Notation Specificity in Numerical Cognition: A critique of Noël and Seron's (1992) reappraisal of the Gonzalez and Kolers (1982) study.

Iris van Rooij
Department of Psychology
University of Victoria, Canada, 2000

Abstract

Gonzalez and Kolers (1982) performed a study on mental arithmetic with arabic and roman numerals, and interpreted their findings as indicating *notation specificity* of mental arithmetic operations. This conclusion has been criticized, among others, by Noël and Seron (1992) on the argument that the effects of notation observed by Gonzalez and Kolers (1982) may arise at an "encoding stage". I show that this "encoding locus" hypothesis is contradicted by Noël and Seron's (1992) own data. Therefore, Noël and Seron's findings do not refute, but actually *support*, the original claim made by Gonzalez and Kolers (1982).

Consistent with common representational assumptions in research on semantic memory in general, theories of numerical cognition have tended to assume that number processing is mediated by internal (abstract) codes (e.g. see McCloskey, 1991, for an overview). As a consequence, number processing is often assumed to be *notation independent*. Kolers (Kolers & Roediger, 1984; Kolers & Smythe, 1979, 1984) criticized the distinction between semantic and procedural memory, and proposed that number related processes should be viewed as a matter of skill in operating on symbols, not on "meanings" abstracted from them.

Gonzales and Kolers (1982) performed a study on mental arithmetic that addressed the issue of notation specific influences in number processing. Their study is an important for-runner of the debate on notation specificity in numerical processing today; mainly between Campbell (e.g., Campbell, 1994a; 1994b; 1997; 1999; Campbell, Kanz, & Xue, 1999), on the one hand, and McCloskey (e.g., McCloskey, 1992; McCloskey, Macaruso and Whetstone, 1992; Sokol et al., 1991) on the other. Opponents of notation specificity, however, have disregarded the study of Gonzalez and Kolers, on the argument that the effects observed by Gonzalez and Kolers can be understood as arising at an "encoding stage" (e.g., McCloskey, 1992; Noël & Seron, 1992; Sokol et al., 1991). Noël and Seron (1992) explicitly defended this hypothesis in an article where they reported findings that, on their view, contradict Gonzalez and Kolers' interpretation. Here I re-evaluate Noël and Seron's study, and set out to show that their approach and arguments are not viable. I will start with a short summary of the original Gonzalez and Kolers (1982) study. After that I will discuss Noël and Seron's reappraisal of Gonzalez and Kolers' study.

Skill versus Abstraction

Gonzales and Kolers (1982) measured reaction times (RTs) in a simple arithmetic verification task. Problems were presented in arabic (A + A = A) or roman format (R + R = R), and in six permutations of arabic/roman combinations (e.g. A + R = A, R + A = R, etc.). The subject's task was to indicate whether the equation was true or false. Gonzales and Kolers found that RTs varied with the format of the symbols; e.g., 2 + 2 = 4 was verified faster than, for example, 2 + 2 = IV and 2 + II = 4. Gonzales and Kolers also observed that the *position* of the roman numeral in the equation was important. Verification was faster for ARA, RAA as compared to AAR, and faster for RRA as compared to ARR and RAR. In other words, roman numerals were more disruptive as operands than as solutions. Based on their findings Gonzales and Kolers argued that arabic and roman symbols "do not seem to share a common abstract representation for which the two kinds of symbols are interchangeable surface expressions" (Gonzales & Kolers, 1982, p. 318). Instead, roman and arabic symbols represent quantity in a different way, and doing mental arithmetic with them requires different mental operations.

Because Gonzales and Kolers were aware of a possible criticism that their effects were mainly due to differences in familiarity of the two different kinds of symbols, they performed a follow-up study in which subjects were trained in naming roman numerals[1]. Again, in the verification task following this procedure, Gonzales and Kolers found the same effects as they found in their first experiment. On Gonzales and Kolers' view this finding shows that the differences, between arabic and roman numerals, in time to carry out arithmetic operations were not due to lack of familiarity of the roman symbols, but due to lack of skill in carrying out the appropriate operations on them required for the arithmetic verification task. Namely, training in naming did not transfer to mental arithmetic with roman numerals, because processes involved in naming are not the same as in mental arithmetic (a view related to the notion of *transfer-appropriate processing* (e.g. Roediger, Weldon, & Challis, 1989) and the *encoding specificity principle* (e.g. Tulving, 1984)).

