SWITCHING INSTRUCTION IN SYMBOLIC COMPARISON

A thesis submitted to
the Faculty of Graduate Studies and Research
in Partial Fulfillment of the requirements for the degree

Masters of Arts

by

Ian Mongeon

Department of Psychology
Carleton University

May 2009

©2009 Ian Mongeon

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-52024-6
Our file Notre référence
ISBN: 978-0-494-52024-6

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Previous symbolic comparison research has indicated that increased memory demands associated with accessing the comparative instructions can enhance the size of the semantic congruity effect (SCE). In the first experiment, that work was extended by examining whether the size of the SCE differs for trials in which the instruction had been switched or repeated in a predictable from the previous trial. Moreover, in a second experiment, the effect of both expected and unexpected instruction switches and repeats on the SCE was examined. In general, no evidence was obtained to indicate that the process of switching instructions leads to enhancements of the SCE. It is suggested, though, that such a result might be expected if participants are either actively inhibiting the irrelevant instruction or potentially unable to reconfigure the relevant one in a timely fashion.

I would like to take this time and thank all the people who have been an influential part of my life while I was in the M.A program in Psychology at Carleton University. Craig Leth-Steensen has always been a superior advisor and a great teacher throughout my academic life here at Carleton. I really appreciate all he has given me, whether it be his knowledge, motivation or simply his patience. Bill Petrusic, I would also like to thank you for your remarkable teachings and all your support through these years. Finally I would like to thank my parents for all of their support and patience they have provided me from day one. Anybody else who has directed me in the direction I am presently heading, I thank you as well.

Table of Contents

Introduction	1
<i>Theories of the Semantic Congruity Effect</i>	2
<i>Manipulating the Size of the Semantic Congruity Effect</i>	7
<i>Task Switching Theories</i>	12
Experiment 1.....	16
Method	16
<i>Participants</i>	17
<i>Stimuli and Apparatus</i>	17
<i>Procedure</i>	18
Results	21
<i>Response Times</i>	21
<i>Accuracy</i>	25
Experiment 2.....	27
Method	27
<i>Participants</i>	27
<i>Stimuli and Apparatus</i>	28
<i>Procedure</i>	28
Results	30
<i>Response Times</i>	30
<i>Accuracy</i>	32

General Discussion.....	33
<i>Inhibition Account</i>	35
<i>Faulty Reconfiguration Account</i>	36
Conclusion.....	42
References.....	46

List of Figures

Figure 1A. Split 1 response time (RT) cell means for Experiment 1a.....	50
Figure 1B. Split 1 response time (RT) cell means for Experiment 1b.....	51
Figure 2A. Split 2 response time (RT) cell means for Experiment 1a.....	52
Figure 2B. Split 2 response time (RT) cell means for Experiment 1b.....	53
Figure 3A. Split 1 proportion correct (PC) cell means for Experiment 1a.....	54
Figure 3B. Split 1 proportion correct (PC) cell means for Experiment 1b.....	55
Figure 4A. Split 2 proportion correct (PC) cell means for Experiment 1a.....	56
Figure 4B. Split 2 proportion correct (PC) cell means for Experiment 1b.....	57
Figure 5. Response time (RT) cell means for the Split 1 end-term pairs in Experiment 2.....	58
Figure 6. Proportion correct (PC) cell means for the Split 1 end-term pairs in Experiment 2.....	59
Figure 7A. Split 1 response time standard deviations (RTSD) for Experiment 1a.....	60
Figure 7B. Split 1 response time standard deviations (RTSD) for Experiment 1b.....	61
Figure 8A. Split 2 response time standard deviations (RTSD) for Experiment 1a.....	62
Figure 8B. Split 2 response time standard deviations (RTSD) for Experiment 1b.....	63

List of Appendices

Appendix A. Anticipated Results for Experiment 1a, Experiment 1b and Experiment 2.....	64
Appendix B. Response Time Results for the Eight Participants in Experiment 2.....	66

How is magnitude information about the sizes of things represented in semantic memory and what are the processes involved in retrieving and using it in everyday life? According to Moyer (1973), such information is represented in perceptual analogue form and processed in the same manner as the corresponding perceptual information (i.e., in terms of an "internal psychophysics"). According to Banks (1977), such information is accessed in categorical form and processed propositionally (i.e., in terms of "semantic codes").

One empirical response-time phenomenon that has historically been deemed to hold great promise with respect to understanding these issues is the *semantic congruity effect* (SCE). This effect is invariably observed whenever people are asked to compare the relative sizes of things in semantic memory (e.g., which of two named animals is the smaller). It is characterized by faster responding whenever the sizes of the things being compared are congruent with the direction of the comparison, as in deciding which of two small things is the smallest. Its presence was originally deemed to provide strong support for Banks' (1977) view that such comparisons are being made by generating and comparing categorizational-based semantic codes. This support has weakened somewhat, though, since the time of Petrusic and Baranski's (1989) original demonstration that SCEs can be obtained for comparisons of purely perceptual stimuli such as the lengths of lines (as long as there is some level of difficulty involved).

One recent approach to understanding the basis for the SCE that has been taken by Shaki and colleagues (Shaki & Algom, 2002; Shaki, Leth-Steensen, & Petrusic, 2006; Petrusic, Shaki, & Leth-Steensen, 2008; Leth-Steensen, Petrusic, &

Shaki, in preparation) has been to determine the conditions that moderate the size of this effect. In that work, a number of experimental manipulations that lead to enhancements in the size of the SCE have been revealed and, as such, these findings have been extremely informative theoretically. Although those results and their implications will be discussed more fully in the following, the gist of the conclusions arising from them is that any manipulation that serves to enhance the difficulty of keeping track of the required direction of the comparative (e.g., to choose the smaller or to choose the larger) on any particular trial also serves to enhance the SCE. The present work will continue in this vein by examining the size of the SCE when the direction of the comparative has either been switched from or stayed the same as that on the previous comparison trial. As such, the full theoretical and empirical framework of the (currently extremely popular) task-switching paradigm can be brought to bear upon the problem of understanding the SCE.

Theories of the SCE

Semantic Coding View. A highly influential theory of the SCE is that it is based on a categorizational (or semantic) coding of the magnitudes of the stimulus items (Banks, 1977). One ramification of such coding is that comparison responses can be made fairly quickly if the items are oppositely coded (e.g., as small and large, say; Banks, 1977; Cech & Shoben, 2001). Two other ramifications are that if the items are similarly coded more time is required to discriminate between their magnitudes and, of most relevance to the SCE, that such codes must also be translated to their corresponding opposites if they do not match the semantic code associated with the direction of the comparative instruction (Banks, 1977; Cech & Shoben,

2001).

Cech and Shoben have been instrumental in marshalling a body of evidence for the semantic coding view of the SCE (Cech & Shoben, 1985, 2001; Cech, Shoben, & Love, 1990). The main thrust of their approach has been to show that manipulations involving the range of the items can influence the way that the items become categorized, thus changing the form of the SCE. For example, Cech and Shoben (1985) demonstrated that if only a set of rather small animals are presented for categorization within an experimental session, animals at the upper range of this set that would normally be coded as small (e.g., rabbit or beaver) now become coded as large and, hence, are responded to faster when choosing the larger than when choosing the smaller. Moreover, Cech et al. (1990) then demonstrated that if two such ranges of items are present within an experimental session (i.e., five small animals and five large animals that are paired for comparison within each range only), dual SCEs arise one for each range of items. Finally, Cech and Shoben (2001) then also extended these findings to the infinite-set case in which no items are ever repeated within an experimental session. Cech and Shoben (2001) argue that such findings are highly supportive of the notion that participants are re-categorizing the items on the fly in order to maximize the number of items that are oppositely coded within an experimental session (i.e., in line with what they refer to as an efficiency principle) and go on to explore the mechanisms responsible for such recategorization (such as sensitivity to the available range or co-occurrence information provided within the experimental session).

Reference Point View. Another influential view of the SCE is that

comparisons are always made with respect to an ideal point of reference at either end of the continuum under consideration, where the particular reference point used on any particular comparison trial is the one specified by the comparative instruction. Such a notion was first introduced by Jamieson and Petrusic (1975) and later formalized in terms of both a specific cognitive model by Holyoak (1978) and a specific mathematical model by Deheane (1989). The essence of the reference point view is that item magnitudes are represented in memory in an analog continuum fashion and that the comparison process involves comparing the ratio of the distances of each of the stimulus magnitudes from the reference point. The SCE arises because the value of this ratio is larger the closer the items are to the reference point, which enhances the discriminability between their magnitudes, and reduces the time needed to compare them.

