NOTE TO USERS

This reproduction is the best copy available.

UMI[®]

Expanding SNARC for temperature, negative numbers and intelligence

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

Doctor of Philosophy

by

Joel A. Lucas

Department of Psychology
Carleton University

November 2009

©2009 Joel A. Lucas

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: 978-0-494-63862-0
Our file *Notre référence*
ISBN: 978-0-494-63862-0

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.

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

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

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

Canada

Abstract

Previous experiments suggest that the direction of SNARC is fixed for numbers based on a mental number line. This theory is challenged as our experiments show that for all symbols representing semantic quantities, including numbers, the direction of SNARC is dependent on instruction such that faster right hand and upward responses were always associated with the highest measure corresponding to the instruction, and faster and downward responses were always associated with the lowest measure corresponding to the instruction. In Experiment 1, an instruction-dependent SNARC effect was obtained for animal size, and Experiment 2 demonstrated that SNARC could be obtained for a more abstract form of magnitude such as animal intelligence. In Experiments 3, 4, and 5, participants were required to pick the warmer or colder digit from positive and negative pairs representing relatively high or low temperatures. Presentation of pairs and responses were horizontal in Experiment 3 and vertical in Experiments 4 and 5. An instruction-dependent SNARC effect occurred such that responses to the “coldest” and “warmest” pairs were always fastest to the right, or upward depending on the instruction, but left or downward responses were faster for pairs least compatible with the instruction. These results show that numbers based on a specific semantic property may be processed differently than nonspecific numbers. Additionally, a thermometer appeared on the screen for half of the trials in Experiments 4 and 5, which elicited a stronger SNARC effect demonstrating the importance of the semantic properties of stimuli. In Experiment 6, the instruction dependent pattern of SNARC continued using the speed of runners represented by positive and negative numbers. These results suggest that the direction of SNARC is caused by two additive factors including the absolute value of numbers, as well as the semantic quantity represented by numbers defined by the instruction. In the absence of one factor, the other factor dominates.

SNARC in these experiments is best explained by the assignment of codes to the poles of stimulus and response dimensions and the compatibility of those codes. When response and stimulus codes are compatible responses are fast, but slower when they are incompatible.

Acknowledgements

Thank-you to my supervisor, Bill Petrusic, for teaching me to stand on my own two feet.

Thank-you to my committee members, Craig Leth-Steensen, Deepthi Kamawar, Jim Davies and Tony Marley for your careful reading and thoughtful feedback.

Thank-you to Etelle Bourassa, June Callender (J.D.C.), Robin Beasley, Joanne Sonley, Payal Chadha, and Natalie Pressburger for putting up with me.

This research was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada, administered to William Petrusic.

Table of Contents

List of Tables – vii

List of Figures – x

Introduction – 1

The Mental Number Line – 1

SNARC – 3

SNARC for Various Tasks – 5

The Development of the Mental Number Line and the Direction of SNARC – 7

The Locus of SNARC – 9

A Model of SNARC – 12

SNARC is contextually-driven and multidirectional – 14

SNARC for Two-Digit Numbers? – 16

Does the mental number line include negative numbers? – 17

The Semantic Congruity Effect – 20

The importance of sleep to consolidate memory – 23

The Current Thesis – 25

Experiment 1 – 28

Method – 29

Results

Session 1 – 32

Session 2 – 39

Discussion – 46

Experiment 2 – 48

Method – 48

Results – 50

Discussion – 58

Experiment 3 – 60

Method – 60

Results – 62

Discussion – 73

Experiment 4 – 77

Method – 78

Results – 80

Discussion – 97

Experiment 5 – 99

Method – 99

Results – 100

Discussion – 115

Experiment 6 – 119

Method – 121

Results – 121

Discussion – 129

General Discussion – 130

References – 150

Appendix A – 150

List of Tables

Table 1.	Average ratings for the intelligence of each animal included in the experiment.....	51
Table 2.	Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 3.	142

Table 3. Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 4. The X shows an incorrect prediction.

.....143

Table 4. Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 5. The Xs show predictions that were wrong.

.....144

Table 5. Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 6.

.....145

List of Figures

Figure 1.	Figure 1. Mean response times for each animal pair for each instruction in Session 1, with the smallest pair at 1 (bee, ant), progressing to the largest at 6 (hog, cow).	33
Figure 2.	Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of words representing animal size, in Session 1. The smallest pair is represented by 1 progressing to the largest pair represented by 6.	36
Figure 3.	Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair of pictures representing animal size, in Session 1. The smallest pair is represented by 1 progressing to the largest pair represented by 6.	37
Figure 4.	Mean response times for each animal pair, for each instruction in Session 2, arranged with the smallest pair at 1 (bee, ant), progressing to the largest at 6 (hog, cow).	40
Figure 5.	Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of words representing animal size, in Session 2. The smallest pair is represented by 1 progressing to the largest pair represented by 6.	42

Figure 6.	Mean left-hand response time subtracted from mean right- hand response times as a function of each stimulus pair of pictures representing animal size, in Session 2. The smallest pair is represented by 1 progressing to the largest pair represented by 6.	43
Figure 7.	Mean response times for each animal pair, for each instruction, arranged with the least intelligent pair at 1 (worm, trout), progressing to the largest at 6 (dolphin, human).	52
Figure 8.	Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of words representing animal intelligence.	56
Figure 9.	Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of pictures representing animal intelligence.	57
Figure 10.	Mean response times collapsed across thermometer presentation, for the instructions “colder” and “warmer” in blocked and intermixed conditions, as a function of each stimulus pair. The x-axis represents the order of temperature with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9,8)	64

Figure 11.	Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “colder”, for the blocked and intermixed conditions. The left panel represents negative and the right panel represents positive pairs.	67
Figure 12.	Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “colder”, for the blocked and intermixed conditions.	69
Figure 13.	Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “warmer”, for the blocked and intermixed conditions. The left panel represents negative and the right panel represents positive pairs.	70
Figure 14.	Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “warmer”, for the blocked and intermixed conditions.	71
Figure 15.	Mean response times collapsed across thermometer presentation, for each instruction and condition, as a function of each stimulus pair. The x-axis represents the order of temperature with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).	82

Figure 16.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	87
Figure 17.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions.	88
Figure 18.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	89
Figure 19.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions.	90
Figure 20.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	91

Figure 21.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions.	92
Figure 22.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	93
Figure 23.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions.	94
Figure 24.	Mean response times collapsed across thermometer presentation, for each instruction and condition, as a function of each stimulus pair. The x-axis represents the order of temperature with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).	102
Figure 25.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	106

Figure 26.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions.	107
Figure 27.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	108
Figure 28.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions.	109
Figure 29.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	110
Figure 30.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions.	111

Figure 31.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	112
Figure 32.	Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions.	113
Figure 33.	Mean response times, for each instruction and condition, as a function of each stimulus pair. The x-axis represents pairs of “runners” with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).	122
Figure 34.	Mean left-hand response time subtracted from median right hand response times as a function of each stimulus pair, for the instruction “slower”, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	126
Figure 35.	Mean left-hand response time subtracted from median right hand response times as a function of each stimulus pair, for the instruction “faster”, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.	127

Expanding SNARC for temperature, negative numbers and intelligence

Numbers seem to be processed in such a way that they are represented spatially. A spatial representation is arguably needed when one is picturing numbers in the mind when trying to solve mathematical problems. Otherwise, the reason for an evolutionary basis for a spatial representation of numbers is not immediately clear. However, much evidence exists to suggest that numbers do in fact have a spatial representation when processed as shown by behavioural studies. For instance, Grade 6 students gifted in math tend to use more visuospatial techniques compared to other children (van Garderen & Montague, 2003).

Neurological studies also support the connection between visuospatial skills and mathematical ability. The intraparietal sulci of the cortex are necessary for both numerical and spatial tasks (see Hubbard, Piazza, Pinel, & Dehaene, 2005 for review). Patients with Gerstmann syndrome share spatial deficits such as left-right disorientation and finger agnosia as well as numerical deficits such as dyscalculia (Gerstmann, 1940). The fact that numerical and spatial tasks share common neural areas is convincing evidence for the spatial properties of number processing. Potentially, understanding the aspects of spatial components of number processing are an important first step to generate practical applications for people with mathematical disabilities.

The Mental Number Line

Behavioural evidence suggests that the spatial representation of numbers occurs specifically in a linear manner. The idea of a mental number line can be traced to Moyer and Laundauer (1967) who found that when comparing two single digits, the participants' response time was inversely related to the numerical distance between the digits.

Comparing the pair (1, 9) for instance was faster than comparing (1, 2). The distance effect, as it is known, is an instance of Weber's Law, and supports the idea of a mental number line because it mimics response times one would make comparing numbers on a ruler.

The distance effect also occurs for two-digit numbers. When participants compared two-digit numbers to a reference number of 65, the response time was inversely proportional to the numerical distance between the two numbers (Dehaene, Dupox and Mehler, 1990). Additionally, there was a progressive increase in time from one decade to another, instead of an abrupt increase, suggesting that participants were processing the full number rather than simply attending to the decades column and ignoring the digits column, when comparing.

Calabria and Rossetti (2005) presented lines to normal participants made up of the words "neuf" or "deux" (e.g., neufneufneufneufneufneuf). Participants were asked to bisect the line at the midpoint. As predicted, they bisected the line more to the left for the "deux" line compared to the "neuf" line, indicating that perhaps the stimulus activated a mental number line representation, which in turn biased their responses in a direction corresponding to the its spatial location.

The distance effect is very robust, whether it be for Arabic numeral comparisons or for any physical comparisons including line-length (Henmon, 1906), pitch (Buckley & Gillman, 1974), animal size (Moyer, 1973), and even more abstract concepts like the ferocity of animals (Kerst & Howard, 1977).

Additional support for a mental number line comes from neurological studies. Some patients with brain damage in the parietal lobe will experience hemi-spatial neglect,

which is a tendency to ignore one side of their visual space. Zorzi et al. (2002) found that when these patients were asked to bisect a line, they would ignore part of it and therefore bisect it off of the midpoint. When these same patients are asked for the midpoint between two numbers, they tend to bias their answers towards one side of the midpoint (e.g., saying that 6 is the midpoint between 1 and 9).

Additional evidence comes from a study in which the experimenters utilized repetitive transcranial magnetic stimulation, a technique that mimics brain lesions in participants by temporarily activating the neurons in a certain area, which then disrupts tasks dependent on that area. When transcranial magnetic stimulation is applied to the angular gyrus, located in the parietal lobe, visual search tasks are disrupted and number comparisons are slowed as well (Göbel, Walsh and Rushworth, 2001). When specifically the *left* angular gyrus is disrupted, numerical comparisons *larger* than the standard are slowed. When the *right* angular gyrus is disrupted, numerical comparisons *smaller* than the standard are slowed. This suggests that the neural correlate of the mental number line could span across both parietal lobes.

SNARC

Most of the cognitive research regarding the mental number line hypothesis has centred on the SNARC effect (Spatial-Numerical Association of Response Codes), which is the difference in response time that occurs between left and right response times in response to magnitude. As magnitude changes, the difference in response time changes, which is hypothesized to be reflective of a mental number line.

Initially, Dehaene, Dupoux, and Mehler (1990) asked participants to indicate, by pressing a left or right key, whether a number (1-99) was smaller or larger than the

reference number (55). Participants who used the left key to indicate “smaller” responses and the right key to indicate “larger” responses were faster than participants who were assigned to respond using the alternate mapping (i.e., left-larger and right-smaller). These results suggest that the difference in response time was because of a mental number line, oriented from left to right, which may have facilitated the group that responded in a direction congruent with the mental number line but inhibited participants who responded in an incongruent direction.

In a subsequent study, Dehaene et al. (1993) examined this effect more directly. Participants were required, on each trial, to judge if a digit from 0-9 was odd or even (i.e., parity task), by pressing a left key to respond “odd” or a right key to respond “even”, as well as the alternate order for the other half, such that there were both left and right responses for each digit. They found that when the digit was relatively small (e.g., 1, 2) responses were faster than right, but when digits were relatively large (e.g., 8, 9) right hand responses were faster compared to the right.

They argued that this effect was the result of the spatial coding of numbers, on a mental number line with the smallest numbers on the left, progressively increasing to the right like a ruler. Somehow, responses were facilitated when the spatial coding and direction of response were congruent (i.e., smaller to the left and larger to the right) as opposed to incongruent. The effect didn’t change when participants crossed their hands, suggesting that SNARC wasn’t related to handedness or the hemispheric control of hand movement, but rather the direction of response.

There is also evidence to suggest that SNARC develops at an early stage in life. Berch, Foley, and Hill (1999) tested children from Grades 2, 3, 4, 6, and 8 (ages 7-13) on

a parity task and found that SNARC effects could be observed in children in the third grade.

SNARC for various tasks

To attain a SNARC effect, the particular task does not seem to be that important, as long as participants make some leftward and rightward responses to each number in the set. In addition to odd and even decisions (parity task), the SNARC effect has also been found for comparative judgements (e.g., “is the digit smaller or larger than 5?”, Shaki & Petrusic, 2005b). In fact, it seems as though the SNARC effect can be attained without any intentional processing of numbers at all. Fias, Lauwereyns, and Lammertyn (2001) asked participants to make left and right responses to tasks which required them to distinguish the orientation of a shape or a line which was superimposed on a digit. They managed to find a strong SNARC effect even though the digit was totally irrelevant to the task.

SNARC does not seem to be dependent on the mode of presentation, or the mode of response. Petrusic (personal communication) found a SNARC effect using left and right mouse clicks with a single hand. Fischer (2003a) found that touching the left or right area of a screen elicited a SNARC effect, and Schwarz and Keus (2004) found a SNARC effect using saccadic eye movements in response to digits. These results suggest that SNARC is a reflection of processing prior to that involved with the motor response. The mode of presentation also doesn't seem to alter the SNARC effect. Nuerk, Wood and Willmes (2004) found SNARC when numbers were presented audibly, as visual words (also see Dehaene et al. 1993, Experiment 8), and as a dot matrix.

The SNARC effect is not exclusive to numbers but can be found for other symbols of magnitude, as well as nonnumerical stimuli. Gevers, Reynvoet, and Fias (2003, but see Dehaene 1993) found a SNARC effect using the 12 months of the year in both order-relevant (“Is the month presented before or after July?”) and irrelevant (“Does the month end with an ‘r’?”) categories. In a second experiment, they found SNARC using letters in both an order-relevant (Does the letter occur before or after ‘O’ in the alphabet”) and an order-irrelevant task, “Is the presented letter a vowel or a consonant?”). These results call into the question of the dependency of the SNARC effect on magnitude per se as opposed to familiarity of sequence.

SNARC has also been found for symbols of magnitude other than numbers. Shaki and Petrusic (2005b) found a SNARC effect using size comparisons between animal names (Experiment 2). Pairs of animal names were either relatively small (snail-mouse) or large (moose-whale). However, one interesting aspect of the findings that differed from results of SNARC experiments using numbers, was that the direction of the SNARC effect depended on the instruction. When participants were instructed to pick the smaller animal, as expected, they found that leftward responses were faster than right for relatively small pairs but rightward responses were faster than left, for relatively large pairs. However, when participants were instructed to pick the larger animal, the direction of SNARC was reversed. Leftward responses were faster for larger pairs and rightward responses were faster for smaller pairs. They found a similar pattern of results using both line lengths and female names representing line lengths.

The reason why the direction of SNARC is instruction-dependent for forms of magnitude but is not for numbers is of particular interest. It was suggested by Shaki and

Petrusic (2005b) that the instruction acts as an “anchor” starting with the smaller magnitude on the left when the instruction is smaller or with larger on the left when the instruction is “larger”. From this anchor the other stimuli are sequentially ordered from left to right. It is possible, however, that the reason this change in direction does not occur for numbers is because numbers are continuously presented and overlearned in an array, ordered smallest to largest numbers, from left to right, and therefore are resistant to change direction based on instruction. However, since animals and line lengths are not normally learned in a particular array, their mental representation is more easily influenced by the context of the to-be-completed task.

