; Gevers, Ratinckx, De Baene, & Fias, 2006a; Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006b; Schwarz & Heinze, 1998).

One popular paradigm to study automaticity in numerical cognition is the size congruity paradigm (Henik & Tzelgov, 1982). In a typical size congruity experiment, two numbers are presented simultaneously. The numbers differ with regard to their numerical size as well as with regard to their physical size, and participants are asked to decide which of the two digits is physically larger (or smaller). Size congruity entails that responses are faster on trials in which the physically larger digit is also numerically larger (i.e., congruent trial, e.g., 2–6) than on trials in which the physically larger digit is numerically smaller (i.e., incongruent trial, e.g., 2–6). The occurrence of a size congruity effect demonstrates that the digit's numerical size is processed despite the fact that it is irrelevant for the task (which is to respond to physical size). This suggests that the numerical size of digits can be processed automatically, at least in the sense that it is processed without the conscious intention to do so.

F. Van Opstal (✉) · W. Fias · T. Verguts
Department of Experimental Psychology,
Ghent University, H. Dunantlaan 2,
9000 Ghent, Belgium
e-mail: Filip.VanOpstal@UGent.be

A. Moors
Department of Experimental Clinical
and Health Psychology,
Ghent University, Ghent, Belgium

At least two views have been proposed with regard to the mechanisms underlying the size congruity effect. One view is that the numerical sizes of the digits in each pair are automatically compared online (i.e., at the time of stimulus presentation) and that the result of this automatic comparison (e.g., 2 is numerically *smaller* than 6) interferes with the result of the explicit comparison of their physical sizes (e.g., 2 is physically *larger* than 6). We will call this the *comparison view* of size congruity. A typical finding is that the size congruity effect exhibits a congruity × numerical distance interaction, meaning that the congruity effect is larger for larger numerical distances. For example, the congruity effect is larger for the pair 2–6 than for the pair 2–3 (e.g., Schwarz & Ischebeck, 2000). From the comparison view, this can be easily explained by the fact that number comparison is faster for larger numerical distances (Moyer & Landauer, 1967), leading to more intrusion than for smaller distances.

Choplin and Logan (2003, 2005) recently questioned the conclusion that a comparison process is responsible for the size congruity effect. In line with Logan's (1988) theory of memory-based automaticity, they proposed an alternative *direct retrieval view*. This theory holds that the size congruity effect does not originate from an automatic comparison of the two numbers, but rather from an automatic retrieval of long-term memory tags associated with the individual numbers (see also Tzelgov, Meyer, & Henik, 1992). Smaller numbers are more strongly connected to the label "small", and larger numbers are more strongly connected to the label "large". If the task is, for example, to choose the physically larger of two numbers, the pair 2–6 is easier than 2–6. According to the direct retrieval view, this is because choosing the physically larger number 2 in the latter pair is hindered by the label "small" that it automatically activates. Choplin and Logan (2005) argued that these labels arise during the course of a lifetime as a result of association. For example, in daily life number 1 is associated more often with "small" than larger numbers are. This leads to a gradient of labels in memory, with a very strong association between 1 and "small", a somewhat less strong association between 2 and "small", and so on.

Choplin and Logan (2003, 2005) reasoned that if the size congruity effect would still be obtained in trials where only a single number is presented, an explanation in terms of a comparison algorithm would be falsified because a comparison algorithm requires the presentation of two digits. They obtained size congruity effects using single digits.[1] From this, they concluded that the size congruity effect was due to direct retrieval of the labels "small" and "large" associated with individual stimuli.