Perceptual Structure Hypothesis

Gonzalez and Kolers' (1982) study has been criticized on the grounds that their findings may simply reflect differences in time to *encode* roman and arabic numerals (McCloskey, 1992; Noël & Seron, 1992; Sokol et al., 1991). For example, Sokol et al. (1991) claimed that "most if not all of Gonzalez and Kolers findings may be interpreted [by] assuming that roman numerals take longer to comprehend (i.e., to translate into abstract internal representations) than do arabic digits" (p. 366). Noël and Seron (1992) performed a study to test this claim. They replicated most of the effects observed by Gonzalez and Kolers, but argued that the findings may reflect peculiarities of the roman code, leading to a "perceptual structure effect." They analysed the characteristics of the roman numerals I through XII, and identified roughly three different ways in which roman numerals represent quantity: (1) "analogical" structures, including I, II and III, (2) "symbolic" structures, including V and X, and (3) "complex" structures, which are combinations of the analogical and symbolic structures. The latter subclass can again be divided in combinations that involve addition (VI, VII, VIII, XI and XII) and combinations that involve subtraction (IV and IX). Noël and Seron hypothesized that these different structures "present a gradient of difficulty to be comprehended, with the simplest one being the analogical structure, followed by the symbolic structure, and then the additive and subtractive complex ones" (p. 455).

To obtain estimates of 'comprehension time' for each arabic and roman numeral Noël and Seron performed two experiments, Experiment 1 involved a number transcoding task (e.g., IV = 4? or III = 5?) and Experiment 2 a parity judgement task (e.g., Is IV odd or even? Is 5 odd or even?). In both tasks three different symbol systems were used: arabic numerals, number words and roman numerals. Based on the estimates of "encoding time" obtained by averaging over RTs observed for a specific numeral in Experiment 1 and 2, they predicted the time required to encode the problems that Gonzalez and Kolers used in their experiments. By simply assuming "that the time necessary to encode an equation of the type $p + q = n$ would be a function of the simple addition of the estimated encoding times for each of the three numerals presented" (p. 467) the pattern of reaction times observed by Gonzalez and Kolers (for ties equations) could be more or less simulated. Hence, Noël and Seron argued, the effect of notation can be explained in terms of encoding alone.

To further support this claim, Noël and Seron (1992; Experiment 3) performed a variant of the arithmetic verification task used by Gonzalez and Kolers. Addends were presented first, either in arabic code (AA) or roman code (RR) or a combination of the two (AR, RA), and subjects had to indicate when they had identified the two numerals (RT_i). "[RT_i] was taken as a measure of the comprehension time for the addends" (Noël & Seron, 1992, p. 469). Subsequently, an operation sign (+, - or x) and an answer in word form (W) were presented and subjects had to verify the truth of the equation (RT_v). "[RT_v] was taken as the time to encode the arithmetic sign and the number word, perform the arithmetic operation, and verify whether the presented response was correct or incorrect" (p. 469).

---

[1] Training proceeded until performance was within 10% of the speed in naming arabic numerals. McCloskey, Cohen and Aliminosa (1991) have argued that the residual 10% difference in naming time, arguably still allows for a "lack of familiarity" interpretation.

Noël and Seron argued that if the effects that Gonzalez and Kolers observed were only due to differences in encoding time between the different symbols no interaction should be seen in RT_v between *problem size* (as indexed by either the size of the smallest addend, or the size of the solution) and notation. No Problem Size x Notation interaction was found by Noël and Seron, and again they saw support for their claim that notational characteristics do not influence processes beyond the assumed encoding stage.