Evidence Accrual View. Petrusic (1992) has furthered the view that the congruity between the sizes of the stimuli in a comparison pair and the direction of the comparison affects the dynamics of an evidence accrual decision process. The particular model considered by Petrusic (1992; his slow-and-fast-guessing model) assumes that evidence favouring one response or the other is gradually accumulated in discrete samples until a criterion (or threshold) amount of evidence has been reached for one of the responses (or for neither response, at which point a guess response is emitted). In support of this view, Petrusic presented experimental evidence that the SCE occurs at the level of each pass through the evidence accrual process with slower accruals for comparisons of semantically incongruent pairs. Most notably, he showed that the properties of the SCE mirror the properties of the

underlying correct and error response times under conditions in which speed and accuracy have been differentially emphasized (e.g., errors are typically slower than correct responses under accuracy emphasis and also show more of an SCE). As discussed by Petrusic (1992), while it is possible to cast this view in terms of degradations in discriminability for stimuli that are further from the activated reference point, it is not specifically necessary to invoke such a notion.

Semantic Interference View. Within the context of their fully developed, connectionist, evidence accumulation model of symbolic comparison, Leth-Steensen and Marley (2000) have assumed that the SCE arises due to semantic (Stroop-like) interference between information pertaining to the relative sizes of the stimulus items in a pair (as indexed by how close they are to the end points of the representational continuum) and the direction of the comparative instruction (see also Banks & Root, 1979). Their model assumes the existence of two different sets of opposite instructional pathways for the processing of information under each of the two possible directions specified by the form of the comparative instruction. Because the activation of the relevant set of pathways on any comparison trial is assumed to be imperfect, these pathways are assumed to always be in competition to some extent and, moreover, that the degree of competition (or interference) is modulated by the relative sizes of the items being compared. This state of affairs serves to enhance the competition between the instructional pathways and slow responding when the relative sizes of the pairs of items and the direction of the comparison are incongruent and, conversely, to attenuate this competition and speed responding for semantically congruent items and comparative instructions. One form of support for this view was

provided by the excellent fit their model provided to the full set of symbolic comparison, response time data obtained by Leth-Steensen and Marley (2000) for pairwise comparisons of the items from a six-term, learned linear ordering. This result indicated that, computationally, the semantic interference assumption worked very nicely in unison with all of the other processing assumptions of their model, allowing it to capture all of the (sometimes very subtle) interactions between the SCE and the other two major symbolic comparison effects (namely, the end and distance effects) that were present in their data. As will be discussed, some further empirical support for this view of the SCE has been provided by Leth-Steensen et al. (in preparation), Petrusic et al. (2008), and Shaki et al. (2006).

Differential Bias View. Another fairly influential view of the SCE that also has its basis within an evidence accumulation model framework is one that posits that a random walk decision process can be biased depending on the congruity between the relative sizes of the items being compared and the direction of the comparison (Birnbaum & Jou, 1990; Link, 1990). More specifically, it is assumed that an initial assessment of the size of the items allows for a probabilistic assessment of the likely correct response (in the context of the direction of the comparative instruction used on that trial), which then motivates a strategic movement of the starting point of the random walk process towards the corresponding response boundary (note that random walks assume that sampled evidence is accumulated for one response and against the other and, hence, that the decision process involves a drift of an evidence accumulation signal between two absorbing response boundaries). This movement reduces the distance between the starting point and the eventual correct decision

boundary reducing response times accordingly. Although, this bias is generally assumed to be stronger for stimuli nearer to each end of the continuum under consideration, it is also assumed that it is even stronger for semantically congruent stimuli. The major bit of support for this view (other than its intuitive appeal) has been the excellent fit obtained by Birnbaum and Jou (1990) to a set symbolic comparison data they collected using a mathematical model which emulated these differential bias assumptions.

Manipulating the Size of the SCE

One recent approach to distinguishing between the various currently popular views of the SCE has been to determine the kinds of experimental manipulations that affect the size of this effect or, in other words, to determine the factors that interact with the SCE. With respect to this kind of endeavour, note that because the SCE itself is defined in terms of a two-way interaction between the overall size of the items in a pair and the type of comparative instruction used, this approach necessarily involves a search for three-way interactions involving the manipulated factor, pair size, and instruction type.

Manipulating semantic encoding. Shaki and Algom (2002; see also some associated earlier work by Cech, 1995) performed a study that used picture-word compounds as comparison stimuli (e.g., the word "cat" embedded within a picture of an elephant, where the actual physical sizes of the pictures themselves were held constant). In one session of trials in their Experiment 1, participants were asked to compare the sizes of the animals referred to by the words while ignoring the pictures, whereas in the other session, they were asked to make comparisons between the the

sizes of the animals referred to by the pictures. As expected, Stroop-like effects occurred for comparisons based on the word names in that response times were slower, by about 70 ms, for Stroop-incongruent pairs containing at least one non-matching, picture-word combination than for Stroop-congruent pairs containing matching, picture-word combinations for both items (no such effect, however, was found for the picture comparisons, which were about 200 ms faster overall). The major finding from this experiment, which was then replicated with an infinite-set stimulus design in their Experiment 2, was that the SCE for the word-name comparisons did not interact with Stroop congruency (i.e., its size did not depend on whether the pairs being compared were Stroop congruent or incongruent).

Because the Stroop congruency manipulation was assumed to degrade the semantic encoding of the stimuli (specifically with respect to identifying them and retrieving their associated magnitude information) and, hence, could be regarded as a form of semantic interference, Shaki and Algom (2002) had hypothesized that if the SCE is also due to stimulus-driven semantic interference, as in the Leth-Steensen and Marley (2000; Banks & Root, 1979) model, then these two effects "ought to coalesce" (Shaki & Algom, 2002, p. 4) and interact with each other. Hence, the (statistical) additivity of these two effects that was observed led Shaki and Algom (2002) to conclude that the locus of the semantic congruity cannot be at the (semantic) encoding stage and, further, that any view that ascribes a role for conflicting semantic information in the SCE could not be supported by their results.

Manipulating access to the comparative instruction. As mentioned, though, some subsequent work by Shaki et al. (2006), Petrusic et al. (2008), and Leth-

Steensen et al. (in preparation) has served to provide support for both the semantic interference view of Leth-Steensen and Marley (2000) and the evidence accrual view of Petrusic (1992). The general approach used in all three of these studies was to examine manipulations which could be assumed to interfere with either the maintenance or the retrieval (or both) of the relevant comparative instructional context. For example, Shaki et al. (2006) found that the size of the SCE was enhanced for blocks of trials of both animal-size and line-length comparisons in which the two forms of the comparative instructions were randomly presented within a block (i.e., mixed blocks) in comparison to blocks of trials in which only one of the two instructions was presented throughout (i.e., pure blocks). They argued that this result follows naturally from the Leth-Steensen and Marley (2000) view because randomly varying the instructions within blocks would serve to increase the level of competition (or interference) between the instructional pathways in that model, which the model would then predict should lead to an enhancement of the SCE. However, they also concluded that this result is consistent with Petrusic's (1992) view if it is assumed that the enhanced memory demands associated with the mixed instructional blocks serves to further slow the rate of each evidence accrual event of the comparative decision process.

Next, Petrusic et al. (2008) found that the SCE was enhanced for conditions in which the comparative instructions were not perceptually available but needed to be retrieved from memory. More specifically, their participants first learned to associate each of the two instructional words (i.e., "Larger" and "Smaller") with four of eight consonant-vowel-consonant (CVC) nonsense syllables (e.g., "GUF", "BIX", etc.).

During the subsequent animal-size or line-length comparison phase in each of two respective experiments, either the standard instruction word or one of the CVCs was presented as the comparative instructions. In the latter condition, participants had to retrieve the instructional association in order to perform the comparison, and this requirement then served to increase the size of the SCE. These authors argued that such a finding was consistent with the notion that the contextual information provided by the instructions was less salient in the CVC-instructions condition which could then be assumed to decrease the strength of the relevant instructional pathway in the Leth-Steensen and Marley (2000) model, resulting in enhanced pathway competition (or interference) and a larger SCE. Again, this result was also deemed to be consistent with Petrusic's (1992) view because slowing memory access to the instructions would then almost certainly also slow down the evidence accrual process affecting semantically incongruent comparisons more than semantically congruent ones. Interestingly, the fact that the instructional manipulation seemed to be having a rather dramatic effect on semantically incongruent trials but, in most cases, only a very minimal effect on semantically congruent ones, also led these researchers to propose that it is possible that the SCE arises because the size of the pair affects the speed of instructional access given the instruction word. If so, further degrading instructional access by using CVC-instructions would then serve to greatly affect semantically incongruent comparisons where instructional access is impeded to begin with.

Finally, in Leth-Steensen et al.'s (in preparation) study, the direction of the comparison to use on each trial was made contingent on the size of the items being compared. Namely, on some trials participants were asked to choose the smaller

animal (in Experiment 1 or the shorter line length in Experiment 2) if the pair contained the names of small animals but on other trials to choose the larger animal if the animals in the pair were small (these were the semantically congruent and incongruent cases, respectively, where analogous cases were derived for pairs of large animals). Once again, the SCE was much larger for category-contingent comparison trials than for trials in which the comparative instructions had been presented in the usual fashion. Again, this effect was attributed by these researchers to either to enhanced instructional pathway competition or to slower evidence accruals when the instructional access was degraded by making them contingent on an initial categorization of the size of the stimulus items.