A new light can be shed on Choplin and Logan's (2003, 2005) finding from the observation that some numerical effects, such as the SNARC effect (Dehaene et al., 1993), are dependent on the range of numbers used. The SNARC effect consists in faster left-sided responses to small numbers and faster right-sided responses to large numbers. Range dependency entails that the direction of the SNARC effect depends on the range of numbers used in an experiment (Dehaene et al., 1993; Fias, Brysbaert, Geypens, & d'Ydewalle, 1996). For example, in the range 1–5, the numbers 4 and 5 are preferentially responded to with the right hand, whereas in the range 4–9 the numbers 4 and 5 are preferentially responded to with the left hand. Recently, Gevers et al. (2006a, b) have proposed a computational model in which the range dependency effect can be explained if one assumes that a number is compared to the mean of the range of numbers in the experimental set. A computationally similar mechanism leading to essentially the same predictions is that the target number is not compared to the mean of the range, but rather to the whole number range at once. The crucial aspect is the same in both mechanisms: A comparison is made, either to the mean of the range or to the whole range at once. A variety of SNARC-related phenomena can be explained by such comparison mechanisms (see Gevers et al., 2006a, b, for further details). It is possible that a similar effect was at work in Choplin and Logan's (2003, 2005) experiments. Even when a single number is presented on the screen, it may be automatically compared to the mean of the numbers in the experimental set or to the whole range of the numbers in the set, thus leading to the size congruity effects described by Choplin and Logan. If so, the conclusion drawn by Choplin and Logan that a single number cannot activate a comparison mechanism, does not hold. Hence, the issue of whether size congruity is based on comparison or on direct retrieval is open again. Our experiment directly

[1] Other researchers already obtained size congruity effects with single digits (e.g., Algom, Dekel, & Pansky, 1996; Schwarz & Ischebeck, 2000; Tzelgov et al., 1992), but in their experiments, trials on which numerical size was irrelevant were always mixed with trials on which it was relevant (either blocked or not). On the relevant trials, participants were instructed to compare the numerical size of the presented digit with the numerical size of a fixed standard. Because of this, one could argue that participants continued engaging in a numerical comparison with the standard on the irrelevant trials, despite the instruction to respond on the basis of physical size.

tests whether a number is compared to the mean of the set of numbers from which the number is sampled. To do so, we used a design similar to Choplin and Logan's (2003) Experiment 2, but with a varying number range.

Experiment

As in Choplin and Logan (2003), the task was to press a button depending on whether a square surrounding the number was small or large. The number appearing in the square was always irrelevant to the task. The performance on two numbers, 4 and 5, was considered. Both numbers were presented in two different ranges: either 1–5 or 4–8. The comparison view holds that on each trial, the presented number is compared to the mean of the total range of the numbers in the experimental set. That is, in the range 1–5, numbers 4 and 5 are compared to number 3 (i.e., the mean for the set 1, 2, 3, 4, 5), and they should facilitate a "large" response. Conversely, in the range 4–8, numbers 4 and 5 should facilitate a "small" response because they are both smaller than 6 (i.e., the mean of 4, 5, 6, 7, and 8). The comparison view predicts a size congruity effect for numbers 4 and 5 in both ranges, but the direction of the effect should be different depending on the range: an advantage for response "large" in the range 1–5, and an advantage for response "small" in the range 4–8. The direct retrieval view as outlined by Choplin and Logan assumes that numbers evoke long-term memory labels that have been associated with them throughout the participants' life. The numbers 4 and 5 should consistently activate the label "small" or consistently activate the label "large". This view predicts that the direction of the congruity effect caused by these labels will not shift between ranges.

Method

Design and stimuli

On each trial, a square containing a number was shown. Importantly, participants were not informed that numbers would be presented in the square. They were instructed to judge the size of the square (3 or 4 cm) as small or large by pressing one of two buttons on a response box. The number was presented in Courier-30 font (following Choplin and Logan's, 2003, Experiment 2). Participants received eight practice trials to determine which square was small and which was large. Feedback was given during the practice trials but not during the experimental trials. Each participant received both ranges (1–5 and 4–8). Hand mapping (small-left or small-right) and order of range (small range first or large range first) were counterbalanced across participants. After the practice trials, participants received 800 experimental trials, 400 for each range, in a quasi-random order. The experiment lasted about 20 min.

Participants

Sixteen undergraduate psychology students from Ghent University participated for monetary reward. None was familiar with the purpose of the experiment.

Results

Reaction times faster than 200 ms or slower than 1,000 ms were removed from the analysis. Accuracies were very high (97%). Consequently, no effects were significant for the accuracy data and so these data are not discussed further.