I believe there are serious conceptual and methodological problems with the approach Noël and Seron took in this issue. First of all, the finding that differences between format are more pronounced when comparing certain kinds of roman numerals (complex ones) with arabic numerals (and number words) in itself does not contradict Gonzalez and Kolers' view and interpretation. This idea, I believe, rests on a misconception. Second, the rationale of Noël and Seron's method, in Experiment 1 and 2, for obtaining estimates of encoding time rests on the assumption that (1) a task independent encoding stage exists, and (2) notational characteristics do not influence processing beyond the encoding stage. I show that these assumptions are unsupported by the very data they used for obtaining the estimates. Further, the absence of a problem Size x Notation interaction in their Experiment 3, is only indirect (and questionable) evidence for an encoding locus interpretation, and a more direct prediction that would follow from the "perceptual structure hypothesis" is falsified by the same data. Finally, there is still the position effect observed by Gonzalez and Kolers that remains unexplained by the "perceptual structure hypothesis." I will discuss each of these four shortcomings of the Noël and Seron study in turn.

Sometimes V ≈ 5, but IV ≠ 4

Recall that Noël and Seron set out to show that the effect of code observed by Gonzalez and Kolers may be due to peculiarities of the roman code, which is constituted of some analogical and symbolic structures, or complex combinations of the two. Their analysis, however, does not contradict Gonzalez and Kolers view at all. On the contrary, Gonzalez and Kolers were well aware of the special nature of the roman number system. For example, they observed that reaction times were faster when the smallest addend in a problem presentation was in roman than when it was in arabic, except for when the roman numeral was 'IV'. Gonzalez and Kolers argued that people use the tallylike nature of the roman numerals using strokes as analog counts and found support for this conclusion in the fact that the advantage of roman numerals as smallest addend was not found when the smallest addend was IV, where the analog property is destroyed (p. 317).

If Noël and Seron thought that Gonzalez and Kolers viewed all roman numerals as equivalent and all to the same extent different from arabic numerals, I think they missed the point. Gonzalez and Kolers were not out to show simply differences between, presumably homogeneous, notational systems per se. They wished to show that, because symbols may differ in formal, as well as perceptual, properties mind operates on them differently.

Gonzalez and Kolers were aware of the possibility of criticism like the one Noël and Seron gave, as they wrote:

> A strict serial-processing theorist might propose that the actual reaction times can be divided into at least two components, an acquisition phase followed by a computation phase (…) In our view this is an erroneous account, for we hold that the form of the symbol is part of the computation. (Gonzalez & Kolers, 1982, p. 318)

For example, the finding that the truth of the equation "4 + II = 6" is verified faster than "2 + IV = 6", Noël and Seron may interpret as reflecting a "perceptual structure effect" (viz., both are of the form R + A = A). However, this finding can just as well be interpreted as reflecting the fact that mental arithmetic with an analog numeral (in this case 'II') allows for operations that a complex numeral (in this case 'IV') does not. For the same reasons, Noël and Seron's finding that, for example, in a number identification task (Experiment 1, RT_1) 'V' and '5' were identified just a quickly, while '4' was identified faster than 'IV', does not contradict Gonzales and Kolers view. The comparable speed in identifying 'V' and '5' may indeed by due to the fact that these symbols represent quantity in a similar way, whereas '4' and 'IV' do not. Gonzalez and Kolers (1982, p. 318) emphasize, however, that "[e]quivalences among different symbol systems should be seen as due not to communality of [internal] representation but rather to communality of interpretation and use".

If differences in performance indeed reflect differences in interpretation and use, then *task specific* effects can be expected. Interestingly, Noël and Seron (1992) observed such effects, although, for apparent reasons, they did not interpret them accordingly. For example, they observed that '2', 'two' and 'II' were identified equally fast (viz., the RTs estimated from the graph in (Noël and Seron, 1992) are 500ms, 510 ms and 490 ms, respectively) while the judgement that 'II' is an even number takes much longer (850 ms) than judging '2' and 'two' as even (700ms and 690 ms respectively). On Noël and Seron's view encoding time is included in the identification response, and apparently no differences in encoding time between '2', 'two' and 'II' are found. How then can it be that it takes longer to make a parity judgement for 'II', as compared to '2' and 'two'? This effect becomes insightful when one acknowledges that the mental operations that will allow one to judge whether or not a numeral is even or odd are different from the mental operations that will allow one to judge the numerical quantity that a symbol represents.