The present study hopes to continue in this same vein by examining how the size of the SCE is affected by conditions in which the direction of the required comparison has either been switched from the previous trial or has not been switched. To discern the relevance of this approach, it is important to note that each form of the comparative instruction in a symbolic comparison experiment essentially serves to cue a different comparison task set. The task sets signaled by each of the two forms of the comparative instructional cues are surely related but also differ in the sense that it is essential to be in the correct set in order to respond correctly. In this light, the act of switching between instructions across comparison trials can be likened to switching tasks which is a phenomenon which has received quite a bit of empirical attention over the past 12 years. After a review of that work in the following, some specific hypotheses will be derived with respect to how the SCE might be expected to be affected by a switch or repeat of the form of the comparative instruction.

Task Switching Theories

The main phenomenon of interest within the task switching paradigm is the presence of switch costs, which reflect the finding that response times are generally slower on trials where the task to be performed has been switched from the previous trial in comparison to trials in which the same task has been repeated. It is now generally accepted that switch costs involve both explicit/active/voluntary/intentional/top-down and implicit/passive/involuntary/unintentional/bottom-up control-based processing components.

Task-set Reconfiguration. One very influential view concerning the explicit component is that switch costs represent the time needed to reconfigure the system to perform the new task (Rogers & Monsell, 1995; such a view has also specifically been implemented in the formal production-system, task-switching model of Sohn & Anderson, 2001). Such reconfiguration would be assumed to involve activating the appropriate set of stimulus-response rules and enabling the required sensory and motor systems. One finding which has been deemed to be highly consistent with such a view is the fact that switch costs typically become reduced whenever more preparatory time between trials is made available. For example, switch costs were observed to drop (but not completely disappear) as Rogers and Monsell (1995; Experiment 3) increased the response-stimulus interval (RSI) between each response and the presentation of the stimulus for the next trial from 150 to 1200 ms across blocks of task-switching trials. One very important aspect of Rogers and Monsell's (1995) study is that they also came up with a methodology for studying switch costs

which avoided some of the problems with previously used task switching paradigms, such as amount of workload in pure and switch blocks, by presenting both task switches and repeats within the same blocks by alternating the task every two trials (i.e., their AABBAABB alternating-runs procedure). Importantly, this procedure allows participants to predict the upcoming task on the next trial so that they can make full use of any preparation time to either actively reconfigure the task set on task-switch trials or to actively maintain the commitment to the current task set on task-repeat trials.

Task-set Inertia. One very influential view concerning the implicit component is that switch costs reflect the need to overcome the persisting activation of the task set from the preceding trial (Allport, Styles, & Hsieh, 1994; Allport & Wylie, 2000). Such carryover effects of the previous task set are assumed both to compete with the instantiation of the now-relevant task set on task-switch trials and to facilitate the maintenance of the still-relevant task set on task-repeat trials. Although fervent proponents of the task-set inertia view do not deny that intentional, goal-driven control operations must be involved in the instantiation of task sets, they regard the task-switch process itself as involving a more gradual activation buildup and decay of the current and previous task sets, respectively. One very important paradigm that has allowed for an examination of the separate effects of the explicit and implicit components of the switch cost was developed by Meiran (1996). In this paradigm, the task to be performed on each trial is indicated by the presentation of task cues and, hence, both the length of the time interval between the previous response and the presentation of the task cue (i.e., the amount of time available for the activation of the

previous task set to passively dissipate) as well as the interval between the presentation of the task cue and the presentation of the upcoming stimulus (i.e., the amount of time available to actively prepare for the next task) can be separately manipulated. This research has indicated that increasing the length of each type of interval leads to reductions in switch costs but more so for the increases in the latter preparatory-based interval. Nonetheless, because of these (and other) findings, even fervent proponents of the task-set reconfiguration view acknowledge that switch costs (especially with respect to the residual switch costs that remain even after quite extensive preparation time has been provided) must to some extent be due to the workings of task-set inertia (and, hence, it is generally felt that a role for it should be included in any model of task switching; e.g., Sohn & Anderson, 2001).

Task-set Inhibition. Still another view of task-switching is one which postulates a major role for inhibitory processing (Mayr & Keele, 2000). This view arises out of the notion that the maintenance of high-level control settings (like that used to determine a task set) likely requires sustained activation, which then makes passive dissipation of that activation over the short term seem rather unlikely, and suggests that active inhibition is likely needed in order to reduce the influence of just-previously-active task sets. Mayr and Keele (2000) suggested that at least part of the switch cost arises as a result of having to overcome this (what they called) backward inhibition. The major empirical support for the presence of such inhibition is the finding that when three tasks (e.g., Task A, B, and C) are used in a task switching study, it is harder to switch back to a task that has been performed, and hence has been inhibited, more recently (i.e., responses to the final task in an ABA task

sequence are slower than for a CBA task sequence; Mayr & Keele, 2000). Note that such a view is not actually inconsistent with either the reconfiguration or inertia views. With respect to the former view, it seems quite reasonable to consider that one aspect of reconfiguring a task set would involve attempts to disable the previous one. With respect to the latter view, one important source of evidence for the task-set inertia notion has been deemed to be the finding that it is harder to switch to the stronger (or more dominant) of two tasks than to the weaker task (e.g., switch costs are larger for word naming than for color naming in the Stroop paradigm, Allport & Wylie, 2000). The typical account for this phenomenon is that the control biases that serve to either increase the level of activation of the weaker task or suppress (i.e., inhibit) the activation of the stronger task (or both), are stronger whenever the weaker task needs to be performed in the face of competition from the stronger task. These control biases are then assumed to carryover into subsequent trials making it harder to reinstantiate the stronger task.

Experiments 1a and 1b

The review of the task switching paradigm just undertaken indicates that the switching of task sets involves the use of control operations that serve to reconfigure the system for the new task. Another function of these control operations could be to minimize carryover effects of the activation of the previous task set through inhibition. With respect to symbolic comparisons, it is of interest to determine the extent to which both the control operations and the task-set-inertia effects that occur on switch trials (where, here, the two task sets refer to each possible direction of the comparison) affect the size of the SCE.

More specifically, access to the relevant comparative instructional context would certainly be expected to be degraded on trials in which the instructional context has been switched from that on the previous trial and, hence, is in the process of being reconfigured. Under such conditions, the Petrusic view of the SCE would likely prescribe that SCEs should be larger. Similarly, carryover of the previous instructional task set on task-switch trials could also be assumed to interfere (or compete) with the process of comparing of the stimulus items on the basis of the current instructional task set. As discussed previously, the Leth-Steensen and Marley view of the SCE would then likely predict that such interference would lead to larger SCEs on task-switch than on task-repeat trials (or analogously, carryover effects on task-repeat trials would be assumed to induce a repetition benefit for that instruction which would then attenuate the degree of instructional interference on those trials leading to smaller SCEs). However, such effects could also be assumed to rely on the amount of available preparation time and might only become manifest for very short RSIs.

To test these notions, in the following experiment the direction of the comparison is predictably switched within blocks of symbolic comparison trials (according to Rogers and Monsell's, 1995, alternating runs procedure) and the effect of switching instructions on the size of the SCE is examined. Moreover, in half of the blocks the RSI used is very short (i.e., 200 ms between each response and the presentation of the next stimulus pair) and in the other half it is comparatively much longer (i.e., 2000 ms). The stimuli used are names from a six-term linear ordering that are to be learned in an initial learning phase (i.e., in the same manner as in Leth-

Steensen & Marley, 2000).

Method

Participants. Respectively, 43 and 36 introductory psychology participants at Carleton University were recruited to participate in Experiments 1a and 1b for extra course credit. Each participant performed the task individually, and each session took approximately 90 min.

Stimuli and Apparatus. Both experiments were programmed and run using MEL v2.0. During each trial in the initial learning phase, the instruction for comparison (i.e., "Taller?" or "Shorter?") was first presented at the top of the screen, after which participants were presented with a pair of three-letter names, presented side by side in the centre of the computer screen (in MEL System48 font size). During the subsequent test phase, no comparative instructions were explicitly presented and the presentation of the pairs of names were rotated through the four quadrants of the computer screen (but otherwise presented in the same fashion as in the learning phase). In both the learning and the test phases, responses were made by pressing either the "z" or "/" response keys on each side of the bottom row of the computer keyboard (marked in yellow).

Each name was a label that stood for an imaginary "person". Six "people" were used in these experiments, all of whom were assumed to differ in height. Four different orderings of these names were derived a priori. From tallest to shortest, the four orderings were (a) Pat, Bob, Ted, Dan, Jim, and Mel, (b) Dan, Ted, Jim, Pat, Mel, and Bob, (c) Ted, Dan, Mel, Bob, Pat, and Jim, and (d) Bob, Mel, Dan, Jim, Ted, and Pat. These orderings were assigned randomly to participants.