The Development of the Mental Number Line and the Direction of SNARC

The basis for the development of the SNARC effect and its left-to-right orientation is currently debated in cognitive psychology. It has been proposed that the direction of SNARC is the result of writing habit (Dehaene et al., 1993). Dehaene et al. showed that French participants, who had written from left to right their entire lives, showed a more robust SNARC effect compared to Iranian participants, who had written from right to left earlier in life, and had switched when moving to France. The degree of SNARC was dependent on the amount of time participants had been writing from a left to right direction. Although Iranians with no exposure to French culture were not tested, the regression analysis conducted by Dehaene predicted that Iranian participants who did not have any exposure to French culture would show a SNARC effect in the opposite direction.

Recently, however, Shaki, Fischer and Petrusic (2009) compared SNARC for Canadian participants as well as for Israeli and Palestinian participants, In contrast to

Canadian participants, Palestinian participants read and write from right to left as opposed to left to right. Additionally, Israeli participants also read and write from right to left, however, they also read Arabic numbers from left to right. If reading or writing direction is the cause for the orientation of the mental number line, then Canadians and Palestinians should show SNARC in opposing directions but Israelis would potentially show a diluted SNARC effect. Indeed, this is what occurred. Canadians showed SNARC from left to right and Palestinians showed SNARC from right to left. Israeli participants did not show a significant SNARC suggesting that their opposing reading and writing directions diluted SNARC. These results are strong evidence that the orientation of the mental number line, and therefore, the direction of SNARC is heavily influenced by both reading and writing direction.

One source of evidence that challenges the notion that SNARC is caused by writing habit is the finding that SNARC occurs vertically from bottom to top, as well as horizontally, (Schwartz and Keus, 2004, which will be discussed more later). Assuming that most students were Dutch or European because the experiment was conducted in Holland, it would not be expected that their participants would have a bottom-to-top orientation caused by reading or writing habits. Additionally, Ito and Hatta (2004) found a SNARC effect in the vertical direction for Japanese participants (to be discussed later) that was oriented from bottom-to-top, even though their writing habit is from top to bottom.

Even if writing habit per se is not the cause of the direction of SNARC, it is quite possible that cultural influences, in general, cause the mental number line to orient from left to right by default. People in the Western world generally have a left to right

orientation, not only for writing, but also reading, and the display of numbers on a keyboard.

Hemispheric dominance is another possible explanation of the direction of the SNARC effect. However, the direction of SNARC does not seem to be dependent on handedness. Dehaene et al. (1993, Experiment 6) found that when participants crossed their hands so that leftward responses were made on the keypad with the right hand, and rightward with the left-hand, a robust SNARC effect occurred. Additionally, they found that it made no difference if all participants were right-handed (Experiment 1) or left-handed (Experiment 5). Therefore, it can be said that hemispheric dominance related to handedness is not the cause of the direction of SNARC.

The Locus of SNARC

Even if the SNARC effect originates from a mental number line, it is not clear as to how or why this would affect response times for number tasks. Recently, researchers have tried to determine the stage of processing that SNARC originates from. Three possibilities that have been considered are that SNARC originates at an early perceptual stage of processing, or at a more central response selection stage, or at the motor stage.

Some evidence has accumulated to suggest that SNARC originates at an early perceptual stage. Keus and Schwartz. (2005) explored the possibility of a perceptual contribution to SNARC in more detail. They reasoned that if SNARC originates at the perceptual level, then the effect should occur no matter what the mode of response is, because SNARC in this case would be the result of congruency and incongruency of the *side of stimuli presentation* with the numerical representation of the mental number line. Alternatively, the SNARC effect may occur during the response selection stage as the

result of congruencies and incongruencies of *side of response* and the numerical representation on the mental number line. If the latter is the case, it should not matter how stimuli are presented to the participant. Immediately, this hypothesis seems flawed because almost all SNARC experiments have been completed using stimuli presented in the centre of the screen, not to either side. However, Keus et al. wanted to explore the possibility that there was some perceptual influence on the SNARC effect.

They tested participants for SNARC using both laterally presented digits in a parity task, with both vocal and manual responses. When digits were presented laterally, they found no SNARC-like effect when participants made vocal responses, and the SNARC effect that did occur, when participants made manual responses, was unaffected by the side of presentation. These results suggest that SNARC occurs at some point after digits are represented on the mental number line, since they were unaffected by perceptual manipulations. However, the results are not unequivocal because when digits were presented laterally, there was an interaction of digit magnitude and side of response only when the side of presentation and side of response were incongruent, suggesting that side of presentation had *some* kind of effect on responses.

Keus, Jenks and Schwarz (2005) tried to determine if the post-perceptual stage of processing responsible for SNARC was the response-selection stage or the response-execution stage by using electrophysiological measures of brain waves also known as event-related potentials. Participants completed a parity task while electroencephalograph (EEG) recordings were made. EEGs are measures of neural activity, which can be attributed to mental events related to perception or motor responses. EEG recordings were averaged once for the period of time that was time-locked to the stimulus

presentation, but also from a second period, time locked to the manual response and averaging the EEG waves during the time leading up to it. In order to indicate the stage of processing for SNARC, the EEG deflection related to SNARC should be clear for one averaged period and not the other. EEG recordings were analyzed by subtracting mean waveform activity of incongruent trials from mean waveform activity of congruent trials, thus revealing EEG activity specific to SNARC. The EEG deflections indicating SNARC were the most robust during the response stage recordings, indicating that that SNARC originates at this stage. However, it is difficult to tell from this analysis whether SNARC is generated during the response-preparation or response-execution stage.

To make a distinction between SNARC activity at the response-preparation versus execution stages, Keus et al. focussed on another aspect of the EEG recordings, which is a particular wave, called the lateralized readiness potential (LRP). The LRP is specifically indicative of preparation and execution of a motor response and can be seen for any SNARC trial with a manual response. They reasoned that if the origin of SNARC was at the motor execution stage then there should be a latency of the start of the wave's occurrence between incongruent and congruent trials, especially for the response-locked measurements. Alternatively if it did not occur at this stage, then there should be no LRP latency difference between congruent and incongruent trials. Indeed, they found no difference in latency of LRP onset between congruent and incongruent trials, which suggests that the SNARC conflict was resolved prior to the motor execution stage. Overall, their results suggest that the origin of SNARC occurs sometime after the stage of processing for stimulus presentation but before response execution, probably during response selection.

Finally, Fischer (2003a) found evidence to suggest that the internal conflict generating SNARC occurred not only at the response planning stage but also the response execution stage. Participants pointed left or right on a touch screen to indicate “odd” or “even” in response to single digits. Fischer measured the reaction time, which was the amount of time from when a stimulus was presented until the participant lifted his or her finger from a starting pad. Additionally, he measured the movement time which was the time it took the participant to move his or her finger from the pad, once lifted, to the side on the touch screen, indicating their answer. He found that both variables interacted with number magnitude and direction of response, thus demonstrating the SNARC effect. Consistent with the previously mentioned studies, these results suggest that SNARC occurs at least as early as the response selection stage. However, the fact that movement times were not constant, but rather showed a SNARC effect suggests that the mental number line continues to influence the response at a motor stage, because the response had already been initiated.

A Model of SNARC

Gevers, Verguts, Reynvoet, Cassens and Fias (2006) have developed a model of SNARC to explain the difference in right and left response times for digits. Their model has 3 stages. In the first stage, a comparison is made to determine if the number is “small” or “large”. The number that is presented on a given trial is compared to either the digit specified in the instruction for magnitude comparison tasks (e.g. “Is the digit less than or greater than 5?”) or, for tasks in which there is no specified digit to compare (e.g., parity task) the average of all the digits presented in the task is compared to the presented digit. Based on this comparison, a node for “small” or “large” is activated in Stage 2,

depending on whether the presented number is smaller or larger than the standard comparison number. In a magnitude comparison task these are the only nodes that can become activated, but if the task is different, then additional nodes will become active to account for these responses. For instance, during a parity task, the digit will also activate an “odd” or “even” node. From Stage 2 there are 2 routes leading to Stage 3. The automatic route will always lead from either the “small” node or the “large” node to activate a “left response” node or a “right response” node, respectively in Stage 3. The task-dependent pathway, however, will lead to the left or right response node based on the instruction, (e.g., “Press the left key if the number is odd and press the right key if the number is even”). A response occurs when one of the response nodes reaches a threshold. If both routes activate the same response, there will be a fast response, but if the response selected by the automatic route is different than the task-dependent route, there will be a delayed response. This leads to the difference in left and right response times for different numbers that are observed for the SNARC effect.

Gevers et al. have claimed that their model accounts for aspects of the SNARC effect in several ways and they show evidence for these aspects and the fit of the model, using their own data. The first property of SNARC that the model accounts for is that the SNARC effect becomes larger over increasing response times. Second they show that the SNARC effect is categorical for the magnitude task such that the difference in response time between left and right responses tends to be about the same for relatively smaller numbers and also about the same for relatively larger numbers. However, in the parity task, SNARC tends to be continuous, such that the difference in response time between left and right responses gradually changes across magnitude. These points are supported

by their data. Third, their model assumes that magnitude of the number is coded automatically. There is ample evidence to support this point (Fias et al., 2001). Their model also assumes that SNARC is generated during a response selection stage, which seems to be true, as discussed previously. Finally, their model accounts for the results of previous studies (Dehaene et al., 1993) that have shown that the relative rather than absolute magnitude of digits is responsible for SNARC, which will be discussed in the next section

SNARC is contextually-driven and multidirectional

One important aspect of the SNARC effect is that it is dependent on the context in which numbers are presented. As mentioned, Dehaene et al (1993) found that the SNARC effect depended on the relative size of numbers, rather than their absolute magnitude. The numbers 4 and 5 elicited faster right hand responses compared to left when they were part of the set 0-5 but faster left-hand responses compared to right, when they were part of the set 4-9. This suggests that a portion of the mental number line is activated based on the task at hand.

Bächtold, Baumüller, and Brugger (1998) found that when participants were asked to visualize numbers on a ruler during left and right responses to numbers, a SNARC effect was obtained. However when participants were asked to imagine the numbers on a clock face, a reverse-SNARC effect was found, such that leftward responses were now faster for larger numbers and rightward responses were faster for smaller numbers. Although the SNARC effect is generally observed in a left-to-right direction, these results suggest that the mental number line is not fixed but can be reoriented in the opposite direction if a task demands it, as in the animal-size comparison study of Shaki and Petrusic (2005b).

In addition to the horizontal number line, some studies have shown evidence for a vertical spatial representation of numbers. Seron, Pesenti, Noël, Deloche and Cornet (1992) surveyed participants regarding the mental representation of numbers and found that 9 out of 15 reported a bottom to top ordering of numbers from small to large, and one was top to bottom oriented.

Ito and Hatta (2004) also surveyed their Japanese participants to determine if they possessed a vertical representation of numbers. Fifty participants were asked to orient the numbers 0-9 to on a vertical line. They found that 36 participants reported a bottom to top orientation and 9 reported a top-to bottom orientation and 5 were undetermined. From this they hypothesized that most of their participants had a bottom to top orientation of numbers and that this may be reflected in the SNARC effect.

They tested participants in a parity task in which they were required to press buttons vertically to indicate their responses. Indeed, they found a SNARC effect representative of a number line, which was directed from bottom to top. Participants were quicker to make lower responses when the numbers were smaller but quicker to make upper responses when the digits were larger.

At around the same time Schwarz and Keus (2004) argued that SNARC may occur vertically as a separate mental number line, in addition to the horizontal number line, or may be a reflection of a single mental number line that can change position in response to the relevant task. Indeed they found a vertical SNARC effect, using saccadic eye movements as responses in a parity task. Replicating Ito and Hatta, their results showed that smaller digits facilitated an advantage for bottom responses and larger digits for the top, supporting the notion of a vertical mental number line. Again these results suggest

that the mental number line is flexible and can be reoriented. Alternatively, participants may orient numbers on an array or map of numbers rather than a single line.

SNARC for Two-Digit Numbers?

If there really is a mental number line, it seems logical that it would include multi-digit numbers. As mentioned, Dehaene et al (1990) did find that participants who made comparisons of two-digit numbers to the reference number (55) were faster when responding to the left for “smaller” and to the right for “larger”, compared to the alternate mapping. Although this was not a direct test of the SNARC effect, it does suggest that the mental number line includes two-digit numbers.

Lucas and Petrusic (2006, unpublished) found SNARC using a two-digit numbers in a memory task. On each trial, participants were presented with a picture of an unseen face and a face from a previously studied set and were required to choose the previously seen face by pressing a left or right key. Presented in the middle of the screen for each trial was an irrelevant two digit number, which was not mentioned in the instructions. Participants showed a robust SNARC effect when picking the familiar face, however, it occurred when SNARC was regressed across the units digit of each number, but not the whole number. Unfortunately, in several attempts to replicate the results from my study I have not been successful. My initial results contrast with the study of Dehaene et al. (1990), which suggests the whole two-digit number may lie on the mental number line. It seems possible that the mental number line includes two-digit numbers, although there is little in the literature to support this.

Does the mental number line include negative numbers?

Few studies have been done to test if SNARC exists for negative numbers, which would indicate that the mental number line extends past zero. In fact the literature on negative numbers, in general, is sparse. SNARC effects generated by negative numbers would be helpful in understanding how they are processed. A question of interest is whether the magnitude of negative values is processed according to their absolute value (i.e., polarity is ignored) or according to their numerical value, such that, the more negative a number gets, the smaller in magnitude it becomes. Fischer (2003b) attempted to answer this question. On each trial in his study, participants chose the larger or smaller of two numbers by pressing a left or right key. The pairs were spatially congruent with the direction of the mental number line, (-8, -3) or incongruent (-3, -8). Responses for negative pairs on congruent trials were faster than incongruent trials, suggesting that participants viewed negative numbers on a mental number line that extends past zero.

Shaki and Petrusic (2005a) questioned the results of Fischer, explaining that it was possible that he could have attained his results due to a confounding semantic congruity effect (to be discussed later in this Introduction). Additionally, Fischer had found a reverse SNARC effect for positive number pairs, which challenged the validity of his results because it was an unexpected result that was difficult to explain.

Fischer and Rottmann (2005) found that participants seemed to treat negative numbers like positive numbers during a parity task. SNARC for both polarities was reflective of processing of numbers according to their absolute magnitude.

Shaki and Petrusic further investigated SNARC for negative numbers, using pairs in a comparison task. Unlike the experiment of Fischer, pairs were constructed such that, each

digit was always paired with its adjacent digit. Therefore, much like single digits, the pairs in a way, acted as single stimuli that ranged from small to large. Pairs were classified according to their numerical value as opposed to their absolute value as relatively small negative (e.g., -8, -9), large negative (e.g., 0, -1), small positive (e.g., 0, 1) and large positive (e.g., 8, 9). On each trial, participants were required to choose the “Larger” or “Smaller” number, according to numerical magnitude.

Participants completed both a blocked condition in which they only completed positive-number trials or negative-number trials, in a single block, as well as a mixed condition in which trials randomly alternated between positive and negative pairs within a single block. They found that when participants made judgements of just negative numbers or just positive numbers within a block, a SNARC effect was found for both which was dependent on the absolute value of the numbers. It was as if polarity was ignored and the direction of response was facilitated by the absolute magnitude of the pairs. However, when negative digits were mixed with positive digits, the direction of the SNARC effect for negative pairs was now dependent on polarity. That is, for negative pairs, left responses became faster, compared to right, as the pairs became more negative, but right responses became faster, as the pairs approached zero.

These results support what Fischer had concluded, that the mental number line extends leftward past zero. However, this result only occurs when negative and positive numbers are mixed together. When participants are tested with negative numbers alone, they seem to treat them like positive numbers. The question immediately becomes, “Why is polarity seemingly unimportant for the direction of SNARC, when all the digits in the pairs are negative?” Of course, in actuality, participants cannot simply ignore polarity because it is

essential to correctly choose the “Smaller” or “Larger” of pairs such as (-8, -9). However, Shaki and Petrusic suggested that if polarity is known ahead of time, as it is in the blocked condition, polarity doesn’t need to be processed immediately with magnitude and therefore potentially, may not affect the SNARC effect. Polarity may be processed at a later stage of processing, after SNARC processing has occurred. Additionally, it can be argued that if a participant knows ahead of time that all the forthcoming trials will be negative, then an efficient rule can be adopted, which participants may be unaware of, to choose the digit of smaller absolute value when the instruction is “Larger”, or the digit of larger absolute value when the instruction is “Smaller”. The idea though that processing of polarity may come after the processing of SNARC may contradict the evidence that suggests that SNARC processing occurs at a response selection stage. However it may be possible that initially magnitude is processed at the response selection stage followed by polarity in quick succession.