The reaction time data were analyzed with a 2 (response: small/large) × 2 (range: small/large) × 2 (number: 4/5) repeated measures ANOVA and are plotted in Fig. 1. This analysis revealed a significant interaction between response and range, $F(1,15) = 4.95$, $MSe = 485$, $P < 0.05$. As predicted by the comparison view, digits 4 and 5 received faster "large" than "small" responses in the small range, but faster "small" than "large" responses in the large range.

There was also a significant interaction between response and number, $F(1,15) = 5.28$, $MSe = 335$, $P < 0.05$. This effect is equivalent to the observation typically found in size congruity experiments (e.g., Schwarz & Ischebeck, 2000) that the congruity effect is larger at distances far from the mean than at distances close to the mean. It can indeed be shown that a congruity × distance interaction is mathematically equivalent to the obtained response × number interaction.[2] In our experiment, the congruity effect was larger for 5 in the small range (with mean 3) and for 4 in the large range

[2] The interaction test between response (small/large) and number (4/5) is equivalent to a contrast that assigns a weight of +1 to the two "response = small, number = 4" conditions and the two "response = large, number = 5" conditions, and a weight of -1 to the other four conditions. Alternatively, the eight conditions can be rearranged in a congruity (congruent/incongruent) × distance (small/large) × response (small/large) format. The interaction between congruity and distance is equivalent to a contrast that assigns a weight of +1 to the two "congruity = incongruent, distance = small" conditions and to the two "congruity = congruent, distance = large" conditions, and a weight of −1 to the other four conditions. The weight assignments are equivalent for the two situations, and therefore the response × number and congruity × distance interactions are equivalent.

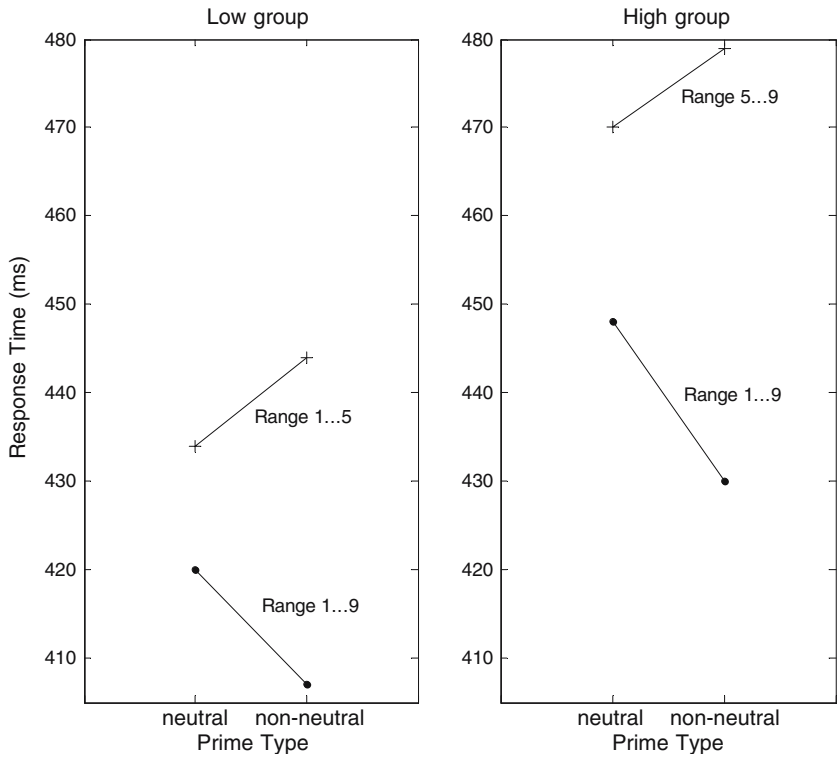


Fig. 1 Reaction times on numbers 4 and 5 in the small range (*left panel*) and in the large range (*right panel*)

(with mean 6) and in these cases, the distance from the standard was largest. No other effects or interactions were significant (all Fs < 1).