Procedure. The first part of each experiment was a learning phase, in which the participants were presented with each of the five comparison pairs consisting of the stimuli that were adjacent to each other in the relevant ordering (i.e., the Split 1 pairs). Each comparison trial began with a blank screen for 1000 ms. Before a pair of names appeared, the relevant comparative instruction ("Taller?" or "Shorter?") was displayed on the computer screen for 1000 ms just above a plus sign, which acted as a temporary fixation point for the location of the name stimuli. On each comparison trial, the participants were presented with two of the six names and had to choose which person (i.e., the name presented on the left or the name presented on the right) was either the shorter or, respectively, the taller of the pair. They were asked to use the index fingers on their left and right hands to make their responses.

The participants were not expected to know the relative heights of the stimuli at the very start of the experiment and, hence, would initially be guessing. After each response in the learning phase, though, feedback was immediately provided on the computer screen that indicated whether the response had been correct or incorrect and also showed the correct name for that comparison trial. This feedback helped the participants eventually learn which person was the taller and which was the shorter in each pair. They were able to examine this feedback for as long as they wanted and could initiate the next trial with a press of the space bar. Both the comparative instruction and the pair of names remained on the screen throughout the trial (including during the presentation of the feedback).

Learning trials were presented in blocks of 20 trials (i.e., five pairs with each of the two instructions in each of the two left-right spatial presentation of the pairs)

with breaks after the first two learning blocks only. In the first two blocks, the pairs were presented in an ordered fashion and then presented randomly within the remaining blocks. The learning phase took anywhere from 15 to 35 minutes to perform and was completed when the participant had been correct for a full block of 20 learning trials. Participants who did not reach this learning criterion within 10 blocks of learning trials were given a shorter version of the experiment by the experimental program, and their results were not included in the following analysis.

The second part of each experiment was the test phase and the procedure for it was similar to that of the learning phase except for the following differences. First, all of the 15 possible pairs of names were now included in the stimulus set and no feedback was provided. Second, the pairs of names were now presented in the center of one of the four quadrants of the computer screen (i.e., top left, top right, bottom right, and bottom left) and rotated systematically through those four quadrants from trial to trial (in the order just specified). Third, no actual comparative instructions were presented in the test phase and the participants were instructed that the direction of the comparison to be made on each trial was to be specified by the location of the pair on the screen. Namely, in Experiment 1a they were asked to choose the taller member of the pair when the names had been presented in either the top-left or top-right quadrants but the shorter member of the pair when the names had been presented in the bottom-right and bottom-left quadrants. In Experiment 1b, however, they were asked to choose the shorter member of the pair when the names had been presented in either the top-left or top-right quadrants and the taller member of the pair when the names had been presented in the bottom-right and bottom-left quadrants.

(this was actually the only difference between the two experiments). For both experiments, a card with the instruction-to-quadrant mappings specified on it was also placed in front of them throughout.

Fourth, within each block a blank screen was presented during the interval between the response to each trial and the presentation of the stimulus pair for the next trial (i.e., during the RSI). For half of the blocks of trials, the duration of this blank screen was 200 ms (i.e., the short-RSI condition) and for the other half of the trials its duration was 2000 ms (i.e., the long-RSI condition). The order of performance of these two RSI conditions was counterbalanced across participants. Each RSI condition contained four blocks of 60 comparison trials (made up of the 15 pairs presented twice in each left-right stimulus presentation order) preceded by a short block of 12 practice trials. The stimulus pairs used on each trial within a block were randomly chosen (without replacement) from the set of 60 possible trials and the 12 trials used in the practice phase were randomly selected from that same set. The participants were informed at the end of the practice phase and were given an optional rest period between each block. Finally, participants were asked to be accurate with each decision without taking too much time to respond.

Results

Of the 43 participants in Experiment 1a, 11 failed to reach the aforementioned learning criterion within 10 blocks of learning trials and were excluded from the rest of the analyses. The mean number of blocks taken to reach the learning criterion by the remaining 32 participants was 6.53 (SD = 1.92). In addition, two of these participants were subsequently dropped from the Experiment 1a analyses because

their overall accuracy levels in the test phase were very low (i.e., 63 and 64%, respectively). Of the 36 participants in Experiment 1b, 10 failed to reach the aforementioned learning criterion within 10 blocks of learning trials and were excluded from the rest of the analysis. The mean number of blocks taken to reach the learning criterion by the remaining 26 participants was 5.31 (SD = 1.23). In addition, two of these participants were subsequently dropped from the Experiment 1b analyses, one because of a very low overall accuracy level in the test phase (i.e., 65%) and the other because of a very slow mean RT in the test phase (i.e., 3089 ms, which was more than 3 SDs from the mean RT for the remaining 26 participants of 1808 ms).

In the following Experiment 1a and 1b analyses, only responses for non-practice trials in the test phase were used (also excluding, however, the first trial in each block because it could not technically be designated as either a repeat or switch trial). The dependent variables were the median correct RTs and the arc-sine transformed proportion correct (PC; although note that all PC values reported in the text and figures are the corresponding untransformed ones) for each individual participant in each cell of the following designs. All missing RT and PC values and also any RT values that were more than 3 SDs above the mean of their cell were replaced by the corresponding cell mean + participant mean – grand mean. All results that were significant at the .05 (Greenhouse-Geisser adjusted) level of significance are reported.

Response times. SCEs were examined in separate analyses for the Split 1 and Split 2 pairs, respectively. The independent variables for each analysis were RSI (200

ms, 2000 ms), instruction transition (repeat, switch), instruction type (shorter, taller), and pair. Plots of the Split 1 RT data (averaged over participants in each cell) obtained in Experiments 1a and 1b are given in Figures 1a and 1b, respectively. Plots of the averaged Split 2 RT data from these two experiments are given in Figures 2a and 2b, respectively.

In the Split 1 ANOVA results for Experiment 1a, the main effect of instruction transition was significant, $F(1, 29) = 19.542, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to slower on average (2114 ms) than were trials where the instruction had been repeated (1902 ms). There was also a significant main effect of pair, $F(4, 116) = 31.141, p < .001$, where the overall mean RTs for each of the increasing pairs (1, 2), (2, 3), (3, 4), (4, 5), and (5, 6) were 1604, 2135, 2542, 2300, and 1479 ms, respectively. The presence of an overall SCE was indicated by a significant interaction between instruction type and pair, $F(4, 116) = 11.269, p < .001$. Although the overall three-way interaction of Instruction transition $\times$ Instruction type $\times$ Pair was not significant, the linear trend of this three-way interaction (which essentially characterizes the SCE) was significant, $F(1, 29) = 6.680, p < .015$. As is evident in Figure 1a, the SCE is actually smaller for trials in which the instruction had been switched than for trials in which the instruction had been repeated.

In the Split 1 ANOVA results for Experiment 1b, the main effect of instruction transition was significant, $F(1, 23) = 22.793, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to more slowly (1870 ms) than were trials where the instruction had been repeated (1684

ms). There was also a significant main effect of pair, $F(4, 92) = 22.076, p < .001$, where the overall mean RTs for each of the increasing pairs (1, 2), (2, 3), (3, 4), (4, 5), and (5, 6) were 1411, 1971, 2295, 1913, and 1294 ms, respectively. Moreover, the main effect of instruction type was also now significant, $F(1, 23) = 5.532, p < .028$, such that responding with the (marked) shorter instruction was slower than with the (unmarked) taller instruction (1822 vs. 1732 ms, respectively). The two-way interaction between RSI and instruction transition was also significant, $F(1, 23) = 10.853, p < .003$, corresponding to the fact that the difference in RT between instruction-switch and instruction-repeat trials was now much larger for the 200-ms RSI than for the 2000-ms RSI. The presence of an overall SCE was indicated by a significant interaction between instruction type and pair, $F(4, 92) = 14.988, p < .001$. In this analysis, although the overall four-way interaction of RSI $\times$ Instruction transition $\times$ Instruction type $\times$ Pair was not significant, the linear trend of this four-way interaction was significant, $F(1, 23) = 4.497, p < .045$. As is evident in Figure 1b, for the 200-ms RSI, the SCE seems to be somewhat smaller for instruction-switch than for instruction-repeat trials, whereas for the 2000-ms RSI, the SCE now seems to be somewhat larger for instruction-switch than for instruction-repeat trials.