Recently, Ganor-Stern and Tzelgov (2008) attempted to determine if in fact polarity of negative numbers is processed holistically with magnitude or if it is possible that magnitude and polarity are processed as separate components. In the first task participants compared positive and negative pairs made up of numbers ranging from 3 to 97 or -3 to -97. Replicating the results of Shaki and Petrusic, they found a magnitude effect, which is when response times increase as the absolute magnitude of the compared numbers increase, while maintaining a constant difference between them. Because a magnitude effect existed for negative numbers, it suggested that participants were processing them according to their absolute magnitude. Ganor-Stern and Tzelgov argued that this is evidence that polarity and magnitude are processed as separate components.

The Semantic Congruity Effect

In addition to the SNARC effect, an important effect related to magnitude of numbers and stimuli is the semantic congruity effect. This effect is a result of an interaction between *instruction* and *relative magnitude of pairs* and demonstrates that symbols are processed in the mental context of a linear ordering. Unlike the SNARC effect it theoretically, has nothing to do with the direction of responding. Further, it exemplifies that specific instructions determine a context, which further affects the ease, or difficulty (i.e. speed) in which a comparison can be made, depending on the relationship of the instruction and the location of a stimulus on the dimension of interest.

Stated differently in relation to number comparisons, the semantic congruity effect is defined by the decrease in response time that occurs on trials in which the instruction and relative size of the digits in a pair are congruent, compared to trials in which the instruction and relative size of the pair are incongruent. For instance, in a subjective phenomenal study, Shipley, Coffin and Hadsell (1945) reported that participants judged the more pleasant of a pair of relatively preferred colours faster than the less pleasant colour. However, in a subsequent study Shipley Norris and Roberts (1946) showed participants were faster to report the more unpleasant colour when the pair was relatively less preferable. Therefore whether the pleasant or unpleasant colour was selected faster depended on the relative level of the preference of the pair.

In the first psychophysical study reporting a semantic congruity effect, Audley and Wallis (1964) found that when participants were asked to choose the brighter of two lights, responses were faster when the pair was relatively bright compared to when the lights were dimmer. However, when choosing the darker of 2 lights, responses were

faster when the pair was relatively dark compared to bright. They repeated this finding using pitch discrimination. Participants were faster to choose the higher of two high frequency tones compared to the lower, but faster to choose the lower of two low frequency tones, compared to choosing the higher (Wallis and Audley, 1964). They named this effect the “cross-over” effect because when graphing response times for each instruction across the range of magnitude, the lines would actually crossover.

Banks, Clark and Lucy (1975) found that participants who were required to choose the higher or lower circles attached to lines which were presented as “yo-yos” or “balloons” were faster to choose the higher of the balloons compared to the lower, but faster to choose the lower of the yo-yos compared to the higher. Banks reasoned that a congruity effect occurred because participants thought of balloons as travelling up and were therefore congruent with the instruction “up” but yo-yos were going down and were therefore congruent with the instruction down. They argued that the stimulus was coded according to their semantic meaning, such that balloons were coded as “High:High+” and yo-yo’s were coded as “Low:Low+”. The plus sign indicates an increase in the dimension for one stimulus over the other. In a separate condition, when participants were required to compare the longer or shorter string length, which was semantically neutral, they found no semantic congruency effect. This suggested that the semantic congruity effect only occurred when the perceptual stimuli represented an underlying linguistic ordering.

Thus the “crossover effect” has since become known as semantic congruity effect. The name “semantic congruity effect” is better than the “crossover effect” because it indicates the nature of the interaction but also includes instances of the interaction when the lines do not actually crossover on a graph. This may occur because responses for one

instruction can be systematically faster than the other, resulting in what was formerly known as the “funnel effect”.

The semantic congruity effect has since been found in a variety of situations including pictorial comparisons of age (Ellis, 1972), skin colour (Friend (1973), audible comparisons of loudness (Banks and Root, 1979), numbers (Banks, Fujii, & Kayra-Stuart, 1976), and words representing length, quality, and temperature (Holyoak and Walker, 1976). The semantic congruity effect has also been previously found using the instructions “Smaller” and “Larger” for animal names, (Jamieson and Petrusic, 1975, Banks and Flora, 1977) as well as animal pictures (Banks and Flora, 1977).

There has been some debate as to whether or not SCEs occur for purely perceptual stimuli (e.g. line length) or only for stimuli that are symbolic of a remembered dimension (e.g., numbers, CVCs, etc.) Specifically, in a series of five experiments, Marschark and Paivio (1981) failed to find semantic congruity effects for comparisons of squares and circles. Additionally, Banks, Mermelstein and Yu (1982) found a semantic congruity effect with CVCs associated with different circle sizes, but not with the circles themselves. They argued that the processing of symbolic and perceptual stimuli is fundamentally different, and that physical stimuli will not generate a semantic congruity effect unless they are conceptually ordered according to a semantic dimension. Although, Audley and Wallis (1964) and Wallis and Audley (1964) had already found semantic congruity effects using brightness of light and pitch, Banks et al. (1982) argued that participants could have based their decisions on mappings of these stimuli onto the corresponding dimension, with linguistic codes, as described in Banks’s model (described later), and therefore were not necessarily purely physical comparisons.

Petrusic and Baranski (1989) have since debated that discrepancy of stimuli may have been an important factor in previous studies because all of the experiments that failed to find semantic congruity effects had accuracy rates close to 100%. They did a series of experiments in which participants made judgements on perceptual dimensions including line length, and weights. They found that when accuracy was near perfect there was little evidence of a semantic congruity effect but for less discriminable pairs there was in fact a robust SCE. Discriminability seems like the best explanation for why comparisons with physical stimuli sometimes elicited a semantic congruity effect and sometimes did not. Semantic congruity effects are clearly attainable with physical stimuli and there is no clear reason why participants would use linguistic codes when the stimuli are less discriminable but not do so when the stimuli are more discriminable, as Banks' argument would suggest.

The importance of sleep to consolidate memory

In my first experiment, two sessions occurred for each participant on separate days to allow participants to sleep in between training, in order to promote memory consolidation. Therefore some discussion on the relationship between sleep and memory is needed. There is an abundant amount of research that suggests that sleep is important for memory consolidation. The first of two basic forms of sleep is rapid eye movement (REM) sleep, which is characterized by high frequency, low amplitude oscillations on EEG (electroencephalograph) recording as well as muscle atonia and ocular saccades. The second form of sleep is slow wave sleep (SWS), which is characterized by less frequent, higher amplitude oscillations on EEG recordings. It has been shown that it is SWS stages

of sleep that are particularly important for declarative memory and REM stages which are particularly important for procedural memory (Plihal & Born, 1997).

Sleep is hypothesized to facilitate the consolidation of memory traces by means of promoting synaptic plasticity (see Maquet, Smith, & Stickgold, 2003 for a review). Hypothetically, exposure to new events creates new memories, but they exist in a fragile state. It is during sleep that the memories are incorporated into longterm memory.

Many studies have shown the harmful effects of sleep deprivation on the effects of learning. Ekstrand (1967) first demonstrated that participants, who were allowed to sleep during an 8 hour period after learning word lists, displayed better recall compared to participants who spent 8 hours awake after learning. Similar types of studies have demonstrated the deleterious effect on recall for participants with a disrupted sleep cycle, compared to those who were undisrupted (Ficca, Lombardo, Rossi & Salzarulo, 2000).

Stickgold, Whidbee, Schirmer, Patel and Hobson (2000) tested participants on a visual discrimination task, in which participants typically improve slowly with practice. After a single training session participants were retested after varying amounts of time, with or without sleep. Only participants who attained at least 6 hours of sleep after the first session performed better in the second session. In fact, the amount of improvement was proportional to the amount of slow-wave sleep in the first quarter of the night, and the amount of REM sleep in the last quarter. Participants who performed the task on the same day as training, without sleep, did not significantly improve on the task.

In another study, conducted by Wagner, Gais, Haider and Born (2004) participants were trained on a cognitive task involving strings of digits in which they were required to follow assigned rules, in order to come up with a final solution. Participants became

faster with practice, but they could become much faster suddenly, if they realized an underlying rule that was not taught, but could be applied to all trials. In the next session, more than twice as many participants were able to apply the underlying rule after 8 hours of sleep compared to 8 hours of wakefulness. Therefore sleep is not only helpful for consolidating remembered items but also for implied rules on a task.

The Current Thesis

The conditional boundaries of the SNARC effect have been largely unexplored. The experiments in this thesis will explore and expand on several aspects of the SNARC effect discussed, including instruction-dependent SNARC for symbolic stimuli, and vertical as well as horizontal SNARC for negative and positive numbers in the context of temperature. It is still unknown what causes the direction of SNARC. What are the conditions that determine if stimuli are ordered in one direction or another? Number experiments seem to suggest that participants order stimuli from left to right, smallest to largest. However, the ordering of numbers is so overlearned that they might act differently than other stimuli. As the results of the clock/ruler experiment of Bächtold, Baumüller, and Brugger (1998) showed, it could be that SNARC can be reversed when the semantic quality of the numbers change. Additionally, Shaki and Petrusic (2005b) showed that SNARC can occur in opposing directions for forms of magnitude other than numbers. This research will try to answer the following questions:

1. What causes the direction of SNARC?
2. When numbers take on semantic properties, what influence does instruction have on the direction of SNARC, particularly for negative numbers?

3. Why did SNARC change direction for negative numbers between blocked and intermixed conditions in Shaki and Petrusic (2005a)?
4. Does SNARC occur in the vertical plane for negative numbers?
5. SNARC can occur for magnitude but what about an abstract quantity such as intelligence?

The first two experiments will expand on the results of Shaki and Petrusic (2005b), using animal size and animal intelligence but no numbers. As mentioned, SNARC had been found for symbols of magnitude other than numbers, but the direction of SNARC in those experiments was dependent on the instruction. It is unclear why this instruction-dependent effect exists.

In the first experiment of my thesis, participants were tested for SNARC using animal pictures representing various magnitudes, as stimuli. After a baseline SNARC test, they were given practice trials on a task, encouraging the development of a mental linear ordering of animal size. In a second session, the practice trials were repeated and then a second SNARC test was conducted. Participants were tested on separate days, with at least one day separating experimental sessions. The reason for this was to ensure that participants had at least one night of sleep. I reasoned that having at least one night of sleep could increase the chance that the ordering of animal size presented during the training phase would be effective in the subsequent SNARC test. I predict that in the first session SNARC will be dependent on instruction, but in Session 2, SNARC will be either independent of instruction or change significantly in a direction that reflects a shift towards independence of instruction.

In Experiment 2, participants made comparisons of animal pairs, with level of intelligence, based on ratings, acting as the dimension of magnitude. I expected to find SNARC results similar to that of Shaki and Petrusic (2005b). This experiment is important because even though SNARC can be found for other forms of magnitude, it is unclear how important magnitude is to attain SNARC. Gevers et al. (2003) found SNARC for letters of the alphabet, but Dehaene et al. (1993) did not. A SNARC effect for ordering of animal intelligence would show that a broad definition of magnitude can be used for the SNARC effect.

The remaining four experiments consisted of experiments with positive and negative numbers in a horizontal as well as a vertical plane. In Experiment 3, participants were required to choose the “warmer” or “colder” from pairs of relatively small or large positive and negative pairs of numbers presented as temperatures. This experiment was designed to replicate the results of Shaki and Petrusic in the context of temperature.

Experiment 4 was a replication of Experiment 3, except that stimuli were presented vertically and participants were required to make responses on keys positioned vertically. Participants were required choose the “warmer” or “colder” of digits from relatively cold or warm temperatures. SNARC for negative numbers in a vertical plane have not been previously found. In one condition, a picture of a thermometer was be presented on trials to see if it elicited codes in long-term memory regarding numerical position, enough to enhance or alter the SNARC effect.

Experiment 5 was a replication of Experiment 4 except that I changed the way the instruction was presented to try to augment the SNARC effect.

Experiment 6 was much like Experiment 3 because responses and number presentation were both in the horizontal plane. However, participants were required to choose the “Faster” or “Slower” of positive and negative numbers representing runners of various speeds. The experiment was designed to test if the actual magnitude of number was important or if it was the perceived value of a number that determined SNARC.

Experiment 1

In a pilot study I tested to see if I could find SNARC, using pictures representing animals of various physical size. I was particularly interested to see if the SNARC effect would be instruction-dependent like other forms of magnitude (e.g., line length) or instruction-independent like that of numbers, both results found by Shaki and Petrusic (2005b). I hypothesized that the reason symbols of magnitude were different from numbers is because numbers have an overlearned orientation of magnitude spanning from left to right. Generally, this orientation seems to be robust enough that it is not affected by instruction. However, because other forms of magnitude do not have an overlearned sequence of magnitude, they seem to be influenced by instruction such that, instruction acts as an anchor, invoking a left-to-right orientation when the instruction is “Smaller”, but a right-to-left orientation when the instruction is “Larger”. On each trial in the practice session of my experiment, I included an array of animal pictures across the top of the screen, presented from left-to-right, with the smallest animal on the left and the largest on the right. Below the array were two of the animals’ names to the left and right of centre, and participants were required to choose the larger or smaller of the two animals, depending on the instruction given. In the test session, I repeated the experiment without the animal array across the top. I hoped that by including this array during

practice trials, it would create a lasting orientation of magnitude similar to that of numbers, and the SNARC effect would be independent of instruction. However, the SNARC effect I found was still instruction-dependent.

I realized that simply having the array of animals across the top of the screen probably did not have much influence on the orientation of magnitude because the array could simply be ignored as participants concentrated on the relevant stimuli.

In this experiment, I replicated my pilot experiment except for including a practice session prior to SNARC testing. The practice session was designed to try to force participants to interact with the animal array such that they would have to concentrate on its ordering which I hoped would result in memorization of the orientation of the array. In the first session participants were tested for SNARC as a baseline and then completed the practice phase. In the second session, which occurred on a second day, participants were required to complete the practice session again and then tested for SNARC using both the animal pictures and the names of the animals that appeared in the array. Only data for the test phases were analyzed.

Method

Participants

Thirty-eight Carleton University students participated in two sessions, each on a separate day, lasting approximately 45 minutes each, in exchange for course credit.

Stimuli and Design

The seven animal pictures used in the experiment include: ant, bee, rat, cat, dog, hog, and cow. The coloured pictures were an artist's rendition of the animals and were collected from various websites. The pictures were presented such that the physical size

of the picture was smallest for the smallest animal, progressively getting larger for the largest animal. This scaling, however, is not intended to represent a ratio true to life but rather used to avoid any confusion as to the ordering of the magnitude (e.g., the cat appeared smaller than the dog).

This experiment consisted of a practice phase and a test phase, and both occurred in two separate sessions. During the practice phase, animal pictures were presented across the upper third of the screen, from smallest to largest. On each trial, a red arrow appeared under each of two animals with the name of one of the animals printed below, in the centre of the screen. The participants' task was to indicate whether the named animal is the left of or right animal of the two pictures indicated by red arrows. Participants press a key to indicate the animal on the left or right. The next trial followed 500 ms after the participant made a response. Every combination of animal was presented an equal number of times. There were seven animals and 21 possible combinations, therefore there were 42 trials in a single block and this was repeated 20 times for each of the 2 practice sessions.

In the test phase, the array of animals did not appear across the top and instead, pairs of animals (pictures in one condition and animal names in another), were presented (left and right) in the upper third of the screen, with an instruction below ("smaller" or "larger"). Only adjacent pairs of animals were presented for comparison and include (ant-*bee*; *bee*-rat; rat-cat; cat-dog; dog-hog; and hog-cow, as well as the reversed ordering: bee-ant, rat-*bee*, etc.). Each block was repeated 12 times and consisted of 24 random trials consisting of six each pairs, in two orders, with two instruction (larger, smaller). The first two blocks of each test phase were treated as practice. Half of the participants

were tested using animal names in the first half of the session and pictures in the second, and half the participants were tested with pictures first, and then names.