Task x Notation Interaction

As we have seen, Noël and Seron's analysis of the special characteristics of the roman number system does not in itself contradict Kolers' procedural account. The fact that some of the roman numerals I through XII represent quantity in a fundamentally different way from the arabic symbols 1 through 12, is even prerequisite for arguing differences in use and interpretation that are task dependent. The burden of proof for the "perceptual structure hypothesis," then, is to show that these characteristics *only* influence encoding; i.e., to show that they do *not* influence any processes that on a "separate stage" account occur *after* the encoding stage. The example of differences between identification time and parity judgement time for the roman numeral 'II' already suggested that such a claim may not hold. Here I discuss Experiment 1 and 2 performed by Noël and Seron (1992) in more detail, and will argue that the whole idea of being able to obtain clean, task independent, estimates of the time to encode a numeral is flawed and contradicted by their own data.

In the number transcoding task two reaction time measures were obtained: first a symbol was presented and subjects had to indicate when they had identified the symbol (RT_1). Subsequently the second symbol was presented and subjects' had to indicate whether or not it represented the same numerical value as the first (RT_2). Though, Noël and Seron acknowledged that "it is difficult to judge what the basis for responding was" (p. 462) in the identification part of this task, they do assume that this judgement at least involves "semantic encoding." Hence, RT_1 can be taken as the time it takes to encode a symbol *plus* whatever else is required to make the response; say, Time(encode + X). RT_2 can be seen as corresponding to the time required to encode the second symbol *plus* the time for the verification operation *plus* response related processes; say, Time(encode + Y). The first and second column of Table 1 give the mean RT_1 and RT_2 that Noël and Seron (1992) observed for the different symbol systems.

Table 1 *Mean reaction times obtained by Noël and Seron (1992) for identification (RT_1), verification (RT_2) and parity judgement (RT_3) for Arabic, Word and Roman symbols.*

Symbols	RT_1	RT_2	RT_3
Arabic	494.0	638.0	723.0
Word	493.0	671.5	752.0
Roman	538.0	757.5	969.0

Table 2 *Differences between reaction time measures obtained by Noël and Seron (1992) for identification (RT_1), verification (RT_2) and parity judgement (RT_3) for Arabic, Word and Roman symbols.*

Symbols	RT_2 - RT_1	RT_3 - RT_1	RT_3 - RT_2
Arabic	144.0	229.0	85.0
Word	178.5	259.0	80.5
Roman	219.5	431.0	211.5

Noël and Seron (1992) performed analyses on RT_1 and RT_2 separately and concluded that roman numerals take overall longer to encode as indicated by differences between roman numerals, on the one hand, and arabic and number-words on the other, both in RT_1 and RT_2 separately. But, if indeed RT_1 would represent Time(encode + X) and RT_2 would represent Time(encode + Y) than the difference between these two reaction time measures would represent (the difference between) processes that occur *after* encoding, i.e. RT_2-RT_1 = Time(Y - X). If indeed notation does not influence processing other than encoding, Time(Y - X) should be the same for all three symbol systems. The first column of Table 2 shows that this is not what Noël and Seron found: RT_2-RT_1 is larger when the second symbol is a roman numeral, than when it is an arabic numeral or number word. Though it's not possible to test the significance of these differences, the difference between RT_2-RT_1 for arabic and roman numerals seems sufficiently large to argue that the assumption that a task-independent encoding time exists may not be warranted.