In the Split 2 ANOVA results for Experiment 1a, the main effect of instruction transition was again significant, $F(1, 29) = 41.945, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to slower on average (1909 ms) than were trials where the instruction had been repeated (1656 ms). There was again also a significant main effect of pair, $F(3, 87) = 18.955, p < .001$, where the overall mean RTs for each of the increasing pairs

(1, 3), (2, 4), (3, 5), and (4, 6) were 1622, 2054, 1996, and 1457 ms, respectively. Moreover, the main effect of instruction type was also significant here, $F(1, 29) = 4.245, p < .048$, although it was responding with the (marked) shorter instruction that was now faster than was responding with the (unmarked) taller instruction (1752 vs. 1813 ms, respectively). The presence of an overall SCE was signaled by a significant interaction between instruction type and pair, $F(3, 87) = 8.750, p < .001$. Finally, the overall four-way interaction of RSI $\times$ Instruction transition $\times$ Instruction type $\times$ Pair was significant here, $F(3, 87) = 3.362, p < .025$, which was accompanied by a significant linear trend for this four-way interaction, $F(1, 29) = 8.679, p < .006$. As is evident in Figure 2a, for the 200-ms RSI, the SCE is almost non-existent for instruction-switch but clearly present for instruction-repeat trials, whereas for the 2000-ms RSI, the SCE now seems to be somewhat larger for instruction-switch than for instruction-repeat trials.

In the Split 2 ANOVA results for Experiment 1b, the main effect of instruction transition was again significant, $F(1, 23) = 31.714, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to slower (1645 ms) than were trials where the instruction had been repeated (1442 ms). There was also a significant main effect of pair, $F(3, 69) = 16.229, p < .001$, where the overall mean RTs for each of the increasing pairs (1, 3), (2, 4), (3, 5), and (4, 6) were 1380, 1745, 1772, and 1278 ms, respectively. Moreover, the main effect of instruction type was again also significant, $F(1, 23) = 9.607, p < .005$, such that responding with the (marked) shorter instruction was again slower than with the (unmarked) taller instruction (1612 vs. 1475 ms, respectively). The

presence of an overall SCE was indicated by a significant interaction between instruction type and pair, $F(3, 69) = 8.331, p < .001$. Finally, the overall three-way interaction of Instruction transition $\times$ Instruction type $\times$ Pair was significant here, $F(3, 69) = 4.793, p < .006$, which was accompanied by a significant linear trend for this three-way interaction, $F(1, 23) = 15.315, p < .001$. As is evident in Figure 2b, the SCE is smaller for trials in which the instruction had been switched than for trials in which the instruction had been repeated.

Accuracy. SCEs were again examined in separate analyses for the Split 1 and Split 2 pairs, respectively. The independent variables for each analysis were RSI (200 ms, 2000 ms), instruction transition (repeat, switch), instruction type (shorter, taller), and pair. Plots of the Split 1 PC data (averaged over participants in each cell) obtained in Experiments 1a and 1b are given in Figures 3a and 3b, respectively. Plots of the averaged Split 2 PC data from these two experiments are given in Figures 4a and 4b, respectively.

In the Split 1 ANOVA results for Experiment 1a, only the main effect of instruction transition was significant, $F(1, 29) = 13.493, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to less accurately on average (.928) than were trials where the instruction had been repeated (.952).

In the Split 1 ANOVA results for Experiment 1b, the main effect of instruction transition was significant, $F(1, 23) = 9.222, p < .006$, indicating that trials where the instruction had been switched from the previous trial were responded to less accurately (.919) than were trials where the instruction had been repeated (.937).

The two-way interaction between RSI and instruction transition was also significant, $F(1, 23) = 4.980, p < .036$, corresponding to the fact that the difference in accuracy between instruction-switch and instruction-repeat trials was much larger for the 200-ms RSI than for the 2000-ms RSI.

In the Split 2 ANOVA results for Experiment 1a, the main effect of instruction transition was again significant, $F(1, 29) = 21.222, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to less accurately on average (.937) than were trials where the instruction had been repeated (.966). Moreover, the main effect of RSI was also significant here, $F(1, 29) = 10.730, p < .003$, whereby responding was less accurate for the 200-ms than for the 2000-ms RSI (.940 vs. .962, respectively).

In the Split 2 ANOVA results for Experiment 1b, only the main effect of instruction transition was again significant, $F(1, 23) = 23.184, p < .001$, indicating that trials where the instruction had been switched from the previous trial were responded to less accurately (.940) than were trials where the instruction had been repeated (.973).

Experiment 2

Experiment 2 expands upon Experiment 1 by sometimes presenting unexpected instruction switches (i.e., TTSSTS) and unexpected instruction repetitions (i.e., TTSSS; Ruthruff, Remington, & Johnston, 2001). Such a manipulation allows for a further determination of the effect of explicit, instructional expectancies on the size of the SCE and also allows for a dissociation of such effects from the more implicit ones driven by task-set inertia. Namely, when the RSI is very short, any effects of switching or repeating the comparative instruction on the size of the SCE should due solely to instructional carryover effects and, hence, should be very similar for both expected and unexpected task-switch and task-repeat trials (i.e., larger SCEs for switch than repeat trials in both cases). On the other hand, when the RSI is rather long, participants should be in a fairly high state of preparation for the wrong instruction on both unexpected task-switch and task-repeat trials. Hence, an enhancement in the size of the SCE should indeed be observed for both unexpected task-switch and task-repeat trials with long RSIs. (Note that both Experiments 1 and 2 were ran simultaneously. Hence, the results for Experiment 1 were not known when developing the motivation for Experiment 2, which essentially follows the same logic used to derive the hypotheses for Experiment 1. The fact that the results of Experiment 1 were not actually consistent with those hypotheses will be discussed extensively in the Discussion as well as the degree to which the results of the following Experiment 2 served to corroborate the Experiment 1 findings.)

Participants. Ten other Carleton University students were recruited for pay to perform three 90-min sessions.

Materials and Apparatus. The materials and apparatus were the same as for Experiment 1 except for the fact that stimulus pairs were now always presented in the center of the screen throughout and the instructional task sets in the test phase were now signaled by the colour of the stimuli (i.e., either red or green).

Procedure. The procedure for the first part of the experiment (i.e., the learning phase) was identical to that used in Experiment 1. The procedure for the second part of the experiment (i.e., the test phase) differed from that used in Experiment 1 in the following ways. First, the direction of the comparison was now signalled by the colour of the stimulus pair. Namely, the participants were instructed to choose the taller member of the pair when the names had been presented in green but the shorter member when the names had been presented in the red (a card with these mappings specified on it was also placed in front of them throughout). Second, only the 9 Split 1 (i.e., adjacent) and Split 2 (i.e., separated by two) pairs were used instead of all 15 possible pairs. Hence, each block now contained 36 trials (9 pairs presented twice in each left-right stimulus presentation order).

Third, as in Experiment 1, the order of the tasks within the blocks still alternated in runs of two. However, every block now contained two unexpected switch or repeat trials (i.e., one quarter of the blocks contained two unexpected switches, one quarter contained two unexpected repeats, and the remaining half contained both one unexpected switch and one unexpected repeat). The placements of these unexpected switch and repeat trials within the blocks were determined non-randomly (i.e., hand-picked to make sure that they covered the space of possible placements across the blocks) for 24 different blocks of trials. As well, care was taken

to always keep the two-by-two alternation of the task sets going after these unexpected trials (e.g., TTSSSTTSS etc. and TTSTTSS etc. for unexpected repeat and switch sequences of trials, respectively). Moreover, although the pairs used on all of the expected trials were determined randomly, only the pairs (1, 2), (2, 1), (5, 6), and (6, 5) were used on unexpected trials (essentially by writing MEL code that replaced whatever pair was going to be randomly chosen for that trial with one of those 4 pairs in a manner such that each pair was equally represented across the unexpected repeat and switch trials). The reason for using only these pairs on unexpected trials was because they are the pairs which typically show the clearest SCE and, hence, it was important to make sure that at least they were well-sampled on the unexpected trials. Fourth, because of the paucity of unexpected trials within each block, it was necessary to have participants do more than one session. Hence, each participant did three sessions across different days. Day 1 involved the learning phase and 16 blocks of trials, 8 with the long 2000 ms RSI and 8 with the short 200 ms RSI. Day 2 involved doing the learning phase once more (starting at the third block of that phase where the pairs began to be randomized) followed by all 24 blocks of trials (manipulating the RSIs used in a long-short-short-long fashion 6 blocks at time). Day 3 also involved doing the learning phase again followed once more by all 24 blocks of trials but with the RSIs used on them switched from Day 2.

Results

Of the 10 participants in Experiment 2, Participants 1 and 4 did not wish to return after the first session. Of the 8 remaining participants, 2 did not actually meet the learning criterion within 10 blocks of learning trials in Session 1 but were still

included in the rest of the analyses (and still received the same amount of comparison trials in the testing phases as the others). The number of blocks taken to reach the learning criterion in Session 1 by the remaining 6 participants were 7, 4, 9, 4, 6, and 9 for Participants 2, 3, 7, 8, 9, and 10, respectively. These 6 participants also spent only about 1 – 3 blocks on the relearning phases on Days 2 and 3 (the other 2 participants spent 7 and 8 blocks, respectively, relearning on Day 2 but only 2 and 4 blocks, respectively, relearning on Day 3).