The pairs appeared on the left and right of centre on the monitor in black, Arial font (36) in the approximate upper third of a white background on a 17-inch (43 cm) View Sonic monitor. The experiment was programmed and executed using SuperLab 2 software on a PC with Pentium III microprocessor. The keypad used was the Cedrus model RB-620 (keys 2 and 5).

Participants completed two sessions with at least 24 hours in between. Participants were tested for SNARC twice, once in each session. In the first session, participants completed the test phase first to attain a baseline, followed by a practice session. In the second session, participants completed the practice phase followed by the test phase. The final test phase was used as an indication of whether or not two practice sessions had any effect compared to the baseline.

Procedure

For each trial, the instruction appeared on the screen and the pair of pictures followed 500 ms later. Participants were required to make their decision by pressing the left key to indicate the animal on the left or the right key to indicate the animal on the right. The instruction and names remained on the screen until participants indicated their decision. The next trial will followed 500 ms later. Participants were encouraged to respond accurately but also quickly.

Results

Prior to analysis, data were eliminated from eight participants who attended the first session but not the second. Therefore the data used in my analyses is from a total 30 participants. Additionally, the first two blocks of data were deleted to eliminate data. The analysis for Session 1 is presented first, followed by Session 2. The analysis for each session consists of an analysis of variance (ANOVA) based on aggregated median scores followed by a regression analysis based on beta weight for each individual participant.

Session 1

Response time

After combining the data over the two presentation orders, medians for each cell, from each participant were averaged. Cells that exceeded the mean plus three standard deviations were replaced with the mean plus three standard deviations for that condition. Cells that were missing were replaced with (the subject mean + the cell mean – the grand mean). A total of 37 medians were replaced. An ANOVA was then conducted with mean correct response time as the dependent variable and the six animal pairs, two stimulus types (pictures, words), the two instructions (smaller, larger), and two response buttons (left, right) as the within participant factors. Significance levels for analysis of variance (ANOVA) were based on the Huyn-Feldt adjusted degrees of freedom, although the degrees of freedom and error mean squares provided are those specified by the design. The level of significance used for all tests was $\alpha = 0.05$.

The ANOVA revealed a significant effect of stimulus type, $F(1, 29) = 204.671$, $MSE = 901807.181$. As seen in Figure 1, participants responded to pictures (834.88 ms) much faster than words (1550.91 ms). The main effect of magnitude was significant, $F(5, 145)$

= 10.581, $MSE = 135834.619$; response times seemed to change in an idiosyncratic manner across magnitude of stimulus. As seen in Figure 1, this unpredictable pattern was unique for words and pictures, evident from the interaction of magnitude and stimulus type (words, pictures), $F(5, 145) = 9.392$, $MSE = 119283.847$.

Also evident in Figure 1, generally, participants took longer to respond to the instruction “smaller” ($M = 1221.94$), compared to “larger” ($M = 1163.85$ ms), $F(1, 29) = 9.574$, $MSE = 126849.062$.

Semantic Congruity Effects

Figure 1. Mean response times for each animal pair, for each instruction in Session 1, with the smallest pair at 1 (ant-bee), progressing to the largest at 6 (hog-cow).

Semantic congruity effects were present for pictures and words, as seen in Figure 1 and confirmed by the significant interaction between instruction and magnitude of stimulus pair, $F(5, 145) = 8.209$, $MSE = 77489.145$. The semantic congruity effect is evidenced by the fact that for the smallest pair, participants were faster to respond when the instruction was “smaller” compared to when it was “larger”. However, as pair size increased, participants were progressively faster to respond when the instruction was “larger” compared to “smaller”.

The interaction between instruction and response button (left, right) was also significant, $F(1, 29) = 9.777$, $MSE = 93902.947$. Left responses were similar for the instruction “smaller” ($M = 1203.179$ ms) and “larger” ($M = 1195.597$ ms), but right responses showed considerably slower responses for the instruction “smaller” ($M = 1240.694$ ms), compared to “larger” ($M = 1132.110$ ms).

There was also an interaction of stimulus type (pictures, words) and instruction $F(5, 145) = 5.946$, $MSE = 56143.550$. The difference in responses time was greater between the instructions smaller and larger for words ($M = 88.536$ ms) compared to pictures ($M = 27.631$ ms).

There was also a three-way interaction between instruction, magnitude and response button $F(5, 145) = 4.312$, $MSE = 63887.934$. This is potentially important because it is indicative of a SNARC-like effect, however, it is for pictures and words combined. The means show that when the instruction was “smaller”, generally, right-hand responses ($M = 1192.86$ ms) were faster than left-hand responses ($M = 1210.00$ ms) when the pairs were small but left-hand responses ($M = 1196.35$ ms) were faster than right (1288.53 ms) when the pairs were larger. When the instruction was “larger”, however, the pattern changed.

Right-hand responses ($M = 1077.26$ ms) were faster than left ($M = 1158.78$ ms) when the pairs were larger, but when the pairs were smaller right-hand responses ($M = 1186.95$ ms) were still faster than left ($M = 1232.30$ ms), but the difference was attenuated. There were no other significant results.

Individual participant analyses

Next I analyzed the possibility of a SNARC effect in more depth by performing one-sample t-tests of the standardized beta coefficients for each participant, in each condition as described by Lorch and Meyers (1990). After combining the data over order, I first obtained standardized beta regression coefficients for each participant with the SNARC index (right button RT – left button RT) as the dependent variable, and magnitude of stimulus pair as the predictor. I then tested the hypothesis (two tailed) that the mean of these weights differed from zero using the t-distribution. R^2 values were calculated using group means rather than individual means.

As seen in Figure 2, there was a significant SNARC effect for words when the instruction was “smaller”. The mean standardized beta weight ($M = .193$, $SD = .507$), differed reliably from zero, $t(29) = 2.083$, $R^2 = .510$. The direction of SNARC was such that participants responded faster to the right compared to the left when the pairs were small, but faster to the left compared to right when the animals were larger. This SNARC effect is in the opposite direction of what is typically observed for numbers and opposite of what Shaki and Petrusic found for animal pairs, with the instruction “smaller”. The SNARC effect fell short of significance when the instruction was “larger”, ($M = -.082$, $SD = .495$), $t(29) = -0.906$, $p = .37$, $R^2 = .068$.

Figure 2. Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of words representing animal size, in Session 1. The smallest pair is represented by 1 progressing to the largest pair represented by 6.

Figure 3. Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair of pictures representing animal size, in Session 1. The smallest pair is represented by 1 progressing to the largest pair represented by 6.

As can be observed in Figure 3, there were no significant SNARC effects for pictures, either with the instruction “smaller” ($M = .041$, $SD = .495$), $t(29) = .451$, $p =$

.66, $R^2 = .248$, or with the instruction “larger” ($M = .060$, $SD = .462$), $t(29) = .715$, $p = .48$, $R^2 = .010$.

Error Rates

The error rate for Part 1 was 5.8% or (5.36% with the aggregate). The correlation between mean RT and mean error rate for the 48 cells representing stimulus type, instruction, magnitude and button, is .519, $p < .001$. Higher accuracy was associated with lower response times, therefore there was no speed-accuracy trade-off.

An ANOVA was performed on mean error rates as the dependent variable and stimulus type, instruction, magnitude and button as the independent variables. There was a main effect of stimulus type, $F(1, 29) = 15.210$, $MSE = .637$. Participants made over twice as many mistakes with words (7.46%) compared to pictures (3.25%). There was also a main effect of magnitude, $F(5, 145) = 5.546$, $MSE = .389$, as well as an interaction between stimulus type and magnitude, $F(5, 145) = 5.002$, $MSE = .190$. Error rates differed in an idiosyncratic manner across magnitude for both words (5.76%) and pictures (21.00%) although generally, there were 13.38% errors for the second largest pair comparison (hog, dog), which were at least twice as many mistakes compared to any other pair. It seems as though the phonetic familiarity of hog and dog may have been particularly confusing. There were also significant reactions for magnitude and response button $F(5, 145) = 3.057$, $MSE = .004$, as well as for stimulus type, instruction and response button $F(1, 29) = 4.694$, $MSE = .002$.

The interaction between instruction, magnitude and response button was significant, $F(5, 145) = 3.031$, $MSE = .003$, indicating a SNARC effect, which was different between instructions. When the instruction was smaller, responses with the left-hand ($M = .037$)

were more accurate than the right-hand ($M = .051$) for small pairs, but for larger pairs, responses with the right-hand ($M = .063$) were more accurate than the left hand ($M = .078$). However, when the instruction was “larger” the opposite pattern occurred. For smaller pairs, right-hand responses ($M = .033$) were more accurate than left-hand responses ($M = .041$) but for larger pairs, left-hand responses ($M = .063$) were more accurate than right-hand responses ($M = .065$). This effect shows that the speed advantages that were observed in the response time analysis also resulted in more errors, indicative of a speed-accuracy trade-off in these conditions.

Session 2

Participants returned for a second session, at least 24 hours later, but no later than 1 week to complete another training session followed by the SNARC test. It was assumed that participants slept in between sessions although no sleep data was recorded. The ANOVA and regression analyses were carried out exactly the same as they were for the Session 1 data. A total of 31 medians were replaced prior to analysis.

The ANOVA revealed main effects and interactions similar to the first session. Like the first session, responses for the second session were significantly faster for the instruction “larger” ($M = 1001.804$ ms) compared to “smaller” ($M = 1045.098$ ms), $F(1, 29) = 5.536$, $MSE = 121880.694$. Response times varied significantly in an idiosyncratic manner across magnitude $F(5, 145) = 11.245$, $MSE = 112585.273$. Furthermore, magnitude interacted with instruction revealing a significant semantic congruity effect similar to that of the first session data, $F(5, 145) = 15.060$, $MSE = 59805.110$. For the smallest pair, responses were faster when the instruction was “smaller” compared to “larger”, but became faster for the instruction “larger” as the size of the pair increased.

Figure 4. Mean response times for each animal pair, for each instruction in Session 2, arranged with the smallest pair at 1 (ant-bee), progressing to the largest at 6 (cow-hog).

As seen in Figure 4, as expected response times for pictures ($M = 742.63$ ms) were significantly faster than for words ($M = 1304.27$ ms), $F(1, 29) = 116.90$, $MSE = 971402.628$. The interaction between stimulus type (pictures, words) and magnitude was significant, $F(5, 145) = 14.091$, $MSE = 91103.267$. As can be observed in Figure 4, response times seemed to vary in an unpredictable manner for both pictures and words. There was also an interaction between instruction and response button $F(1, 29) = 7.984$, $MSE = 61004.377$. Left response times were comparable for the instructions “smaller” ($M = 1033.36$ ms) and “larger” ($M = 1056.84$ ms) but for right-hand responses, times were

considerably slower with the instruction “smaller” ($M = 1026.85$ ms) compared to “larger” ($M = 976.76$ ms).

The three-way interaction between stimulus (picture, word), instruction (smaller, larger) and magnitude $F(5, 145) = 3.506$, $MSE = 58612.633$ is potentially important because it indicates that the semantic congruity effect I found differs somewhat between words and pictures in Session 2. As seen in Figure 4, the semantic congruity effect was greater for words compared to pictures.

Individual participant analyses

To investigate the possibility of a SNARC effect further, I performed one-sample t -tests of the standardized beta coefficients for each participant, in each condition, the same as the first session.

In contrast to the results of Session 1, the regression analyses revealed that none of the beta weights for each condition were significantly different from zero. Thus, as seen in Figure 5, the SNARC effect evident in the first session, for words when the instruction was “smaller” was not repeated, ($M = -.095567$, $SD = .378739$), $t(29) = -1.382$, $p = .18$, $R^2 = .298$. Again SNARC was also not found for words with the instruction “larger”, ($M = -.023600$, $SD = .406291$), $t(29) = -.318$, $p = .75$, $R^2 = .000$.

Figure 5. Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of words representing animal size, in Session 2. The smallest pair is represented by 1 progressing to the largest pair represented by 6.

Figure 6. Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of pictures representing animal size, in Session 2. The smallest pair is represented by 1 progressing to the largest pair represented by 6.

Figure 6 displays the nonsignificant results for pictures with the instruction “smaller” ($M = -.087833$, $SD = 0.487889$), $t(29) = .986$, $p = .33$, $R^2 = .338$; or “larger” ($M = -.015433$, $SD = .500473$), $t(29) = .169$, $p = .87$, $R^2 = .000$.

Comparison of Sessions 1 and 2

Although, out of all conditions, there was only one significant SNARC effect for words with the instruction “smaller”, in Session 1, it is possible that the difference in slopes between Sessions 1 and 2 may be significantly different than zero for other conditions. To test this, the standardized beta weights from Session 2 were subtracted from those of Session 1 for each participant. A paired-samples t-test was then performed on these scores to test if they differed significantly. The results showed no significant change in slope from Session 1 to Session 2 for pictures with the instruction “smaller”, ($MD = -.04713$, $SD = 0.69468$), $t(29) = -.372$, $p = .71$ or for the instruction “larger”, ($MD = .04490$, $SD = 0.61135$), $t(29) = .327$, $p = .75$. For words there was no significant change when the instruction was “larger” ($MD = -.05833$, $SD = 0.66523$), $t(29) = -.480$, $p = .64$. However, when the instruction was “smaller”, the change in slope from Session 1 to 2 was significant ($MD = -.28833$, $SD = 0.61135$), $t(29) = 2.583$, $p < .02$. This result confirms that participants shifted from faster right hand responses compared to left, for relatively small pairs, and faster left hand responses compared to right for relatively large pairs in Session 1, to faster left hand responses for smaller pairs, and faster right hand responses for larger pairs, in Session 2.

Error Rates

The error rate for the data in Session 2 was 6.05 %. The correlation between mean RT and mean error rate for the 48 cells representing stimulus type, instruction, magnitude and button, was .555, $p < .001$. Higher accuracy was associated with lower response times; therefore there was no speed-accuracy trade-off.

An ANOVA was performed on mean error rates as the dependent variable and stimulus type, instruction, magnitude and button as the independent variables. The same effects were significant for Session 2 as for Session 1. There was a main effect of stimulus type, $F(1, 29) = 10.725$, $MSE = .075$. Again, participants made over twice as many mistakes with words (8.41 %) compared to pictures (3.70 %). There was also an interaction between stimulus type and instruction, $F(1, 29) = 5.674$, $MSE = .007$. For pictures there were more errors for the instruction “smaller” (4.10 %) compared to “larger” (3.33 %), but for words, there were more errors for the instruction “larger” (9.10 %) compared to “smaller” (7.74 %). There was also a significant interaction for stimulus type and response button, $F(1, 29) = 7.986$, $MSE = .003$. For pictures there were more errors for left responses (4.12 %) compared to right responses (3.30 %), but for words there were more errors for right responses (8.80 %) compared to left (8.02 %).

There was also a three-way interaction for stimulus type (words, pictures), instruction (smaller, larger) and magnitude, $F(5, 145) = 2.519$, $MSE = .007$. For words, when the pairs were relatively large, responses were more accurate when the instruction was “smaller” ($M = .083$) compared to “larger” ($M = .103$). However, when the pairs were relatively smaller, the accuracy advantage of the instruction “smaller” ($M = .072$) over “larger” ($M = .079$) decreased substantially. This suggests that the increased response times associated with the semantic congruity effects observed in Session 2 were associated with lower accuracy, indicating a speed-accuracy trade-off in these conditions.

For pictures, however, a different pattern emerged. When the pairs were relatively large, responses were more accurate when the instruction was “larger” ($M = .036$) compared to “smaller” ($M = .049$) but when the pairs were relatively smaller, the accuracy

advantage of the instruction “larger” ($M = .031$) over “smaller” ($M = .032$) disappeared. This indicates that for pictures the decreased response times associated with the semantic congruity effects also resulted in higher accuracy rates.

Discussion

This experiment failed to achieve my primary goal, which was to demonstrate that through the process of learning, types of magnitude other than numbers, could generate SNARC effects that were *independent* of instruction. I expected that in the initial session, I would find SNARC effects that were dependent on instruction, as did Shaki and Petrusic (2005b). In the second session, I expected all of the SNARC effects to reflect an ordering of smallest on the left and largest on the right no matter what the instruction was, as is typically observed with numbers.