Inspection of how the reaction times that Noël and Seron obtained in their parity judgement task relate to RT_1 and RT_2 leads to the same observation. The third column of Table 1 gives the mean reaction times (RT_3) that Noël and Seron observed for the different kinds of numerals in this task. As can be seen RT_3 reflects an effect of symbol, which Noël and Seron again interpreted as being due to differences in encoding time (i.e., the time to perform the parity judgement can be represented by Time(encode + Z)). However, as can be seen in the second and third column of Table 2 both RT_3-RT_1 and RT_3-RT_2 are different for the different kinds of symbols.

Thus, a closer look at Noël and Seron's data from Experiment 1 and 2 makes "the simple assumption that the time necessary to encode an equation of the type $p + q = n$ would be a function of the simple addition of the estimated encoding times for each of the three numerals presented" (p. 467) very questionable. Instead, Noël and Seron's findings should be taken as supporting either (1) the view that notational characteristics do influence processing beyond the encoding stage, or (2) the idea that no such task independent encoding stage exists[2].

[2] Noël and Seron (1992) do mention that "the difference observed between RT's for the three codes varied across task" (p. 474), but just concluded that "the R code *globally* took more time to be processes that did the A and D codes" (p. 474, emphasis added). They also argued that this global effect of code is "very stable across tasks" (p. 467), and suggested that the differences in the effect of notation between tasks may be understood as being "a function of the task's degree of complexity" (p. 474). I do not see how all these remarks fit together, and certainly do not understand how saying that notation specific influences are "a function of the task's complexity" is any different for saying that the influence of notation is task specific, or saying that task performance is notation dependent.

Residual Effect of Notation

As described earlier, in Experiment 3 of Noël and Seron (1992) two RT-measures were obtained: RT_i denoted the time required to identify the two operands (either AA, RA, AR or RR), and RT_v denoted the time required to encode the operation sign (+, - or x), the solution (a number word) and verify the truth of the equation. Noël and Seron argued that no Problem Size x Notation interaction should have been found if effects of notation originate in encoding. Since, the problem size effect is typically associated with retrieval (i.e. central) processes (e.g., McCloskey, 1991) it is reasonable to argue, on an encoding locus assumption, that problem size and notation cannot interact. It is not obvious, however, why the reverse should hold as well; i.e. why would problem size and notation *necessarily* interact if notation effects do not originate in encoding?[3]. Besides, it is not clear why Noël and Seron chose such an indirect test of the "perceptual structure hypothesis" when a much more direct test was available. Namely, Noël and Seron could have simply argued that if indeed notation effects were only due to differences in encoding time of different symbols, an effect of notation should be observed for RT_i, but not for RT_v.

Noël and Seron indeed observed an effect of notation in RT_i, but they also observed a notation effect for RT_v; viz., mean RT_v showed the following pattern AA < AR = RA < RR (p. 470). Such a pattern, one would think, supports the idea of notation specificity in *arithmetic verification*. Surprisingly, Noël and Seron did not think so, and interpreted it as reflecting "a *residual* effect of encoding" (p. 473; emphasis added). That is, they proposed that by the time that the solution and the arithmetic operation appeared on the screen the operands (especially the complex roman numerals) had not yet been fully encoded. Hence, a residual effect of encoding was incorporated in RT_v.

It should be clear that this explanation is inconsistent with the fundamental premises of Noël and Seron. Namely, timing of the appearance of the operation sign and solution was controlled by participants themselves; i.e., by pressing a button when they had *identified* the two operands. How can one identify numerals without having encoded them? On a separate stage account there simply is no way. Hence, the "residual encoding" explanation forces Noël and Seron to give up the premise that identification requires encoding, in which case the whole notion of a task independent encoding stage is lost.[4] Besides, giving up encoding as a stage in the identification process would leave unexplained the effect of notation in RT_i in Experiment 3, as well as RT_1 in Experiment 1.

Noël and Seron cannot have it both ways; i.e., they cannot argue both that a task independent encoding stage exists *and* that RT_v reflects effects of encoding. Thus, as I said, the only interpretation that does justice to the theoretical commitments of Noël and Seron is that the notation effect in RT_v shows notation specificity of (central) verification processes.