In the following Experiment 2 analyses, only responses for test phase trials on Days 2 and 3 were used (and note that no practice trials were given after Day 1). Also excluded from these analyses were the first trial in each block (because it could not technically be designated as either a repeat or switch trial) and also the first trial after an unexpected instruction trial (which itself might have involved some uncertainty about which instruction was to follow that unexpected instruction trial). The dependent variables are the median correct RTs and the arc-sine transformed PCs (although note that all PC values reported in the text and figures are the corresponding untransformed ones) for each individual participant in each cell of the following designs. No cell values were ever above 2.5 SDs of the cell means in any of the following analyses, and given the amount of data collected there were no missing trials. Given the potential lack of power that might result from having only eight participants, all results that were significant at the .10 (Greenhouse-Geisser adjusted) level of significance are reported.

Response times. SCEs for the Split 1 end-term pairs only were examined (because, as mentioned in the Method, these were the only pairs used on unexpected

instruction trials). The independent variables for this analysis were instruction expectancy (expected, unexpected), RSI (200 ms, 2000 ms), instruction transition (repeat, switch), instruction type (shorter, taller), and pair. Plots of the average RT data are given in Figure 5 (and corresponding plots for each of the eight participants are given in the Appendix). Note, as well, that a corresponding ANOVA that was ran with only the six participants who had learned in Session 1 included gave analogous results.

In these ANOVA results, the main effect of instruction transition was significant, $F(1, 7) = 7.650, p < .028$, indicating that trials where the instruction had been switched from the previous trial were responded to slower on average (1104 ms) than were trials where the instruction had been repeated (1048 ms). The two-way interaction between RSI and instruction transition was also significant, $F(1, 7) = 10.486, p < .014$, corresponding to the fact that a difference in RT between instruction-switch and instruction-repeat trials was present only for the 200-ms RSIs (i.e., 1139 vs. 1026 ms, respectively) and not for the 2000-ms RSI (i.e., 1068 vs. 1071 ms, respectively).

There was also a marginally significant main effect of pair, $F(1, 7) = 3.865, p < .090$, corresponding to the fact that, as is evident in Figure 5, the taller end-term pair was generally responded to faster than the shorter end-term pair (and note that both the $RSI \times Pair$ and the $Instruction\ Expectancy \times RSI \times Pair$ interaction were also marginally significant, $F(1, 7) = 3.896, p < .089$, and $F(1, 7) = 3.984, p < .086$, respectively, but also that all of these three pair effects were present only in the analysis that included all eight participants). The presence of an overall SCE was

indicated by a significant interaction between instruction type and pair, $F(1, 7) = 33.687, p < .001$. Finally, the only other significant effect was an Instruction Expectancy $\times$ Instruction Transition $\times$ Instruction Type three-way interaction, $F(1, 7) = 7.895, p < .026$.

Accuracy. Again, only the Split 1 end-term pairs only were examined in this analysis. The independent variables were again instruction expectancy (expected, unexpected), RSI (200 ms, 2000 ms), instruction transition (repeat, switch), instruction type (shorter, taller), and pair. Plots of the average PC data for the cells in this design are given in Figure 6 (but no individual participant plots are provided for the PC results). Note, as well, that a corresponding ANOVA that was run including only the six participants who had learned in Session 1 included gave analogous results.

In these ANOVA results, the main effect of instruction expectancy was significant, $F(1, 7) = 29.194, p < .001$, indicating that trials with unexpected instructions were responded to less accurately (.982) than were trials with expected instructions (.986; note that the arc-sine transformation would have served to exaggerate what looks to be a rather small difference in PC here). There was also a marginally significant main effect of pair, $F(1, 7) = 3.812, p < .092$, corresponding to the fact that, as is evident in Figure 6, the taller end-term pair was generally responded to more accurately than the shorter end-term pair (in addition, note that the RSI $\times$ Pair two-way interaction was significant, $F(1, 7) = 9.468, p < .018$, and also that both the Instruction Transition $\times$ Pair and the Instruction Expectancy $\times$ Instruction Transition $\times$ Pair interactions were marginally significant, $F(1, 7) = 4.719$,

$p < .066$, and $F(1, 7) = 4.007, p < .085$, respectively). The only other significant effect was the Instruction Expectancy $\times$ Instruction Transition $\times$ Instruction Type $\times$ Pair four-way interaction, $F(1, 7) = 27.309, p < .001$, which corresponds to the fact that whereas an SCE (collapsed across both RSIs) was present in accuracy in the expected instruction, task-repeat condition, a reverse SCE (collapsed across both RSIs) was present in the unexpected instruction, task-repeat condition (and note that this effect was still significant for the analysis with only the six participants who had initially learned).

General Discussion

The major goal of the present research was to extend the findings of Shaki et al. (2006), Petrusic et al. (2008), and Leth-Steensen et al. (in preparation) to the situation in which the comparative instructions were either repeated or switched in a predictable manner. In that previous work, experimental manipulations which served to degrade access to the relevant instructions invariably led to increases in the size of the SCE. Such increases were then regarded as providing support for both the Petrusic (1992) slowed accrual and the Leth-Steensen and Marley (2000) instructional-pathway competition view regarding the locus of the SCE. Analogously, similar increases in the size of the SCE were expected in the present experiments for conditions in which relevant instructional context had been switched from the previous trial.

More specifically, consider the case in the present experiment in which participants had to switch from the “Shorter?” instruction on the previous trial to the “Taller?” instruction on the current trial where the stimulus pair on the current trial

was a relatively short one (i.e., semantically incongruent with the current trial instruction). According to the Petrusic (1992) view, access to the relevant instructional context is assumed to be degraded for such semantically incongruent pairs, in general, and it was expected that such access would become even more degraded under conditions in which the irrelevant instructional context had just very recently been the relevant one. According to the Leth-Steensen and Marley (2000) view, the relative difference in the strengths of the (competing) relevant and irrelevant instructional pathways is assumed to be smaller for such semantically incongruent pairs, in general, and it was expected that such instructional pathway strength differences be even more diminished under conditions in which the irrelevant instructional context had just very recently been the relevant one. In contrast, for the alternative semantically congruent case in which participants had to switch from the “Taller?” instruction on the previous trial to the “Shorter?” instruction on the current trial with a relatively short stimulus pair on the current trial, under both the Petrusic and the Leth-Steensen and Marley views, carryover effects of the instructional context from the previous trial were not expected to have nearly as much influence on decisional processing.

Experiment 1

The current Experiments 1a and 1b, however, provided some evidence that, if anything, SCEs are attenuated on trials in which the instruction had been switched. Moreover, the strongest evidence for an attenuation of the SCE with switched instructions occurred in the short-RSI conditions (the very conditions in which the largest enhancement of the SCE had originally been expected). Notably, in all of

those short-RSI conditions, fairly substantial overall instruction-switch costs in both RT and PC were also present indicating that the need to switch instructions did indeed make additional processing demands.

Inhibition account. Obviously, within the context of the hypotheses developed on the basis of the Petrusic (1992) and Leth-Steensen and Marley (2000) views, such a result does not seem to make very much sense. One way to explain these results, though, would be to consider the possibility that the process of switching instructional sets with very short RSIs necessitates that the previously relevant, but now irrelevant, instructional context be rather strongly inhibited (a notion which is analogous to Druey & Hubner's, 2008, proposal that stronger inhibition is needed as the risk of accidental response re-executions increases). Note that in contrast to all other previous task switching work (where half of the time responses associated with the irrelevant task set still end up being correct; e.g., having the digit "2" as the stimulus where the responses "even" and "smaller than 5" are both mapped onto the left response key), in the current paradigm, responses associated with the irrelevant (instructional) task set are always incorrect with respect to the relevant (instructional) tasks set. The presence of strong inhibition of the now-irrelevant instructional context on switch trials would then serve to dramatically reduce the interfering effects of that irrelevant instructional context for both semantically incongruent and congruent stimulus trials thus diminishing the basis for the SCE. Such a view might also be able explain why reductions in the SCE for short-RSI switch trials were sometimes not observed here (i.e., for the Split 1 end-term pairs in Experiment 1b which are relatively much shorter or taller than any of the other pairs), if it could be assumed

that it is harder to apply such inhibition when the size of the pair on the current trial is highly semantically congruent with the previous instruction (but, necessarily, semantically incongruent with the current instruction).

In general, it has been rather hard to provide solid evidence for the presence of inhibition in task switching, with research on this issue typically focusing on the role of inhibition somewhat indirectly using the ABA versus CBA task sequence paradigm described earlier. Interestingly, some very recent work has provided some evidence that inhibition of a previous task can be triggered at the stimulus processing stage of a new task (Sdoia & Ferlazzo, 2008). With respect to the present paradigm and results, if the SCE does indeed arise due to competition between (instructional) task sets, its attenuation on short-RSI switch trials could then actually be regarded as providing a form of evidence for the presence of task-set inhibition in task switching.