In the initial session I did in fact find a SNARC effect, dependent on instruction, for words when the instruction was “smaller”. Based on the direction of SNARC, I would assume that participants had mentally ordered animal pairs with the smallest on the right and the largest on the left, for that condition. This seems to be inconsistent with the results of Shaki and Petrusic (2005b). They found that with the instruction “smaller” participants seemed to order animal size (also presented as words) and line-length, with the smallest on the left and the largest on the right, the opposite of what I found. Shaki and Petrusic (2005b) hypothesized, that instruction seems to act as an anchor or a starting point, which determines the ordering of magnitude for stimuli other than numbers. Therefore it seemed that the instruction “smaller” in my experiment dictated the right side as the “starting point” and animals were ordered with the smallest on the right

progressively getting larger to the left. Why SNARC occurred in the opposite direction compared to their experiment is not clear.

I hoped that after training sessions, more SNARC effects would be observed after testing in Session 2. Unfortunately this did not occur and the SNARC effect that I found in Session 1 was no longer present in Session 2. This may have occurred for several reasons. It is possible that there was not a real SNARC effect present in Session 1 and my result was merely a Type 1 error, which was not repeated in Session 2.

A second possibility is that the effect was real in Session 1 but because of Type 2 errors, it was not observed in Session 2, nor was found in any other condition. In order to resolve this confusion, the experiment would have to be replicated with more practice trials and/or use more effective practice, to generate a more robust SNARC effect.

I did observe significant semantic congruity effects for each session, which were very similar. They show that participants' responses to small stimuli were facilitated by the instruction "smaller" and larger stimuli were facilitated by the instruction "larger". These effects indicate that participants did in fact process animal size in a linear order, however, they do not reveal anything about the direction in which participants may have mentally ordered the stimuli.

In this experiment I failed to find SNARC in all but one condition in Session 1. Based on the inconsistency of my significant finding with any previous SNARC result, as well as my failure to find SNARC in any other condition of my experiment, in either Session 1 or 2, the result seems suspect and is quite possibly a Type 1 error.

EXPERIMENT 2

SNARC has been found for numbers and also forms of magnitude other than numbers such as animal size. This experiment explored the possibility that SNARC may be attained for a more abstract quantity such as animal intelligence.

Method

Participants

Sixty-one Carleton University students participated in a single session lasting approximately 45 minutes, in exchange for course credit.

Stimuli and Design

Stimulus pairs were constructed from the names of seven animals representing varying levels of intelligence. Holyoak and Mah (1981) attained subjective ratings of animal intelligence by asking participants to rate animals between 0 and 10, 10 being the highest level of intelligence. Based partially on this list and partially on what I predicted participants would think, I constructed six pairs, which were be used as stimuli in the experiment (human-dolphin; dolphin-dog, dog-cat, horse-cow, cow-trout, trout-worm).

I noticed a confound with these pairs prior to the experiment, being that intelligence, to a certain degree, increased with the size of the animal. Of course I rated the human and dolphin as the most intelligent, which are not as big as a horse or a cow, but are bigger than all other pairs, and also, the trout and worm, which were considered to be the least intelligent, were also smaller than all other pairs. Also, sometimes in the pairs, the more intelligent animal of the two was also the larger. To discourage any tendency to ignore intelligence and just pick the larger or smaller animal, and to prevent a mapping of animals based on size, I added two additional pairs (cat-cow, dog-horse), made from

animals already being used in the experiment, which appeared as many times as all other pairs but were not included in the analysis.

The order that I used to construct the pairs was verified by the ratings of intelligence participants provided after they had completed testing. Participants were given a list of the animals used in the study in alphabetical order and asked to rate each animal for their intelligence on a scale of 1-10. Table 1 shows the mean scores of rated intelligence for each animal and each pair.

The pairs appeared on the left and right of centre on the monitor in black, Arial font (36) in the approximate upper third of a white background on a 17-inch (43 cm) View Sonic monitor. The experiment was programmed and executed using SuperLab 2 software on a PC with Pentium III microprocessor. The keypad used was the Cedrus model RB-620 (keys 2 and 5).

Each pair was presented in each of the two possible left-right position orders, an equal amount of times, but in a random order. The two instructions, “smarter” and “dumber” appeared equally often underneath the stimuli. The six stimulus pairs by two orders by two instructions were presented 13 times. The first two blocks were counted as practice.

Procedure

Participants were instructed that the presentation of the word “smarter” or “dumber” would serve as a warning for the next trial and would indicate whether or not they are required to choose the more or less intelligent of the two animals. After the instruction was presented, the pair of names or pictures followed 500 ms later. Participants were instructed to indicate their decision by pressing the left response key to choose the animal on the left or the right response key to choose the animal on the right. The instruction and

names or pictures remained on the screen until participants indicated their decision. The next trial followed 500 ms later. Participants were encouraged to respond accurately but also quickly.

Results

Prior to analysis, data were eliminated from one participant because the person admitted they had performed incorrectly in the experiment in a systematic manner. Therefore, the data used in the analysis were from 60 participants. There was no elimination of errors from the data in my analysis because each participant ordered the intelligence of animals differently which were recorded after the study, therefore errors were different for each participant. It would be possible to analyze the data based on the specific intelligence ratings given by each participant, however it was evident that some participants reported an ordering of intelligence that was not reflected in their responses. Additionally, stimuli were presented as pairs, constructed from adjacent levels of intelligence, which did not generally match the adjacent levels of intelligence reported by each participant.

Ratings

I designed the pairs of stimuli in this experiment based on the ratings of intelligence I expected participants to give. The order of intelligence as rated by participants, shown in Table 1, was consistent with the order that I implemented in the design.

Table 1

Average participant ratings, on a scale of 1-10, for the intelligence of each animal included in the experiment.

Animal 1	Rating 1	Animal 2	Rating 2	Mean Rating	Mean Difference
human	9.77	dolphin	7.80	8.79	1.97
dolphin	7.80	dog	6.62	7.21	1.18
dog	6.62	cat	6.12	6.37	0.50
horse	6.05	dog	6.62	6.34	0.57
cow	4.12	cat	6.12	5.12	2.90
horse	6.05	cow	4.12	5.09	1.93
cow	4.12	trout	3.22	3.67	0.90
trout	3.22	worm	1.67	2.45	1.55

Response time

After combining the data over the two presentation orders, medians for each cell, from each participant were averaged. Cells that exceeded the mean plus three standard deviations were replaced with the mean plus three standard deviations for that condition. A total of 42 means were replaced. An ANOVA was then conducted with mean RTs as the dependent variable and the six pairs of varying intelligence, the two stimulus types (words, pictures), the two instructions (dumber, smarter), and the direction of response as the within-participant factors. Significance levels for analysis of variance (ANOVA) were based on the Huyn-Feldt adjusted degrees of freedom, although the degrees of

freedom and error mean squares provided are those specified by the design. The level of significance used for all tests was $\alpha = 0.05$.

The ANOVA revealed, as seen in Figure 7, participants responded significantly faster to pictures ($M = 907.053$ ms) than to words, ($M = 1259.219$ ms), $F(1, 59) = 166.694$, $MSE = 535680.588$. The ANOVA also revealed a significant effect of instruction, $F(1, 59) = 37.850$, $MSE = 457852.161$. As seen in Figure 7, participants responded faster when the instruction was “smarter” ($M = 1005.565$ ms) compared to “dumber” ($M = 1160.707$ ms).

Semantic Congruity Effects

Figure 7. Mean response times for each animal pair, for each instruction arranged from the least intelligent pair (worm-trout), progressing to the most intelligent (dolphin-human).

Semantic congruity effects were partially hidden by the substantially longer response times for “dumber” trials, however the interaction between instruction and magnitude, was in fact significant, $F(5, 295) = 14.661$, $MSE = 1218121.647$. Figure 7 shows that for the least intelligent pair, response times were close to the same for both instructions. However, as the pairs became more intelligent, responses to the instruction “smarter” progressively became faster compared to when the instruction was “dumber”. There was a significant effect of magnitude level, $F(5, 295) = 31.314$, $MSE = 122580.874$. Generally, the least and most intelligent pairs were responded to fastest and the rest seemed to vary in an unpredictable manner. This is somewhat different from comparative judgements of magnitude, which typically show a magnitude effect, or a general increase in response time as magnitude increases. There was also a significant interaction between stimulus type (words, pictures) and magnitude, $F(5, 295) = 16.777$, $MSE = 67609.819$. Response times for both pictures and words seem to follow a similar pattern, however, the interaction seems to be caused by slight discrepancies in response times from one pair to another, with no particular pattern, as can be observed in Figure 7.

The new four-way interaction between instruction (dumber, smarter), stimulus (words, pictures), magnitude (intelligence level) and response button (left, right) $F(5, 295) = 3.047$, $MSE = 36774.374$ is important because it indicates that there was a SNARC effect and it was different between words and pictures, which is consistent with my regression analyses. For words, when the instruction was “dumber” right-hand responses ($M=1385.57$ ms) were faster than left ($M=1385.57$ ms) for less intelligent pairs but left-hand responses ($M=1303.18$ ms) were faster than right ($M=1345.00$ ms) for higher intelligence pairs. However when the instruction was “smarter” the opposite occurred.

Right-hand responses ($M=1050.16$ ms) were faster than left ($M=1062.80$ ms) for higher intelligence pairs and left-hand responses ($M=1265.47$ ms) were faster than right ($M=1316.48$ ms) for lower intelligence pairs. There were no other significant effects.

Consistent with my regression analysis, SNARC for pictures was in the opposite direction compared to words. When the instruction was “dumber”, right-hand responses ($M=1015.82$ ms) were faster than left ($M=1031.47$ ms) for higher intelligence pairs. For lower intelligence pairs, right-hand responses ($M=924.84$ ms) were still faster than left ($M=928.632$ ms), but the difference in response time between left and right responses decreased substantially. However, when the instruction was “smarter”, left-hand responses ($M=805.83$ ms) were faster than right ($M=823.673$ ms) for higher intelligence pairs and right-hand responses ($M=850.87$ ms) were faster than left ($M=875.25$ ms) for lower intelligence pairs. There were no other significant effects.

Individual participant analyses

To investigate the possibility of a SNARC effect further, I conducted repeated measures regression analyses as outlined in Lorch and Myers (1990), separately with each instruction (dumber, smarter), and each stimulus type (words, pictures). After combining the data over order, I first obtained standardized beta regression coefficients for each participant with the SNARC index ($RT(\text{Right})-RT(\text{Left})$) as the dependent variable and magnitude of stimulus pair as the predictor. I then tested the hypothesis (two tailed) that the mean of these weights differed from zero using the t-distribution. R^2 values were calculated using group means rather than individual means.

No statistically significant SNARC effects were found, however, the results approached significance in each condition and high R-squared values were obtained for words. The SNARC effect for the words condition resembled that of Experiment 1. For each instruction, the pair of animals that represented the highest level corresponding to that instruction (dumbest or smartest) was responded to fastest with the right hand. As seen in Figure 8, when the instruction was “dumber” rightward responses were faster compared to left, for the least intelligent pairs, but leftward responses were faster than right for the most intelligent pairs, $M = .116$, $SD = .462$, $t(59) = 1.935$, $p = .058$, $R^2 = .986$. When the instruction was “smarter” the same rule is followed resulting in a SNARC effect in the opposite direction. Rightward responses were faster compared to left, for the most intelligent pairs, but leftward responses were faster than right for the least intelligent pairs, $M = -.113$, $SD = .451$, $t(59) = -1.947$, $p = .056$, $R^2 = .717$.

Figure 8. Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of words representing animal intelligence.

When participants responded to pictures instead of words, the pattern of SNARC reversed, which was not consistent with the words condition. As seen in Figure 9, when the instruction was “dumber”, participants responded faster to the left, not right, for the most unintelligent pairs but faster to the right compared to left for the relatively intelligent pairs, $M = -.069$, $SD = .394$, $t(59) = -1.560$, $p = .124$, $R^2 = .009$. For the instruction “smarter”, again, it was leftward instead of rightward responses that were

faster for the more intelligent pairs but rightward responses were faster than leftward for the less intelligent pairs, $M = .110$, $SD = .445$, $t(59) = 1.923$, $p = .059$, $R^2 = .119$.

Figure 9. Mean left-hand response time subtracted from mean right-hand response times as a function of each stimulus pair of pictures representing animal intelligence.

Discussion

My goal in this experiment was to test if a SNARC effect was attainable on an abstract dimension such as animal intelligence. This experiment was only a partial success. Although a SNARC effect seems to exist for judgements of animal intelligence, at least when represented by words, I can only make hypotheses for future studies because none of the conditions were significant. However, because the p-values in all conditions approached significance, for the sake of discussion, I will treat the observed SNARC effects as if they were real.

The SNARC effects I found for words were instruction-dependent in a way that was consistent with the SNARC effect I found for animal size in Experiment 1. When the instruction was “dumber”, participants responded faster to the right compared to left, for less intelligent animals, but faster to the left, compared to right for more intelligent animals. However, when the instruction was “smarter” the converse was true; participants responded faster to the right compared to left for more intelligent animals, but faster to the left compared to right for less intelligent animals.

It was as if participants mentally ordered stimuli from right to left starting with “the most” on the right. When the instruction was “smarter” it seems as though participants ordered animals from most intelligent to least intelligent, from right to left, but when the instruction was “dumber” participants seemed to order pairs in the reverse order, from the least intelligent to most intelligent, right to left.

I had originally considered the result from Experiment 1 to be a probable Type I error, but in light of these results, it seems the effect was probably real. It seems as though in that experiment, as well as the present experiment, for words, participants were ordering

the stimuli such that the most extreme pair corresponding to the instruction (e.g. the “dumbest” or “smartest” or “smallest”) would always be placed farthest to the right and the rest of the stimuli would fall in place to the left, progressing towards the opposite pole.

The results from the “words” condition are consistent with the results of Shaki and Petrusic (2005b) in the sense that, they too found an instruction-dependent SNARC effect, however, they found that leftward responses were always favoured by the highest ranking stimuli that corresponded to the instruction. Therefore, leftward responses were faster than right for the smallest stimuli when the instruction was “smaller” and for larger stimuli when the instruction was “larger”. In my experiment, for the words condition rightward responses were favoured, not left.

However, the apparent effect I found for pictures was in fact consistent with the results of Shaki and Petrusic. I found that leftward responses were favoured over right with the highest-ranking stimuli that were consistent with the instruction. It is not clear why participants would order animal intelligence in one direction or the other and why the direction would reverse for pictures versus words.

The semantic congruity effects I found are consistent with those found in Experiment 1. Responses to the more intelligent animals were facilitated by the instruction “smarter” and responses to the less intelligent animals were facilitated by the instruction “dumber” indicating that participants’ sequentially ordered animal intelligence.

My results are important because they demonstrate that SNARC not only occurs for numbers and magnitude but is also possible for abstract concepts such as intelligence. It

is unclear at this point why participants seemed to order animal intelligence the way they did. The subsequent experiments using temperature will enlighten my discussion.

EXPERIMENT 3

Shaki and Petrusic (2005) found a SNARC effect for negative digits when participants were required to choose the “smaller” or “larger” of a pair. They found a SNARC effect that corresponds to spatial arrangement of the mental number line when negative and positive digits were mixed together in a block, but a SNARC effect consistent with the absolute magnitude of the digit, when negative and positive digits were presented in separate blocks. It is possible that unlike the results of Shaki and Petrusic (2005), the results of the present experiment will show that, even when negative numbers are presented exclusively within a block they may still elicit a SNARC effect consistent with the spatial arrangement of the mental number line. The reason for this is that a context of temperature may elicit an orientation of negative numbers that would not seem to be important to participants otherwise.

Method

Participants

Forty-one Carleton University students participated in a single session lasting approximately 45 minutes, in exchange for course credit.