Discussion

The results of this experiment show that the size congruity effect depends on the range of the numbers used in the experimental set. This is in line with the predictions of the comparison view that on each trial the presented number is compared to the mean of the range, even if the number is completely irrelevant to the task. The results cannot be explained by a direct retrieval view that each number activates the same long-term memory label throughout the experiment. The conclusion is warranted that the comparison was automatic rather than controlled, as seen by application of the criteria outlined by Moors and De Houwer (2006). In this experiment, numbers were compared and led to the retrieval of their long-term memory labels even though (a) participants had to engage in the demanding task of evaluating the physical size of the squares as quickly as possible, and (b) participants probably had the intention to avoid processing the numerical size of the digits rather than to engage in the processing of it. It can therefore be concluded that the comparison and retrieval processes at work occurred automatically, in the sense of efficient and uncontrolled.

The automatic comparison process can, however, be online or offline. Online comparison holds that the presented number is compared to the mean of the range on every trial. Another possibility consistent with the data is offline comparison, meaning that during the course of the experiment, new associations between numbers and the labels "small" and "large" are formed. For example, participants exposed to the large-number range (4–8) may relabel 4 and 5 as small because 4 and 5 are the smallest numbers in that range. It must be clear, however, that in both the online and offline view, numbers are compared. This is clear in online comparison, but it must also be the case in offline comparison. Consider, for example, the situation in which a participant in the first phase saw numbers from a small range (1–5) and afterwards from a large range (4–8). Because 4 and 5 were processed as large in the first phase but not in the second phase, it must be the case that the numbers 4 and 5 were reassigned a new label before (or during) the second phase. This labeling process requires locating 4 and 5 in the new range of numbers, and this requires a comparison process. This relabeling should occur despite the fact that the numbers are irrelevant; furthermore, the resulting connections should be strong enough to interfere with the subsequent processing of physical size. The offline comparison account is consistent with Choplin and Logan's explanation of the size congruity effect as a result of retrieval of labels from memory, but

inconsistent with their assumption that the labels retrieved are fixed throughout life. At the very least, our data urge the direct retrieval model to allow a rapid reorganization of labels.

We have emphasized the difference between the comparison view and the direct retrieval view. Part of the argumentation invoked by Choplin and Logan (2003) to doubt the possibility of automatic number comparison was that they granted automatic processes the possibility to evoke specific responses upon specific stimuli (which are formally one-place predicates, e.g., *small*{*A*}), but they doubted whether automatic processes could handle two-place predicates (e.g., *smaller-than*{*A*, *B*}). When interpreted in neural network terms, however, the direct retrieval view and comparison view, or one-place predicates versus two-place predicates, are not fundamentally different and can be integrated in a common framework. To see this, consider the graded associations and retrieved instance models proposed by Choplin and Logan (2005) as two specific instantiations of the direct retrieval model. In these models, there is a gradient of connection weights from individual numbers to response labels "Small" and "Large". As a consequence, when a number is presented, it activates the two labels to different degrees. The gradient of associations was invoked by Choplin and Logan to account for parts of their data. But actually, this gradient of associations can be put to use to perform number comparison. To see this, suppose that the standard number has a gradient of connections to the two response labels, but the gradients are reversed, such that, the connections are increasing to one of the response labels but decreasing to the other response label. If this is the case, the comparison task can be solved by simply reading off which of the labels receives the larger input. And in fact, existing neural network models of numbers and of non-numerical orders solve (number) comparison in either this or a formally related manner (e.g., Leth-Steensen & Marley, 2000; Verguts, Fias, & Stevens, 2005). In this way, online comparison can be implemented in a neural network. Implementation of offline comparison is similar except that it must be additionally assumed that the training process occurs very fast during the instructional phase and/or between trials. Hence, although offline and online comparison cannot be distinguished by the present data, both are actually very similar except that offline comparison requires an extra assumption that learning is very fast because appropriate connections must be generated during instruction or between trials early in the test phase. It is possible that other interpretations account for our data as well, but our account at least is consistent both with the data and with the larger framework of connectionist learning models of ordered quantities (see Leth-Steensen & Marley, 2000; Verguts et al., 2005, respectively, for reviews of non-numerical and numerical models, respectively). Whatever the mechanism, our results highlight the surprising flexibility of automaticity.