Effect of Numeral Position

Recall that Gonzalez and Kolers found that roman numerals in the solution were less disruptive than roman numerals as one or more of the addends. Evidently this effect of position cannot simply be explained by assuming differences in encoding times for roman numerals as compared to arabic numerals, because there are, for example, as many roman numerals to encode in AAR problems as there are in ARA and RAR problems. Noël and Seron (1992, p. 457) acknowledged that their perceptual structure hypothesis does not allow for such an effect, but claimed that because the effect was more pronounced

[3] Even if it did, the *absence* of such an interaction effect in the data would only provide very weak support for an "encoding locus" interpretation, especially when power is not reported. Moreover, Noël and Seron did not even observe the problem size effect itself in RT_v.

[4] Note that if one would allow for the possibility that some kinds of decisions (e.g., identification) do not require the same extent, or the same kind, of encoding as other kinds of decisions (e.g., verification) this would inflate the notion of an encoding stage. If the encoding stage is task specific, then no clean estimates of encoding time can reasonably obtained (as Noël and Seron attempted in Experiment 1 and 2), nor would the question of whether or not the effect of notation is specific to encoding be tractable. Because we use differences between tasks to make inferences about the operation of different processes, we cannot tease apart stage-dependent influences if the operation of a stage is task dependent (e.g., Watkins, 1990).

when only one roman numeral was involved, and less so when two roman numerals were involved this effect did not seem as pervasive (p. 457). Also, the fact that the effect seemed to interact with the nature of the roman numerals Noël and Seron interpreted as reflecting encoding issues instead of arithmetic operation (but see section *Sometimes* $V \approx 5$*, but* $IV \neq 4$ for a criticism of this kind of reasoning).

Unfortunately, Noël and Seron's experiment 3 did not allow for replicating this same position effect, because they never used arabic or roman numerals as solution, only number words. Interestingly, though, they did report a different effect of position. That is, they found that for complex roman numerals RT_v in the arithmetic verification task was smaller for RD as compared to DR. This indicates that a complex roman numeral was more disruptive as second addend than as first addend. Such a position effect seems hard to reconcile with the perceptual structure hypothesis.[5]

Discussion and Conclusion

Gonzalez and Kolers viewed "semantic" knowledge, just like skills, as intimately tied to the specifics of the way in which it is learned. On this procedural account there are no separate *stages* of numerical processing (such as an encoding stage and a semantic stage), and hence notation specific effects in numerical tasks *by definition* are taken to show notation specificity of number processing *proper*. On a "separate stage" account, on the other hand, one can argue about whether the effect of notation should be understood as arising at an encoding stage or at a semantic stage; where the former can be seen as *peripheral* process and the latter as a *central* process. Noël and Seron (1992) subscribe to such a "separate stage" view (see also Dehaene & Cohen, 1995; McCloskey, 1992; McCloskey et al, 1992; Noël, Fias, Brysbaert, 1997; Sokol et al., 1991; for "separate stage" views) and proposed the "perceptual structure hypothesis" as an alternative explanation of Gonzalez and Kolers' findings. However plausible the perceptual structure hypothesis may be, merely postulating an alternative explanation is not sufficient to refute the procedural explanation. To refute Gonzalez and Kolers interpretation, Noël and Seron had to show that (1) an encoding stage exists, *and* (2) that notation only influences encoding. The data from Noël and Seron's (1992) seem to falsify at least the second assumption, leaving us with two options: either (1) a task independent encoding stage exists and notation affects both encoding and semantic number processing, or (2) notation affects "semantic" number processing. It should be clear that option 2 is much simpler than option 1. Why assume *more stages* and *more effects* than necessary?