Stimuli and Design

The digits 1-9, excluding 5, were used to construct stimulus pairs that were categorized according to their numerical magnitude (not absolute magnitude) as relatively small positive (1, 2), (2, 3), (3, 4); large positive (6, 7) (7, 8), (8, 9); small negative (-9, -

8) (-8, -7), (-7, -6); and large negative (-4, -3), (-3, -2), (-2, -1). Each pair was presented in each of the two possible left-right position orders, an equal amount of times, but in a random order. Pairs were presented in a condition in which all trials were either positive or negative (blocked) as well as a condition in which trials alternated randomly between negative and positive trials (intermixed). The two instructions, “warmer” and “colder” were blocked and occurred equally often. The 12 stimulus pairs by two orders by two instructions by two conditions were presented eight times. The order of condition (blocked mixed), polarity for the blocked condition and blocked instructions were counterbalanced so that eight possible orderings were possible (e.g., Order1: blocked-warmer-positive, blocked-warmer-negative, blocked-colder-positive, blocked-colder-negative, mixed-warmer, mixed-colder; Order 2: blocked-warmer-negative, blocked-warmer-positive, blocked-colder-negative, blocked-colder-positive, mixed-warmer, mixed-colder).

One temperature of the pair appeared to the left of centre and the other appeared to the right of centre in the approximate upper third of the screen. The instruction appeared below the stimuli in the approximate centre of the screen. The instructions and stimuli were printed in black (Arial font, size 36) on a white background on a 17-inch (43 cm) View Sonic monitor. The experiment was programmed and executed using SuperLab 2 software on a PC with Pentium III microprocessor. The keypad used was the Cedrus model RB-620 (keys 2 and 5).

Procedure

Participants were tested individually seated at a desk, approximately 90 cm from the computer monitor. Participants were told that the presentation of the word “colder” or

“warmer” would serve as a warning for the next trial and indicated whether or not they are required to choose the warmer or colder of the two temperatures. After the instruction was presented, the pair of temperatures followed 500 ms later. Participants indicated their decision by pressing the left key to choose the temperature on the left or the right key to choose the temperature on the right. The instruction and temperatures remained on the screen until participants indicated their decision. The next trial followed 500 ms later. Participants were encouraged to respond accurately but also quickly.

Results

Prior to analysis, data were eliminated from one participant because of a corrupted data file. Therefore, the data used in the analysis were from 40 participants. Separate analyses will be presented for response time data, followed by an analysis for error rates. The analyses for response time consisted of an ANOVA presented first, followed by a one sample t-test based on individual beta weights representing the SNARC effect.

Response time

Medians for each cell, from each participant were averaged. Cells that exceeded the mean plus three standard deviations were replaced with the mean plus three standard deviations for that condition. A total of 76 medians were replaced. An ANOVA was then conducted with mean correct RTs as the dependent variable and the six levels of magnitude, the two polarities of the number pair, two condition (blocked, intermixed) the two instructions (colder, warmer), and the direction of response (left, right) as the within participant factors. Significance levels for analysis of variance (ANOVA) were based on the Huyn-Feldt adjusted degrees of freedom, although the degrees of freedom and error

mean squares provided are those specified by the design. The level of significance used for all tests was $\alpha = 0.05$.

The ANOVA showed that participants responded faster when the instruction was “warmer” ($M = 742.05$ ms), compared to “colder” ($M = 778.30$ ms), $F(1, 39) = 5.624$, $MSE = 224361.933$. Participants also responded faster to positive pairs ($M = 720.89$ ms), compared to negative pairs, ($M = 799.46$ ms), $F(1, 39) = 38.880$, $MSE = 152432.963$. Additionally, participants were faster in the blocked condition ($M = 601.83$ ms) compared to the intermixed condition ($M = 929.67$ ms), $F(1, 39) = 108.039$, $MSE = 891143.050$. There was also a significant effect of magnitude of pair, $F(5, 195) = 3.146$, $MSE = 24508.452$. Response times for pairs seemed to differ in an idiosyncratic manner. There was also a significant effect for response button $F(1, 39) = 10.565$, $MSE = 22122.470$. Right-hand responses ($M = 752.37$ ms), were faster than left-hand responses ($M = 767.97$ ms).

Temperature of pair interacted with polarity $F(5, 195) = 16.020$, $MSE = 31589.033$. The interaction seems to have been caused by the fact that generally, responses were slower as temperature increased for positive pairs. However, in contrast, for negative pairs, responses generally became faster as temperature increased.

Instruction also interacted with polarity, $F(1, 39) = 48.579$, $MSE = 68310.049$. This result signifies a significant semantic congruity effect. As seen in Figure 10, semantic congruity effects were observed as responses to negative numbers were faster when the instruction was “colder” opposed to “warmer”, but for positive numbers, responses were faster when the instruction was “warmer” compared to “colder”. This pattern of semantic congruity effect is very similar to the one found by Shaki and Petrusic (2005).

There was a significant three-way interaction between instruction, polarity and condition (blocked, intermixed), $F(1, 39) = 5.243$, $MSE = 50557.807$. As it appears in Figure 10 there is a greater discrepancy in response times for positive pairs between the instructions “warmer” and “colder” in the intermixed condition compared to the blocked condition.

Figure 10. Mean response times collapsed across thermometer presentation, for the instructions “colder” and “warmer” in blocked and intermixed conditions, as a function of each stimulus pair. The x-axis represents the order of temperature with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).

The ANOVA also showed that instruction interacted with condition (blocked, intermixed), $F(1, 39) = 7.781$, $MSE = 156159.499$. Generally, it took participants the same amount of time to respond in the blocked condition, for the instruction “colder” ($M = 602.17$ ms) and “warmer” ($M = 601.49$ ms). However, when responding in the intermixed condition, participants were substantially slower when the instruction was “colder” ($M = 954.431$ ms) compared to “warmer” ($M = 882.60$ ms).

Polarity also interacted with condition (blocked, intermixed), $F(1, 39) = 19.578$, $MSE = 116259.701$. As seen in Figure 10, even though responses were always faster in the blocked condition, the difference in response time between blocked and intermixed conditions was greater for negative pairs ($M = 365.38$ ms) compared to positive pairs ($M = 267.99$ ms). Instruction interacted with magnitude of pair, $F(5, 195) = 5.311$, $MSE = 17080.339$. Generally response times became slower as magnitude increased when the instruction was “colder” but stayed relatively even when the instruction was “warmer”.

Instruction interacted with response button, $F(1, 39) = 8.028$, $MSE = 16086.070$. Left ($M = 780.30$ ms) and right ($M = 776.30$ ms) response times were comparable when the instruction was “colder”, but when the instruction was “warmer”, left responses ($M = 755.65$ ms) were considerably slower compared to right responses ($M = 728.44$ ms).

There was a three-way interaction between instruction, polarity, and condition, $F(1, 39) = 5.243$, $MSE = 50557.807$. As can be seen in Figure 10, the semantic congruity effect is larger for positive pairs in the intermixed condition compared to the blocked condition.

There was also a significant interaction between instruction, condition, and response button, $F(1, 39) = 6.804$, $MSE = 16974.248$. In the blocked condition, the speed

advantage of right over left responses was maintained between the instructions “colder” ($MD = 12.06$ ms) and “warmer” ($MD = 13.69$ ms). However, in the intermixed condition, left and right responses were more similar for the instruction “colder” ($MD = 4.05$ ms) and right responses were faster than left, once again for the instruction “warmer” ($MD = 41.09$ ms).

Finally, there was a significant interaction between instruction, magnitude and response button $F(5, 195) = 2.810$, $MSE = 27064.950$, thus showing a SNARC effect although it only occurred when data was collapsed across condition and polarity. When the instruction was “colder”, right responses were faster than left responses when the pairs were numerically smaller but left responses were faster than right responses when the pairs were numerically larger. However, when the instruction was “warmer”, right responses were generally faster than left responses when the pairs were numerically larger but the discrepancy between left and right responses was minimal when the pairs were numerically smaller.

Individual participant analyses

To investigate the possibility of a SNARC effect further, I conducted repeated measures regression analyses as outlined in Lorch and Myers (1990), for each participant separately, with each instruction (colder, warmer), each polarity (positive, negative) and condition (blocked, intermixed). I first obtained standardized beta regression coefficients for each participant, representing the SNARC index (rightward response time - leftward response time) as the dependent variable, and temperature of stimulus pair as the predictor. I then tested the hypothesis (two tailed) that the mean of these weights differed

from zero using a one-sample t-test. R^2 values were calculated using group means rather than individual means.

The results showed that a SNARC effect occurred, which seemed to be dictated by the absolute magnitude of both positive and negative numbers in the significant conditions. Significant SNARC effects occurred for positive numbers when the instruction was “warmer” and for negative numbers when the instruction was “colder”. Unlike the results of Shaki and Petrusic (2005a) I found no differences in the direction of SNARC between blocked and intermixed conditions.

Figure 11. Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “colder”, for the blocked and intermixed conditions. The left panel represents negative and the right panel represents positive pairs.

As seen in the left panel of Figure 11, when the instruction was “colder” a significant SNARC effect occurred for negative numbers in the blocked condition, which

seemed to be dictated by the absolute magnitude of numbers. The mean standardized beta weight ($M = .24$, $SD = 0.059$), differed reliably from zero, $t(39) = 4.000$, $R^2 = .349$.

When the temperatures were the lowest, (-8, -9) participants were faster to respond to the right compared to the left. However, as the pairs increased in temperature (-2, -1), participants were faster to respond to the left compared to the right. Also seen in the left panel of Figure 11, the same pattern of significant results occurred in the intermixed condition, ($M = .160$, $SD = 0.492$), $t(39) = 2.053$, $R^2 = .501$. The right panel of Figure 11 shows nonsignificant SNARC effects for positive numbers, with the instruction “colder” in the blocked condition, ($M = -.056$, $SD = 0.482$), $t(39) = -.741$, $p = .46$, $R^2 = .068$ and the intermixed condition, ($M = -.056$, $SD = 0.482$), $t(39) = -.741$, $p = .43$, $R^2 = .278$.

Figure 12. Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “colder”, for the blocked and intermixed conditions.

To explore the possibility that SNARC was represented across the entire range of temperatures, we conducted the same regression analysis including both negative and positive temperatures. As can be seen in Figure 12, for the instruction “colder”, there was a significant SNARC effect for the intermixed condition across the entire range of stimuli, ($M = .116$, $SD = 0.289$), $t(39) = 2.534$, $R^2 = .299$. Responses were faster to the right compared to left for the numerically smallest numbers, but responses to the left

became faster compared to the right as numerical magnitude increased. Also shown is the nonsignificant slope for the blocked condition ($M = .024$, $SD = 0.333$), $t(39) = .653$, $R^2 = .006$.

Figure 13. Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “warmer”, for the blocked and intermixed conditions. The left panel represents negative and the right panel represents positive pairs.

As seen in the right panel of Figure 13, a SNARC effect also occurred when the instruction was “warmer” but only for positive pairs. There was a significant SNARC effect in both the blocked condition ($M = -.243$, $SD = 0.489$), $t(39) = -.3.140$, $R^2 = .565$, as well as the intermixed condition, ($M = -.170$, $SD = 0.365$), $t(39) = -2.950$, $R^2 = .369$. Participants were faster to respond right compared to left, for higher temperature pairs, but faster to respond left compared to right for lower temperature pairs. Again it seems as though the direction of SNARC was dictated by the absolute value of numbers, although

as seen in the left panel of Figure 13, this cannot be verified for negative numbers because SNARC was not significant in either the blocked ($M = .038$, $SD = 0.431$), $t(39) = .551$, $p = .59$, $R^2 = .016$, or intermixed condition, ($M = -.025$, $SD = 0.482$), $t(39) = -.329$, $p = .74$, $R^2 = .101$.

Figure 14. Mean left-hand response time subtracted from mean right hand response times as a function of each stimulus pair for the instruction “warmer”, for the blocked and intermixed conditions.

When a regression analysis was conducted across the entire range of stimuli, there was a significant SNARC effect for the blocked condition, as seen in Figure 14, ($M = -.111$,

$SD = 0.306$), $t(39) = -2.293$, $R^2 = .026$. Left hand response were faster compared to right for the numerically smallest pairs but right hand responses became faster than left as numerical magnitude increased. Also shown is the nonsignificant slope for the intermixed condition, ($M = -.035$, $SD = 0.356$), $t(39) = -.619$, $R^2 = .119$.

Error Rates

The overall error rate was low at 2.6%. The correlation between response time and mean error rate was $r = .375$, $p < .001$. Therefore, lower accuracy was associated with longer response times, so there was no speed-accuracy trade-off.

An ANOVA was conducted with mean error rate as the dependent variable and instruction, polarity, condition, magnitude, and button as the independent variables. There was a main effect of polarity $F(1, 39) = 9.932$, $MSE = .004$. There were actually fewer errors for negative comparisons (2.02%) compared to positive (2.65%). There was a main effect of condition (blocked, intermixed), $F(1, 39) = 9.932$, $MSE = .009$. As expected there were fewer errors in the blocked condition (1.74%) compared to the intermixed condition (2.93%).

The interaction between polarity and magnitude was significant, $F(5, 195) = 7.413$, $MSE = .004$. For negative pairs errors decreased as magnitude increased towards zero, but for positive pairs, errors increased as magnitude increased. The interaction between condition and magnitude was significant, $F(5, 195) = 4.194$, $MSE = .003$. Error rates seemed to vary in an idiosyncratic manner across magnitude for both blocked and intermixed conditions.

There was a three-way interaction for instruction, polarity and condition, $F(1, 39) = 8.075$, $MSE = .004$. Although the intermixed condition always had more errors than the

blocked condition, the error rate between the blocked and intermixed condition was more comparable for negative pairs, when the instruction was “warmer” compared to other pairs. There was also a four-way interaction between instruction, polarity, magnitude and button, $F(5, 195) = 2.874$, $MSE = .003$, but it was not easily interpretable.

Discussion

The results of this experiment show an instruction-dependent SNARC effect, which suggests participants processed temperatures on a full mental number line that includes negative numbers. Furthermore, unlike the results of Shaki and Petrusic (2005a), we found no difference in the direction of SNARC between blocked and intermixed conditions.

This is the first experiment to my knowledge that shows an instruction-dependent SNARC effect using negative numbers. In fact, the only other study I know of where the direction of SNARC changes based on instruction, is that of Bächtold et al. (1998). In both blocked and intermixed conditions, when the instruction was “warmer”, left hand responses were faster than right, for the numerically smallest pairs and right hand responses were faster than left for numerically larger pairs. When the instruction was “colder” however, the direction of SNARC was in the opposite direction such that, for the numerically smaller pairs, right hand responses were faster than left but for the numerically larger pairs, left hand responses were faster than right hand responses.

These results support the hypothesis that negative numbers are included on the mental number line. This is consistent with the results of Fischer (2003b) who also found evidence of a full mental number line. He found that responses to negative pairs, ordered in a way that was consistent with a full mental number line (-9 -4), were faster than pairs

that were presented in an order inconsistent with a full mental number line (-4 -9).

However, as Shaki and Petrusic previously argued, these results may be nothing more than an extended semantic congruity effect.

However, Fischer and Rottmann (Experiment 1, 2005) found that SNARC was reflective of absolute magnitude during paired comparisons for negative numbers in both blocked and intermixed conditions, thus supporting a mental number line that does not include negative numbers.

My results however are partially consistent with the results of Shaki and Petrusic (2005a). I found, as they did, that for negative pairs in the intermixed condition (for the instruction “warmer”), the direction of SNARC was consistent with the hypothesis that negative numbers are processed according to their numerical magnitude, thus supporting the hypothesis that the mental number line does indeed extend past zero to include negative numbers. However, I found that the direction of SNARC for negative numbers in the intermixed condition was the same as the blocked condition, suggesting that mixing positive and negative pairs together does not change how negative numbers are processed. However, Shaki and Petrusic found a SNARC effect for the blocked condition in the opposite direction compared to the intermixed condition, suggesting that participants in their experiment were processing polarity and magnitude separately during the blocked condition, which is not consistent with a full mental number line hypothesis.

They explained the difference in the direction of SNARC they found between blocked and intermixed conditions, by using the model of Banks, Fujii and Kaya-Stuart (1976). According to the model of Banks, Fujii and Kaya-Stuart (1976), during pair comparisons, magnitude is processed in an initial stage in which numbers are coded as relatively large

(L, L+) or relatively small (S, S+), “as quickly as possible”. In a second stage of processing, a decision is made. If the instruction is to choose the “Larger” then a quick response will be made for the compatible codes (L, L+) whereas a slower response will be made for (S, S+) because the items must be recoded to (L+, L) before a decision can be made. In contrast when the instruction is to choose the smaller, a fast response will be made for pairs coded as (S, S+) but a slower decision will be made for pairs coded as (L, L+).