Acknowledgments The authors want to thank Jan De Houwer for helpful suggestions. WF is supported by the Inter University Attraction Pole P5/05 (Belgian Government) and by a GOA grant from the Ghent University Research Fund. The contribution of FVO and TV was supported by grant G.0188.04 from the Fund of Scientific Research—Flanders

References

Algom D., Dekel A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory & Cognition, 24*, 557–572.

Choplin, J. M. & Logan, G. D. (2003). Automaticity of number perception. In R. Alterman, & D. Kirsh (Eds), Proceedings of the Twenty-fifth Annual Conference of the Cognitive Science Society, 252–257.

Choplin, J. M., & Logan, G. D. (2005). A memory-based account of automatic numerosity processing. *Memory & Cognition*, *33*, 17–28.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396.

Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition*, *2*, 95–110.

Gevers, W., Caessens, B., & Fias, W. (2005). Towards a common processing architecture underlying Simon and SNARC effects. *European Journal of Cognitive Psychology*, *17*, 659–673.

Gevers, W., Ratinckx, E., De Baene, W., & Fias, W. (2006a). Implicit number–space links activate spatial responses: Evidence from the LRP. *Experimental Psychology*, *53*, 58–68.

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006b). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 32–44.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*, 389–395.

Lammertyn, J., & Fias, W. (2005). Negative priming with numbers: No evidence for a semantic locus. *Quarterly Journal of Experimental Psychology: A*, *58*, 1153–1172.

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, *107*, 62–100.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519–1520.

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, *80*, 215–229.

Pansky, A., & Algom, D. (1999). Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 39–58.

Schwarz, W., & Heinze, H. J. (1998). On the interaction of numerical and size information in digit comparison: A behavioral and event-related potential study. *Neuropsychologia*, *36*, 1167–1179.

Schwarz, W., & Ischebeck, A. (2000). Sequential effects in number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1606–1621.

Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science*, *15*, 720–728.

Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 166–179.

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*, 66–80.

Psychological Research
DOI 10.1007/s00426-007-0119-1

ERRATUM

Offline and online automatic number comparison

Filip Van Opstal · Agnes Moors · Wim Fias · Tom Verguts

© Springer-Verlag 2007

Erratum to: Psychol Res
DOI 10.1007/s00426-007-0108-4

It has been drawn to our attention that Fig. 1 is incorrect. The correct figure is printed below.

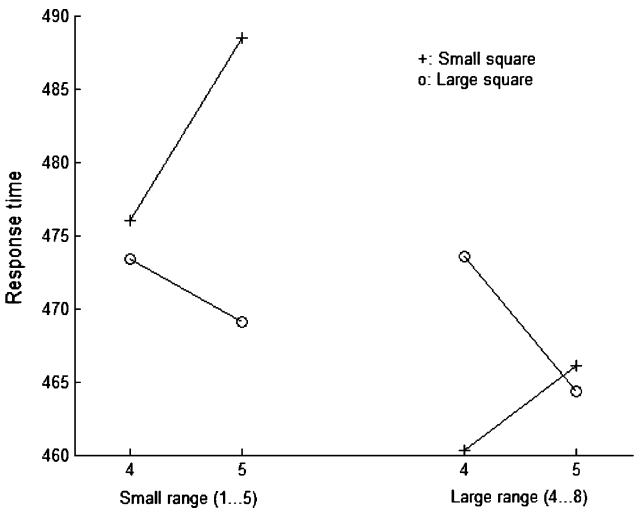


Fig. 1 Data of experiment: mean reaction times on numbers 4 and 5 in the small range (*left panel*) and in the large range (*right panel*)

The online version of the original article can be found under doi:10.1007/s00426-007-0108-4.

F. Van Opstal (✉) · W. Fias · T. Verguts
Department of Experimental Psychology, Ghent University,
H. Dunantlaan 2, 9000 Ghent, Belgium
e-mail: Filip.VanOpstal@Ugent.be

A. Moors
Department of Experimental Clinical and Health Psychology,
Ghent University, Ghent, Belgium