Faulty reconfiguration account. On the other hand, another way to potentially account for the presence of attenuated SCEs in the short-RSI, instruction-switch conditions is to consider that they could have been diluted by cases in which reconfiguration failed because in such cases the nominal and actual instruction set (and, hence, the nominal and actual semantic congruency of the pair) would not actually have been the same (i.e., although the instruction should have been “Shorter?”, the comparison process was running off according to the not-yet-abandoned “Taller?” instructional set). Although, technically, such cases should then have ended up in error responses (and, hence, not actually contribute to the SCE for correct RTs), perhaps in a large number of these cases, reconfiguration did manage to kick in before the response was made allowing participants to make the appropriate

adjustment before responding. Such response adjustments could be regarded as occurring either in terms of a scenario in which the decision process is allowed to run off to completion under the previous instruction but then a stage is added in which the response to be made is mentally flipped, or in terms of a scenario involving an evidence accrual process whereby sometime during the course of that process the mental labels associated with each of the decision boundaries are reversed (with respect to the latter view, however, note that it cannot be the case that reconfiguration would simply reverse the direction of the evidence signal itself midway through the accrual process because that would mean that the process essentially would then have to backtrack from its state of accumulation at the point of reconfiguration which would then be expected to drastically affect trials in which the newly reconfigured instruction was semantically incongruent with the current stimulus pair).

Such a view would also predict that due to the presence of such response adjustments, the overall mean correct RT for the short-RSI, instruction-switch conditions be larger than for the short-RSI, instruction-repeat conditions, which was indeed found. Moreover, it would also predict that the corresponding within-cell RT SDs be much larger given that they would represent a mixture of adjusted and non-adjusted RTs. In light of this point, average within-cell, correct-RT SDs were computed for each of the Split 1 and 2 end-term pairs in the instruction-switch and instruction-repeat conditions of both Experiment 1a and 1b and are given in Figures 7 – 8 (where all missing cell values and outliers were replaced in the same manner as for the corresponding RT cell values when deriving these average SDs). One very interesting aspect of the results for the short (i.e., 200 ms) RSIs in these figures is that

whereas within-cell SDs are typically lower for the semantically congruent end-term pairs in the instruction-repeat conditions, the opposite finding is more typical in the instruction-switch conditions, namely, less variability for the RTs within a cell for semantically incongruent end-term pairs.

Such a finding is consistent with the notion that in the short-RSI, instruction-switch conditions, semantically incongruent pairs (which, as noted, are semantically congruent with the previous trial's instruction on switch trials) were often being processed like semantically congruent ones and vice versa for semantically congruent pairs. Note, as well, that with respect to the previous inhibition explanation, although trial-to-trial variability in the inhibition process could also serve to explain overall increases in within-cell SDs for the short-RSI, instruction-switch in comparison to the instruction-repeat conditions (where overall increases in RT in these conditions would also be expected if the application of inhibition is at all effortful), such a view would have a hard time accounting for the enhanced variability for the semantically congruent end-term pairs in the instruction-switch conditions (i.e., it would have to be assumed that inhibitory processing is more variable when switching to a congruent instructional context and away from one that is incongruent with the size of the current pair).

Unfortunately, if this faulty reconfiguration account is indeed correct, the results obtained would then not actually be very diagnostic with respect to the various theories of the SCE. Namely, if it is assumed that the SCE is reduced on short-RSI, instruction-switch trials in these two experiments because participants are sometimes processing semantically incongruent pairs under the previous semantically congruent

instruction (and vice versa for semantically congruent pairs), such a view really doesn't say very much about why the SCE is occurring in the first place and any of the available SCE theories still remain viable candidates. On the other hand, this account would have something to say with respect to the issue of task switching, namely, that the mode of processing during (instructional) task switches is much more discrete than would likely be prescribed by the standard task-set inertia view. In other words, it does not seem to be the case that instruction switches involve a gradual turnover of instruction sets but rather that the comparative processing of these symbolic stimuli seems only able to occur under one instructional set at a time (but note also that such a state of affairs could be regarded as being especially troublesome for the Leth-Steensen and Marley, 2000, view given that it is specifically assumed by that view that both forms of the instruction are continually activated, to either a greater or lesser extent depending on both their current relevancy and semantic congruency, and contributing to the decision process).

Experiment 2

Probably the most notable thing about the results obtained in Experiment 2 was that the key finding from Experiment 1, namely, the attenuation of the SCE in the short-RSI, instruction switch condition, was not replicated. In fact, for the expected instructions in Experiment 2, the SCE looks almost identical for both instruction-repeat and instruction-switch trials regardless of the length of the RSI (and also looked very similar in size across these same conditions for the unexpected instructions).

Importantly, though, the need to switch instructions with short RSIs was still associated with a substantial switch cost for both expected and unexpected instructions (although in the latter case, Figure 5 indicates that this was mainly due to slower responses for unexpected switches to the shorter instruction). On the other hand, no switch costs seemed to be evident when the RSI was long for either expected or unexpected instructions. Moreover, no real costs were evident for the unexpected instructions in either RSI condition. Although the lack of an instruction expectancy effect for the short RSIs might itself have been expected because there was no time to prepare for the next instruction anyway, clear instruction expectancy effects were to be expected for the long RSIs given that that actual instruction would have been different from the one that had been prepared during the 2000-ms RSI.

These instruction switch and (null) instruction expectancy effects are, however, consistent with the notion that the participants in Experiment 2 were not really aware of the underlying two-by-two sequence of switch and repeat trials. Hence, they were never actually trying to predict the next instruction but were simply were waiting to be exogenously cued (in a bottom-up fashion) by the color of the stimulus pair as to which instruction to use. Note that such a notion does not preclude the use of preparatory processing during the long RSIs but does maintain that such preparation be unspecific to any instruction. Moreover, under such conditions, switch costs would still be expected for short RSIs due to carryover effects from the previous instructional task set whose activation would not have had very much time to decay. Presumably, there was something about the design of this experiment that might not have rendered the underlying task transition sequence salient enough to be picked up

by these participants. Perhaps, the periodic breaking up of that sequence by unexpected instructions (which did happen twice within every block of 36 trials) was enough to make it seem unpredictable.

With respect to the SCE, it is of note that Mayr and Keele (2000) found that backward inhibition only actually occurs when the identity of the upcoming task is known (see also Hubner, Dreisbach, Haider, & Kluwe, 2003), which suggests that participants seem to need to know for sure that a task set is to be abandoned in order to fully inhibit it. Hence, if the participants were indeed waiting to be cued by the color of the stimulus on each trial in Experiment 2, less inhibition of previous instructional task sets (and, hence, a normal-sized SCE) would have been expected under the inhibition view.

On the other hand, with respect to the faulty reconfiguration account of the reduced SCEs for short RSI trials in the previous experiment, given the lack of a similar effect in Experiment 2, it would have to be assumed here that there was something about the conditions in this experiment that actually served to decrease the likelihood of faulty reconfiguration. Hence, perhaps it was the case that the presence of task uncertainty served to greatly enhance the potency of the stimulus colors as task cues.

Conclusion

Clearly, the results obtained in these experiments were not consistent with what had been expected at the outset given either the Petrusic (1992) or the Leth-Steensen and Marley (2000) views of the basis for the SCE. Both views had predicted, a priori, that enhancements in the size of the SCE should be observed for instruction-switch trials. However, in the results for Experiments 1a and 1b, attenuations in the size of this effect seemed to be more the norm on instruction-switch trials, and in the results for Experiment 2, the SCE was seemingly unaffected by the need to switch instructions.

It is truly unfortunate that the present results ended up turning out the way they did. If they had indeed gone the way that had been expected of them, they would have added nicely to the set of results obtained in the three previous studies of Shaki, Petrusic, and Leth-Steensen, all of which showed enhancements in the size of the SCE under conditions of increased memory demands associated with retrieving the relevant comparative instructions. In that case, the body of evidence in support of the validity of the Petrusic (1992) and Leth-Steensen and Marley (2000) views could have been regarded as being fairly strong.

If so, that would indeed have been quite important because it has generally been difficult to resolve the issue of which account of the SCE is the most plausible. In their own right, each of the theories for the SCE described at the start of the Introduction have turned out to be quite difficult to disprove. Hence, the general strategy when researching this effect has been to try to provide sets of results that seem to be most consistent with one view or another (although note that proponents

of the semantic coding model have had nothing more to add since Cech & Shoben, 2001, and also that it is sometimes the case that the bias model account for this effect has simply been accepted as fact on purely intuitive grounds, e.g., Schwarz & Stein, 1998, even though no real evidence for it has ever really been provided).

Returning to what actually did happen in these experiments, the present findings could likely be regarded as being more problematic for the Leth-Steensen and Marley view than for the Petrusic view. Indeed, perhaps, the only way to save the former view would be to assume that instructional pathway interference is actually reduced on instruction-switch trials in comparison to instruction-repeat trials because inhibition of the irrelevant instruction is stronger on instruction-switch trials.