This model accounts for results found by Shaki and Petrusic (2005a) in the blocked condition because when polarity is known prior to the trial, magnitude can be processed initially, and the polar code can be coded in the second stage, resulting in a SNARC effect representative of absolute magnitude but with a correct answer given by participants based on the relative magnitude of digits, processed in the decision stage.

Shaki and Petrusic further suggested that during the intermixed condition, polarity is not known prior to the trial so that it must be encoded with magnitude in the first stage of processing and therefore polarity is reflected in the SNARC effect. Using this same logic, applied to our results, I would have expected the same difference in SNARC for negative numbers between the blocked and intermixed condition. However, there was none.

In addition to the SNARC results, the semantic congruity effects that I found, which were similar to those found by Shaki and Petrusic (2005a), also provide some evidence that participants do in fact order negative numbers on a full number line. If participants processed positive and negative numbers according to their absolute magnitude then positive and negative numbers should each generate the same semantic congruity effect;

numbers with a larger absolute value should be faster with the instruction “warmer” and smaller numbers, faster with the instruction “colder”. However, instead, I found a semantic congruity effect that spanned across negative and positive numbers, replicating the semantic congruity effect found by Shaki and Petrusic (2005a); responses for negative numbers were faster than responses for positive numbers with the instruction “colder” but responses for positive numbers were faster than negative numbers with the instruction “warmer”. This is consistent with the hypothesis that participants process negative numbers according to their numerical value. However this effect was statistically categorized as negative and positive, rather than graded across the entire span of digits, as would be expected as a result of processing a full mental number line, and therefore it weakens the argument for a full mental number, somewhat.

It is not clear why Shaki and Petrusic found differences in SNARC between the blocked and intermixed conditions, but we didn’t. However there may be a reasonable explanation as to why our results were instruction-dependent their results were not. It is possible that in my experiment, the “normal” default left to right, overlearned ordering of numbers, may have been circumvented due to context. Perhaps once numbers have a semantic attachment, their ordering becomes more flexible to allow for efficient processing. My results suggest that temperatures are processed differently from numbers in general. For some reason, the apparent robustness of a left-to-right ordering of numbers usually seen in SNARC experiments, does not apply to temperatures. This suggests that temperatures may not be interpreted as numbers in a pure sense, but rather, as symbols that are mapped on to the dimension of temperature or separate dimensions of warmth and coldness.

An instruction-dependent SNARC effect has in fact been observed for forms of magnitude other than numbers, such as animal size and line-length. However, the instruction-dependent SNARC effect Shaki and Petrusic (2005b) found for nonnumerical magnitude is arguably in the opposite direction of what I found. They found the most extreme comparison corresponding to the instruction favoured leftward responses, compared to rightward. When the instruction was “smaller” the direction of SNARC suggested an ordering of stimuli with the smallest on the left and the largest on the right. However, when the instruction was “larger”, the direction of SNARC reversed such that it was reflective of an ordering from the largest stimuli on the left to the smallest on the right. It is unclear at this point, why participants in my study seemed to order temperatures in the particular directions that were reflected by my results. The next experiment will explore if SNARC exists for positive and negative temperatures in the vertical dimension in both blocked and intermixed conditions.

EXPERIMENT 4

Some studies have shown that the SNARC effect can occur vertically as well as horizontally (Ito & Hatta, 2004). However, it is unclear if it can occur vertically for negative digits. This experiment explores this possibility in the context of temperature. This experiment tests if SNARC occurs for negative numbers in the vertical dimension, and if the results from Experiment 3 which suggest that negative numbers are not represented on the mental number line, will be replicated.

In the second half of the experiment, a picture of a thermometer was added and remained on the screen. I reasoned that by emphasizing the context of temperature, the

strength of SNARC might be enhanced or altered in some conditions. The thermometer always appeared in the second half of the experiment to eliminate the possibility of any carry-over effects caused by the thermometer.

Method

Participants

Thirty-three Carleton University students participated in a single session lasting approximately 45 minutes, in exchange for course credit.

Stimuli and Design

The stimuli and design were the same as Experiment 3 except for several differences:

1) The temperature pairs were presented in a vertical plane. One number was presented up from the centre point and one was presented down from the centre point in “Arial Black” font, size 24 (see Appendix A). There was no obvious centre point indicated on the thermometer and no indication of whether the scale was Fahrenheit or Celsius.

2) I eliminated the written instructions and instead participants were expected to determine the instruction based on the colour of the stimuli. Blue pairs indicated that participants were required to choose the colder temperature and red pairs indicated that participants were required to choose the warmer temperature of the pair.

3) For the second half of the experiment a picture of a thermometer, created in PowerPoint, appeared to the left of the digit pair and remained on the screen for each trial.

4) The twelve stimulus pairs by two orders by two instructions by two conditions (blocked, intermixed), by two thermometer conditions (present or not) were presented six times.

Procedure

The procedure for Experiment 4 was the same as in Experiment 3, except that instead of responding on a keypad, participants used a standard computer keyboard and were instructed to press the upper "Y" key to choose the upper digit on the screen or to press the lower "B" key to choose the lower digit. The keys were covered with silver tape.

The monitor used was a 19-inch (48.3 cm) NEC monitor. The experiment was programmed and executed using SuperLab 4 software on a PC with Pentium IV microprocessor.

For the second half of Experiment 4, the same participants repeated the first half, except that on all trials, a picture of a thermometer appeared on the screen, vertically beside the temperatures. The temperatures remained in the same position as they did in Experiment 4. I reasoned that as in the study of Bächtold, Baumüller, and Brugger (1998), where the direction of SNARC depended on whether participants imagined a clock face or a ruler, adding a contextual feature such as a thermometer may strengthen the influence of mental number line such that the SNARC effect is intensified or somehow altered.

Results

Prior to analysis, data were eliminated from three participants because their response accuracy was below 90%. Therefore, the data used in the analysis were from 33 participants.

Response time

Medians for each cell, from each participant were averaged. Cells that exceeded the averaged median plus three standard deviations were replaced with the averaged median plus three standard deviations for that condition. An ANOVA was then conducted with mean correct RTs as the dependent variable and the six magnitude pairs, two polarities of the number pair (negative, positive), the two instructions (colder, warmer), two levels of thermometer picture (present or not) and the direction of response as the within participant factors. Significance levels for analysis of variance (ANOVA) were based on the Huyn-Feldt adjusted degrees of freedom, although the degrees of freedom and error mean squares provided are those specified by the design. The level of significance used for all tests was $\alpha = 0.05$.

The ANOVA, revealed various main effects and interactions. The ANOVA showed that participants responded faster when the instruction was “warmer” ($M = 949.124$ ms), compared to “colder” ($M = 978.899$ ms), $F(1, 32) = 8.822$, $MSE = 159174.064$. Participants also responded faster to positive pairs ($M = 922.521$ ms), compared to negative pairs, ($M = 1005.502$ ms), $F(1, 32) = 108.009$, $MSE = 100983.243$. Additionally, participants were faster in the blocked condition ($M = 811.424$ ms) compared to the mixed condition ($M = 1116.599$ ms), $F(1, 32) = 274.484$, $MSE = 537445.928$. Interestingly, participants were significantly faster to respond when there

was a picture of a thermometer on the screen ($M = 919.491$ ms) compared to when there wasn't, ($M = 1008.532$ ms), $F(1, 32) = 53.284$, $MSE = 235684.582$.

A semantic congruity effect was evident from the significant interaction of instruction and polarity, $F(1, 32) = 75.315$, $MSE = 154154.466$. As seen in Figure 15, generally, for both the blocked and mixed conditions, when the pairs were negative, responses were faster for the instruction “colder” compared to “warmer”, but when the pairs were positive, responses were faster for the instruction “warmer” compared to “colder”. The pattern of SCE replicates that of horizontal experiment, as well as the results of Shaki and Petrusic (2005).

Response times differed significantly depending on magnitude of temperature, $F(5, 160) = 13.232$, $MSE = 42752.126$. More specifically, as can be seen in Figure 15 and confirmed by a significant polarity by magnitude interaction, $F(5, 160) = 19.336$, $MSE = 57398.566$, generally, response times became faster as negative temperatures increased, but slower as positive temperatures increased. Stated differently, participants were progressively slower to respond as absolute magnitude increased. This finding indicates that participants judged temperatures according to Weber's law, in which response time increases as the ratio of magnitude of digits and numerical distance between compared digits increases.

Semantic Congruity Effects

Figure 15. Mean response times collapsed across thermometer presentation, for each instruction and condition, as a function of each stimulus pair. The x-axis represents the order of temperature with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).

Instruction interacted with condition (blocked, mixed), $F(1, 32) = 8.173$, $MSE = 155659.161$. Response times were about the same for trials with the instruction “colder” (812.14 ms) as well as “warmer” (810.71 ms), in the blocked condition; however, in the intermixed condition, response times were substantially longer when the instruction was “colder” (1145.67 ms) compared to “warmer” (1087.54 ms). Polarity (negative, positive)

interacted significantly with condition (blocked, mixed), $F(1, 32) = 13.019$, $MSE = 114404.974$. The interaction seems to have been caused by the fact that the increase in response time from blocked to intermixed trials is greater for negative numbers (335.84 ms) compared to positive numbers (274.51 ms).

Condition (blocked, intermixed) interacted significantly with thermometer, $F(1, 32) = 11.307$, $MSE = 138320.517$. Trials with a thermometer were always faster than without, however, the difference between trials with and without a thermometer was greater for the intermixed condition (120.46 ms), compared to the blocked condition (57.62 ms). Instruction and magnitude of temperature interacted significantly, $F(5, 160) = 7.783$, $MSE = 36202.591$. Response times changed in an idiosyncratic manner across temperature for each instruction. Condition (blocked, mixed) interacted significantly with response button, $F(1, 32) = 10.408$, $MSE = 26963.929$. There was no difference (0.41 ms) between overall upward and downward responses in the blocked condition, but a 27.03 ms decrease from top to bottom in the intermixed condition. Thermometer also interacted with response button $F(1, 32) = 5.891$, $MSE = 23155.008$. The speed advantage for upward over downward responses was greater without a thermometer on the screen ($MD = 23.00$ ms) compared to with a thermometer on the screen ($MD = 4.44$ ms).

There were several significant three-way interactions including instruction by polarity by thermometer, $F(1, 32) = 6.614$, $MSE = 59223.972$. This seems to have been caused by the fact that the semantic congruity effect was larger for the nonthermometer condition compared to the thermometer condition.

The interaction between polarity, condition and magnitude was significant, $F(5, 160) = 4.572$, $MSE = 32596.431$. Generally, in the blocked condition, response times increased as temperature decreased for negative numbers, but response time increased as temperatures increased for positive numbers. Response times seemed to change in a more idiosyncratic manner in the intermixed conditions.

The interaction between instruction, polarity and magnitude was significant, $F(5, 160) = 5.417$, $MSE = 29016.643$. Response times became generally faster as positive numbers became colder, when the instruction was “colder” and generally faster as negative numbers became warmer when the instruction was “warmer”. Response times seemed to change in a more idiosyncratic manner in the other conditions.

There was a significant interaction between instruction, polarity and response button (top, bottom), $F(1, 32) = 12.651$, $MSE = 82938.179$ thus indicating an instruction dependent SNARC effect, although only when data was collapsed across magnitude, mix condition and thermometer condition. When the instruction was “colder”, upward responses were faster than downward responses when the pairs were negative but downward responses were faster than upward responses when the pairs were positive. When the instruction was “warmer”, the opposite pattern occurred. When the pairs were negative, downward responses were faster than upward responses, but when the pairs were positive, upward responses were faster than downward responses.

Additionally, there was a significant three-way interaction for instruction, magnitude and response button, $F(5, 160) = 5.008$, $MSE = 33911.521$ confirming the instruction-dependent SNARC effect. When the instruction was “colder”, upward response times were faster than downward responses, for numerically smaller pairs, but for numerically

larger pairs, downward responses were faster than upward responses. However, when the instruction was “warmer” this pattern reversed. Downward responses were faster than upward responses when the pairs were numerically smaller, but upward responses gradually became faster than downward responses as the pairs increased in numerical value.

There was also a three-way interactions for polarity, condition, and thermometer $F(1, 32) = 4.592$, $MSE = 47766.264$. In the blocked condition, the speed advantage of trials with positive pairs over trials with negative pairs decreased from the nonthermometer condition ($MD = 45.66$ ms) to the thermometer condition ($MD = 58.98$ ms). However, in the intermixed condition, this speed advantage increased from the nonthermometer condition ($MD = 130.52$ ms) to the thermometer condition.

There was also an interaction for condition, thermometer, and response button $F(1, 32) = 5.107$, $MSE = 18393.076$. In the blocked condition there was almost no difference between upward and downward responses for both the nonthermometer condition ($MD = 1.98$ ms) and the thermometer condition ($MD = 1.17$ ms). However, in the intermixed condition downward responses were considerably faster than upward responses in the nonthermometer condition ($MD = 44.44$ ms) but less so in the thermometer condition ($MD = 10.05$ ms).

There was a significant four-way interaction for instruction, polarity, condition and response button, $F(1, 32) = 10.538$, $MSE = 35734.335$, polarity, condition, magnitude and response button, $F(5, 160) = 2.823$, $MSE = 33911.521$, as well as polarity, thermometer, magnitude and response button, $F(5, 160) = 2.771$, $MSE = 23592.429$.

These interactions could not be easily interpreted in a meaningful way. There were no other significant results.

Individual participant analyses

To further investigate the possibility of a SNARC effect, I conducted repeated measures regression analyses as outlined in Lorch and Myers (1990), separately with each instruction (colder, warmer), each polarity (positive, negative), condition (blocked, mixed) and thermometer condition (present or not). After combining the data over order, I first obtained standardized beta regression coefficients for each participant with the SNARC index (Upward RT – Downward RT) as the dependent variable and magnitude (temperature) of stimulus pair as the predictor. I then tested the hypothesis (two tailed) that the mean of these weights differed from zero using the t-distribution. R^2 values were calculated using group means rather than individual means.

In this experiment I observed SNARC effects that were reflective of number processing on a full number line. Also, in contrast to previous experiments with numbers I found an instruction-dependent SNARC effect, such that more extreme temperature pairs corresponding to the instruction were always responded to faster in the upward direction compared to downward, and more moderate temperatures facilitated downward responses compared to upward.

Figure 16. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

SNARC was significant when the instruction was “colder” for the negative-blocked condition, with a thermometer $M=.213$, $SD=.362$, $t(32)=3.383$, $R^2=.884$. As seen in the left panel of Figure 16, upward responses were faster compared to downward, when the numbers were most negative but as the pairs increase in temperature, downward responses became faster compared to upward. Also shown is the nonsignificant intermixed condition, $p=.89$.

As can be seen in the right panel of Figure 16, for the instruction “colder” there was a significant SNARC effect for positive numbers in the blocked condition, with a thermometer, $M=.153$, $SD=.427$, $t(32)=2.066$, $R^2=.127$. Again, the lower temperature pairs were responded to faster in an upward direction compared to downward, but the

higher temperature pairs were responded to faster in a downward direction. Also shown is the nonsignificant intermixed condition, $p = .17$.

Figure 17. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions.

When SNARC was regressed across the entire range of temperatures it was significant in the intermixed condition, $M = .131$, $SD = .356$, $t(32) = 2.102$, $R^2 = .246$ but not in the

blocked condition $M = .061$, $SD = .323$, $t(32) = 1.077$, $R^2 = .070$, as can be seen in Figure 17.

Displayed in the left panel of Figure 18 are the nonsignificant results for the instruction “colder” without a thermometer for conditions with negative numbers when blocked, $p = .91$, and intermixed, $p = .443$, as well as in the right panel, the nonsignificant results for conditions with positive numbers blocked, $p = .26$ and intermixed, $p = .24$.

Figure 18. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

Figure 19. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions.

The regression analysis across both negative and positive numbers showed significant SNARC in the intermixed condition, $M = .197$, $SD = .306$, $t(32) = 3.698$, $R^2 = .496$, but not in the blocked condition, $M = .084$, $SD = .309$, $t(32) = 1.573$, $R^2 = .082$. As can be seen in Figure 19, upward responses were faster than downward responses for the numerically smallest pairs but downward responses were faster than upward responses for the relatively larger pairs.