In sum, the present re-analysis of Noël and Seron's study shows that one should not disregard Gonzalez and Kolers' study as irrelevant for the study of "semantic" number processing. On the contrary, most of the findings reported by them were replicated by Noël and Seron (1992) and can be understood as showing notation specificity in number processing. Further, the additional data provided by Noël and Seron (1992) show that, on the assumption that a task independent "encoding stage" exists, the effects of notation observed cannot be understood as merely due to differences in "encoding time". Therefore, the findings of both Gonzalez and Kolers (1982) and of Noël and Seron (1992) are more parsimoniously interpreted as showing notation specificity in numerical cognition proper.

References

Campbell, J. I. D. (1992). In defense of the encoding-complex approach: Reply to McCloskey, Macaruso, & Whetstone. In J. I. D. Campbell (Ed.), *The Nature and Origins of Mathematical Skills* (pp. 539-556). Amsterdam: Elsevier Science Publishers.

Campbell, J. I. D. (1994a). Architectures for numerical cognition. *Cognition, 53*, 1-44.

[5] It should be mentioned that Noël and Seron made an attempt to account for this position effect by assuming that the residual encoding effect would particularly occur for *complex* operands in the second position. Besides, the fact that this explanation is clearly ad hoc, I already argued that the "residual encoding" explanation itself is inherently inconsistent with a separate stage account, and hence is not a reasonable explanation for the position effect.

Campbell, J. I. D. (1994b). Numerical Cognition: Evidence for hyperspecific interactive operations. *Cahiers de Psychologie Cognitive, 13*(3), 297-320.

Campbell, J. I. D. (1997). Reading-based interference in cognitive arithmetic. *Canadian Journal of Experimental Psychology, 51*(1), 74-81.

Campbell, J. I. D. (1999). The surface form x problem size interaction in cognitive arithmetic: Evidence against an encoding locus. *Cognition, 70*, B25-B33.

Campbell, J. I. D., Kanz, C. L., & Xue, Q. (1999). Number processing in Chinese-English bilinguals. *Mathematical Cognition, 5*(1), 1-39.

Gonzalez, E. G., & Kolers, P. A.(1982). Mental manipulation of arithmetic symbols. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*(4), 308-319.

Kolers, P. A., & Roediger III, H. L. (1984). Procedures of mind. *Journal of Verbal Learning and Verbal Behavior, 23*, 425-449.

Kolers, P. A., & Smythe, W. M. (1979). Images, symbols and skills. *Canadian Journal of Psychology, 33*(3), 158-184.

Kolers, P. A., & Smythe, W. M. (1984). Symbol manipulation: Alternatives to the computational view of mind. *Journal of Verbal Learning and Verbal Behavior, 23*, 289-314.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44*, 107-157.

McCloskey, M, Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition, 4*, 171-196.

McCloskey, M., Harley, W., & Sokol, S. M. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(3), 377-397.

McCloskey, M., & Macaruso, P. (1995). Representing and using numerical information. *American Psychologist, 50*(5), 351-363.

McCloskey, M., Macaruso, P., & Whetstone, T. (1992). The functional architecture of numerical processing mechanisms: Defending the modular model. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 493-537). Amsterdam: Elsevier Science Publishers.

Noël, M.-P., Frias, W., & Brysbaert, M. (1997). About the influence of the presentation format on arithmetical-fact retrieval processes. *Cognition, 63*, 335-374.

Noël, M.-P., & Seron, X. (1992). Notational constraints and number processing: A reappraisal of the Gonzales and Kolers (1982) study. *The Quarterly Journal of Experimental Psychology, 45A*(3), 451-478.

Roediger III, H. L., Weldon, M. S., & Challis, B. H. (1989). Between implicit and explicit measures of retention: A processing account. In H. L. Roediger III, & F. I. M. Craik (Eds.), *Varieties of memory and consciousness* (pp. 3-41). Hillsdale, NJ: Erlbaum Associates.

Sokol, S. M., McCloskey, M., Cohen., N. J., & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(3), 355-376.

Tulving, E. (1984). Précis of elements of episodic memory. *Journal of Behavioral and Brain Sciences, 7*, 223-268.

Watkins, M. J. (1990). Mediationism and the obfuscation of memory. *American Psychologist, 45*(3), 328-335.