However, at this point such a notion is somewhat speculative and cannot actually be distinguished from the alternate possibility that the present results could also be due to faulty reconfiguration (where, in fact, an examination of the within-cell RTSDs for Experiments 1a and 1b actually provides more support for the faulty reconfiguration account than for the inhibition account).

On the other hand, the current results do not necessarily disconfirm the Petrusic view because it is possible that the memory demands associated with retrieving the relevant instructions in these experiments, even for instruction-switch trials, were just not as demanding as they had been in the previous Shaki, Petrusic, and Leth-Steensen studies (especially the latter two involving the remembered instructions and the category contingent instructions, respectively; although it must be noted that there has been some task switching work involving meaningful vs. abstract cue-task associations that has shown that the size of the switch cost can be moderated

by degree to which the task cues facilitate retrieval of the corresponding task sets from memory; Mayr & Kliegl, 2000).

Interestingly, the present results also provide a counterexample to what has been a particularly troublesome criticism of the Petrusic view, namely that increases in the size of the SCE would naturally be expected to occur for any manipulations that serve to slow down overall RTs. Obviously, the reductions in the size of the SCE on the slower instruction-switch trials in the present Experiments 1a and 1b show that this conjecture is indeed not always true.

Finally, it is also of interest that the present study has not been the only one where reductions in the size of the SCE have been observed for manipulations which had been assumed, a priori, to lead to an enhancement in the size of this effect. Namely, in Shaki and Petrusic (2003) participants performed a condition in which multiple copies of each of the two possible comparative instructions were displayed simultaneously on each trial, and participants had to decide which one to use by determining which instruction had the more displayed copies (which those researchers referred to as their “Stroopy” instruction condition). To their surprise, attenuated SCEs were found for this condition in comparison to one in which no interfering instructional information was used. Presumably, the similarities in the results obtained under the instruction-Stroop conditions of Shaki and Petrusic (2003) and under the instruction-switch conditions of the present experiments suggest the possibility that there might be a common underlying account for both sets of results. In this vein, it is notable that, as discussed in the Introduction, a role for task-set

inhibition has been implicated in previous task switching work utilizing the standard Stroop colour-word paradigm (i.e., Allport & Wylie, 2000).

References

- Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: The MIT Press.
- Allport, A., & Wylie, G. (2000). "Task-switching", stimulus-response bindings, and negative priming. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 35-70). Cambridge, MA: The MIT Press.
- Banks, W. P. (1977). Encoding and processing of symbolic information in comparative judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 11, pp. 101-159). New York: Academic Press.
- Banks, W. P., & Root, M. (1979). Semantic congruity effects in judgments of loudness. *Perception and Psychophysics*, 26, 133-142.
- Birnbaum, M. H., & Jou, J. (1990). A theory of comparative response times and "difference" judgments. *Cognitive Psychology*, 22, 184-210.
- Čech, C. G. (1995). Is congruity due to encoding? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 5, 1275-1288.
- Cech, C. G., & Shoben, E. J. (1985). Context effects in symbolic magnitude comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 299-315.
- Čech, C.G. & Shoben, E.J. (2001). Categorization processes in mental comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 800-816.

- Čech, C. G., Shoben, E. J., & Love, M. (1990). Multiple congruity effects in judgments of magnitude. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 1142-1152.
- Deheane, S. (1989). The psychophysics of number comparison: A reexamination of apparently incompatible data. *Perception and Psychophysics*, 45, 557-566.
- Druey, M. D., & Hübner, R. (2008). Response inhibition under task-switching: Its strength depends on the amount of task-irrelevant response inhibition. *Psychological Research*, 72, 515-527.
- Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cognitive Psychology*, 10, 203-243.
- Jamieson, D. G., & Petrusic, W. M. (1975). Relational judgments with remembered stimuli. *Perception & Psychophysics*, 18, 373-378.
- Hubner, M., Dreisbach, G., Haider, H., & Kluwe, R. H. (2003). Backward inhibition as a means of sequential task-set control: Evidence for reduction of task competition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 289-297.
- Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, 107, 62-100.
- Leth-Steensen, C., Petrusic, W. M., & Shaki, S. (in preparation). Category contingent comparative judgments.
- Link, S. W. (1990). Modelling imageless thought: The relative judgment theory of numerical comparisons. *Journal of Mathematical Psychology*, 34, 2-41.
- Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of

- backward inhibition. *Journal of Experimental Psychology: General*, 129, 4-26.
- Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1124–1140.
- Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 22, 1423-1442.
- Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception and Psychophysics*, 13, 180-184.
- Petrusic, W. M. (1992). Semantic congruity effects and theories of the comparison process. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 962-986.
- Petrusic, W. M., & Baranski, J. V. (1989b). Semantic congruity effects in perceptual comparisons. *Perception & Psychophysics*, 45, 439-452.
- Petrusic, W. M., Shaki, S., & Leth-Steensen, C. (2008). Remembered instructions with symbolic and perceptual comparisons. *Perception & Psychophysics*, 70, 179-189.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207-231.
- Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1404-1419.

- Schwarz, W., & Stein, F. (1998). On the temporal dynamics of digit comparison. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1275-1293.
- Sdoia, S., & Ferlazzo, F. (2008). Stimulus-related inhibition of task set during task switching. *Experimental Psychology*, 55, 322-327.
- Shaki, S. & Algom, D. (2002). The locus and nature of semantic congruity in symbolic comparison: Evidence from the Stroop effect. *Memory and Cognition*, 30 (1): 3-17.
- Shaki, S., Leth-Steensen, C., & Petrusic, W. M. (2006). Effects of instruction presentation mode in comparative judgments. *Memory & Cognition*, 34, 196-206.
- Shaki, S., & Petrusic, W. M. (2003). Instruction interference and the semantic congruity effect. In B. Berglund & E. Borg. (Eds.) *Fechner Day 2003. Proceedings of the Ninetenth Annual Meeting of the International Society for Psychophysics* (pp. 293-298). Larnaca Bay, Cyprus: The International Society for Psychophysics.
- Sohn, M.-H., & Anderson, J. R. (2001). Task preparation and task repetition: Two-component model of task switching. *Journal of Experimental Psychology: General*, 130, 764-778.

Figure 1A. Split 1 response time (RT) cell means for Experiment 1a.

Figure 2A. Split 2 response time (RT) cell means for Experiment 1a.

Figure 2B. Split 2 response time (RT) cell means for Experiment 1b.

Figure 3A. Split 1 proportion correct (PC) cell means for Experiment 1a.

Figure 3B. Split 1 proportion correct (PC) cell means for Experiment 1b.

Figure 4A. Split 2 proportion correct (PC) cell means for Experiment 1a.

Figure 4B. Split 2 proportion correct (PC) cell means for Experiment 1b.

Figure 5. Response time (RT) cell means for the Split 1 end-term pairs in Experiment 2.

Figure 6. Proportion correct (PC) cell means for the Split 1 end-term pairs in Experiment 2.

Figure 7A. Split 1 response time standard deviations (RTSD) for Experiment 1a.

Figure 7B. Split 1 response time standard deviations (RTSD) for Experiment 1b.

Figure 8A. Split 2 response time standard deviations (RTSD) for Experiment 1a.

Figure 8B. Split 2 response time standard deviations (RTSD) for Experiment 1b.

Appendix A
Anticipated Results for Experiment 1a, Experiment 1b and Experiment 2

	<i>200ms</i>	<i>2000ms</i>
<i>Switch</i>	SCE Enhanced	
<i>Repeat</i>		

Figure A1. Anticipated results for Experiment 1a and 1b.

		Expected		Unexpected	
		200ms	2000ms	200ms	2000ms
<i>Switch</i>		SCE Enhanced		SCE Enhanced	SCE Enhanced
<i>Repeat</i>					SCE Enhanced

Figure A2. Anticipated results for Experiment 2.

Appendix B

Response Time Results for the Eight Participants in Experiment 2.

Figure B1. Response time (RT) cell means for the Split 1 end-term pairs for Participant 2 in Experiment 2.

Figure B2. Response time (RT) cell means for the Split 1 end-term pairs for Participant 3 in Experiment 2.

Figure B3. Response time (RT) cell means for the Split 1 end-term pairs for Participant 5 in Experiment 2.

Figure B4. Response time (RT) cell means for the Split 1 end-term pairs for Participant 6 in Experiment 2.

Figure B5. Response time (RT) cell means for the Split 1 end-term pairs for Participant 7 in Experiment 2.

Figure B6. Response time (RT) cell means for the Split 1 end-term pairs for Participant 8 in Experiment 2.

Figure B7. Response time (RT) cell means for the Split 1 end-term pairs for Participant 9 in Experiment 2.

Figure B8. Response time (RT) cell means for the Split 1 end-term pairs for Participant 10 in Experiment 2.