Figure 20. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

When the instruction was “warmer” the direction of SNARC reversed such that upward responses were faster than downward for the higher temperature pairs but downward responses were faster than upward responses for relatively lower temperature pairs. As seen in the left panel of Figure 20, this pattern was reliable for negative pairs in the blocked condition without a thermometer, $M = -.259$, $SD = .463$ $t(32) = -3.21$, $R^2 = .882$, but not in the intermixed condition, $p = .477$. As seen in the right panel of Figure 20, SNARC was not significant for positive pairs in the blocked condition without a thermometer $p = .07$, or in the intermixed condition, $p = .69$.

Figure 21. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions.

The regression analysis across both negative and positive numbers showed significant SNARC in the intermixed condition, $M = -.189$, $SD = .351$, $t(32) = -3.085$, $R^2 = .548$, but not in the blocked condition, $M = -.121$, $SD = .395$, $t(32) = -1.764$, $p = .09$, $R^2 = .069$. As can be seen in Figure 21, downward responses were faster than upward responses for the numerically smallest pairs but upward responses were faster than downward responses for the relatively larger pairs.

Figure 22. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

Interestingly, as seen in the left panel of Figure 22, when a thermometer was present, the test did not quite reach significance in the blocked condition, for negative numbers with the instruction, “warmer”, $M = -.151$, $SD = .454$, $t(32) = -1.92$, $p = .06$, nor for the intermixed condition, $p = .739$. However, for positive pairs in the blocked condition with a thermometer, as observed in the right panel of Figure 22, SNARC was significant, $M = -.152$, $SD = .347$, $t(32) = -2.51$, $R^2 = .605$, but not in the intermixed condition, $p = .51$. Consistent with the other conditions, upward responses were faster than downward for the relatively higher temperature pairs but downward responses were faster than upward responses for relatively lower temperature pairs.

Figure 23. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions.

When the data was regressed across the entire range of stimuli there was a significant SNARC effect in the blocked condition, $M = -.177$, $SD = .326$, $t(32) = -3.124$, $R^2 = .372$, but not in the intermixed condition, $M = -.115$, $SD = .354$, $t(32) = -1.87$, $R^2 = .292$. As can be seen in Figure 23, downward responses were faster than upward responses when for the numerically smaller pairs but upward responses became progressively faster than downward responses as pairs increased in size.

The results from this experiment suggest that the SNARC effects I found in Experiment 1 may not have been representative of the processing of numbers according to their absolute value but rather, processing of numbers according to their value on a full number line. The direction of SNARC however, was influenced by instruction. It seems that the most extreme temperatures corresponding to the instruction facilitated upward responses while more moderate temperatures facilitated downward responses.

Error Rates

The overall error rate was low at 3.97 %. The correlation between response time and mean error rate was $r = .370$, $p < .001$, based on 96 cells representing, magnitude, instruction, polarity, condition, and thermometer. Therefore, lower accuracy was associated with longer response times, so there was no speed-accuracy trade-off. An ANOVA was conducted with mean error rate as the dependent variable and instruction, polarity, condition, magnitude, and button as the independent variables. There was a main effect of condition, $F(1, 32) = 9.043$, $MSE = .048$. As expected there were fewer errors in the blocked condition (3.05 %) compared to the intermixed condition (4.89 %).

Condition (blocked, intermixed) interacted with numerical magnitude $F(5, 160) = 6.224$, $MSE = .004$. In the blocked condition, errors were lowest when numerical magnitude was low and high but highest for the middle pairs. In the intermixed condition, the opposite occurred. Errors were highest for the smallest and largest pairs and lower for the middle pairs. The interaction between polarity and magnitude was significant, $F(5, 160) = 6.736$, $MSE = .041$. Like in Experiment 2, generally, for negative pairs errors

decreased as magnitude increased towards zero, but for positive pairs, errors increased as magnitude increased.

There was a significant three-way interaction for instruction, condition, and magnitude $F(5, 160) = 2.284, MSE = .005$. In the blocked condition, when the instruction was “colder”, errors gradually increased as numerical magnitude increased, however for the instruction “warmer” the error rate stayed relatively the same across numerical magnitude. In the intermixed condition, errors seem to peak somewhere towards the middle across the pairs for each instruction.

There was an interaction for polarity, condition, and magnitude $F(5, 160) = 2.826, MSE = .004$. Errors tended to be greater in the intermixed condition but the differences in error rates between the blocked and intermixed conditions varied in an idiosyncratic manner across magnitude for positive and negative pairs. There was also a three-way interaction between instruction, magnitude, and response button $F(5, 160) = 3.639, MSE = .004$. Differences in error rate between left and right responses varied in an idiosyncratic manner across numerical magnitude for each instruction.

One potentially important result is the significant interaction of instruction, polarity, and response button $F(1, 32) = 4.869, MSE = .004$. Generally, when the instruction was “colder” upward responses ($M = .045$) to negative pairs were less accurate than downward responses ($M = .041$), but for positive pairs, downward responses ($M = .038$) were less accurate than upward responses ($M = .037$). However, when the instruction was “warmer” the opposite pattern occurred and the differences were slightly more pronounced. For positive pairs, upward responses ($M = .033$) were less accurate than downward responses ($M = .029$), but for negative pairs, downward responses ($M = .030$) were less accurate than

upward responses ($M=.025$). This effect is reflective of a speed/accuracy trade off for the response time advantages that occurred due to the interaction of instruction, polarity, and response button described on earlier.

There were two 4-way interactions including that between instruction, polarity, thermometer, and response button, $F(1, 32) = 5.101$, $MSE = .003$; and between instruction, condition, thermometer and response button, $F(1, 32) = 5.865$, $MSE = .002$. There was also a five-way interaction between instruction, condition, thermometer, magnitude and response button $F(5, 160) = 2.979$, $MSE = .003$. These interactions were not easily interpretable in a meaningful manner.

Discussion

This experiment produced several important findings including: 1) SNARC occurs in the vertical dimension, not only for positive digits but negative as well; 2) The direction of SNARC is not fixed but depends on instruction; 3) The mental number line extends past zero to include negative numbers; 4) SNARC was more robust when a thermometer was included.

The SNARC effect is relatively under-established in the literature for the vertical dimension in comparison to the horizontal dimension. To my knowledge this is the first time someone has found SNARC for negative numbers in the vertical dimension. Previously, SNARC effects in the vertical dimension were found that were consistent with a mental number line ordered from bottom to top, as small numbers were responded to faster in a downward direction and larger numbers were responded to faster in an upward direction, Ito and Hatta (2004).

My results are only partially consistent with these studies because they found SNARC effects that were independent of instruction, unlike what I found. I found that upward responses were always facilitated by the most extreme pairs corresponding to the instruction, but less compatible pairs facilitated downward responses. Therefore when the instruction was “warmer” the highest temperature pairs were responded to faster in an upward direction compared to downward and the lowest temperature pairs were responded to faster in a downward direction compared to upward. However, when the instruction was “colder” the reverse occurred; the lower temperature pairs were now responded to faster in an upward direction compared to downward but higher temperature pairs were responded to faster in a downward direction compared to upward. My results do in fact support a hypothesis of a full mental number line as it seems clear that as pairs became more negative, participants judged them as smaller.

Additionally, like in Experiment 3, I found a semantic congruity effect that was consistent with a full mental number line hypothesis. If participants processed numbers according to absolute value the semantic congruity effect would look the same for both positive and negative numbers. Instead I can see that when the instruction was “colder”, responses for negative numbers were faster than for positive numbers, but when the instruction was “warmer”, responses for positive numbers were faster than for smaller numbers. My results are consistent with a hypothesis that the mental number line is flexible in that can change dimensions from horizontal to vertical and that the direction is reversible depending on context. Another way to interpret the results from all of the experiments thus far is that participants always order “the least” on the left or on the bottom and “the most” to the right or on top.

EXPERIMENT 5

As far as I know, the results of Experiment 4 are the first time SNARC effects have been found for negative numbers in the vertical plane. I decided to repeat the experiment with some slight modifications. I reasoned that colouring the stimuli red or blue as a substitute for a written instruction may cause unnecessary work for participants and potentially distort response times. Therefore I returned to the written instructions used in my previous experiments.

Method

Participants

Thirty-two Carleton University students participated in a single session lasting approximately 45 minutes, in exchange for course credit.

Stimuli and Design

The stimuli and design were the same as Experiment 4 except that stimuli was coloured black and the written instruction was included below the digit pairs at approximately the bottom third of the screen in “Arial” font, size 36.

Procedure

The procedure was the same as in Experiment 4 except that participants used the Cedrus response pad (model RB-620). The response pad was positioned and taped to the desk, such that the keys were arranged in a vertical plane. Also, the experiment was executed using SuperLab 2 software.

Results

The data used in the analysis were from all 32 participants. Separate analyses will be presented for response time data, followed by an analysis for accuracy data.

Response time

After combining the data over the two presentation orders, medians for each cell, from each participant were averaged. Cells that exceeded the mean plus three standard deviations were replaced with the mean plus three standard deviations for that condition. Cells that were missing were replaced with (the subject mean + the cell mean – the grand mean). A total of 101 medians were replaced.

None of the interactions that would indicate a SNARC effect were significant from the ANOVA. I explore the SNARC effect further in the next section, “Individual Participant Analyses”. An ANOVA was then conducted with mean correct RTs as the dependent variable and the six magnitude pairs, two polarities of the number pair (negative, positive), the two instructions (colder, warmer), two levels of thermometer picture (present or not) and the direction of response as the with-in participant factors. Significance levels for analysis of variance (ANOVA) were based on the Huyn-Feldt adjusted degrees of freedom, although the degrees of freedom and error mean squares provided are those specified by the design. The level of significance used for all tests was $\alpha = 0.05$.

The ANOVA showed that there was a significant effect for polarity $F(1, 31) = 48.896$, $MSE = 115242.710$. Participants responded faster to positive pairs ($M = 756.722$ ms), compared to negative pairs, ($M = 816.039$ ms). Additionally, participants were faster in the blocked condition ($M = 652.491$ ms) compared to the intermixed condition

($M = 920.270$ ms), $F(1, 31) = 264.702$, $MSE = 416090.418$. As in Experiment 4, participants were significantly faster to respond when there was a picture of a thermometer on the screen ($M = 750.584$ ms) compared to when there was not, ($M = 822.177$ ms), $F(1, 31) = 34.267$, $MSE = 229752.888$.

Generally, response times differed significantly depending on the magnitude of temperature, $F(5, 155) = 13.320$, $MSE = 24823.752$. More specifically, as can be seen in Figure 24, response times were generally faster as magnitude increased for negative numbers but slower as magnitude increased for positive numbers. Therefore responses for pairs generally became faster as they approached zero, resulting in a significant interaction of polarity (negative, positive) and magnitude of temperature, $F(5, 155) = 36.674$, $MSE = 35099.543$. The general decrease in participants' responses as absolute magnitude increased, demonstrates Weber-like properties of participants' response times.

Semantic Congruity Effects

Figure 24. Mean response times collapsed across thermometer presentation, for each instruction and condition, as a function of each stimulus pair. The x-axis represents the order of temperature with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).

Semantic congruity effects were evident from the significant interaction between instruction and polarity, $F(1, 31) = 46.773$, $MSE = 114259.928$. As shown in Figure 24, in both the blocked and intermixed conditions, when numbers were negative, responses were faster for the instruction “colder” compared to “warmer”, but when pairs were positive, responses were faster when the instruction was “warmer” compared to “colder”. Furthermore, the interaction between instruction, polarity, and magnitude of temperature

pair was significant, $F(1, 31) = 28.251$, indicating graded semantic congruity effects. As can be seen in Figure 24, generally, responses were slower as absolute magnitude increased (due to Weber's law) which seems to mask the semantic congruity effects somewhat, but for positive pairs, when the instruction was "colder", and for negative pairs when the instruction was "warmer", a graded increase in response time can be seen as temperature increases or decreases respectively. For negative pairs with the instruction "colder" and positive pairs with the instruction "warmer" the graded semantic congruity effect is more difficult to detect, but the response times decrease more at the extreme temperatures as the semantic congruity effect begins to "over power" the magnitude effect.

Polarity (negative, positive) also interacted significantly with condition (blocked, intermixed), $F(1, 31) = 15.294$, $MSE = 83046.066$. It seems that the interaction occurred because the difference of the increase in response time from blocked to intermixed trials was greater for negative numbers ($M = 296.53$ ms) compared to positive numbers ($M = 239.024$ ms). The interaction of polarity, condition and numerical magnitude was significant, $F(5, 155) = 9.766$, $MSE = 13938.969$. Observation of Figure 24 shows that polarity differences between blocked and intermixed conditions differs across magnitude.

Polarity interacted significantly with thermometer, $F(1, 31) = 7.992$, $MSE = 36415.718$. Trials with a thermometer were always faster than without, however, the difference between trials with and without a thermometer was greater for negative numbers ($M = 85.36$ ms) compared to positive numbers ($M = 57.83$ ms).

Condition (blocked, intermixed) interacted with thermometer, $F(1, 31) = 25.207$, $MSE = 88297.259$. Again trials with a thermometer were always faster than those without

but the difference between trials with and without a thermometer was greater in the intermixed condition ($M = 109.66$ ms), compared to the blocked condition ($M = 33.53$ ms). Instruction (colder, warmer) interacted significantly with magnitude of temperature, $F(5, 155) = 10.621$, $MSE = 23447.819$. Response times across pairs, seemed to change in an idiosyncratic manner for each instruction. There was also an interaction for polarity and response button, $F(1, 31) = 8.212$, $MSE = 17086.935$. Downward responses ($M = 809.33$ ms) were faster than upward responses ($M = 822.75$ ms) for negative pairs, but for positive pairs, upward responses ($M = 753.87$ ms) were faster than downward responses ($M = 759.57$ ms). There was also an interaction for condition (blocked, intermixed) and response button, $F(1, 31) = 5.526$, $MSE = 26519.993$. In the blocked condition, upward responses ($M = 649.54$ ms) were slightly faster than downward response times ($M = 655.54$ ms), however, in the intermixed condition, downward responses ($M = 913.46$ ms) were faster than upward responses ($M = 927.09$ ms).

There was a three-way interaction of polarity, condition, and thermometer, $F(1, 31) = 5.028$, $MSE = 38053.369$. The increase of the difference in response time between thermometer and non-thermometer trials, from the blocked condition to the intermixed condition was greater for negative numbers ($MD = 98.45$ ms) compared to positive numbers ($MD = 53.81$ ms). The interaction between instruction, polarity, and response button (top, bottom) was significant, $F(1, 31) = 28.251$, $MSE = 30094.027$. It seems as though this interaction occurred because upward responses were faster than downward responses in every condition except for positive numbers when the instruction was “warmer”.

Also there was one four-way interaction between instruction, polarity, condition and response direction $F(1, 31) = 9.918$, $MSE = 12305.148$ which could not be interpreted in a meaningful way. There were no other significant results.

Individual participant analyses

To further investigate the possibility of a SNARC effect, I conducted repeated measures regression analyses as outlined in Lorch and Myers (1990), separately with each instruction (colder, warmer), each polarity (positive, negative), condition (blocked, intermixed) and thermometer condition (present or not). After combining the data over order, I first obtained standardized beta regression coefficients for each participant with the SNARC index (Upward RT - Downward RT) as the dependent variable and magnitude (temperature) of stimulus pair as the predictor. I then tested the hypothesis (two tailed) that the mean of these weights differed from zero using the t-distribution. R^2 values were calculated using group means rather than individual means.

Generally, the results of Experiment 5 replicated those of Experiment 4. Again SNARC effects were dependent on instruction such that, generally, participants responded faster in an upward direction compared to downward, for more extreme temperature pairs corresponding to the instruction, but faster to downward pairs compared to upward pairs that corresponded relatively less with the instruction.

Figure 25. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

For the instruction “colder”, SNARC was significant in the blocked condition when a thermometer was present for negative numbers, $M=.152$, $SD=.362$, $t(31)=3.383$, $R^2=.062$. As seen in the left panel of Figure 25, in the blocked condition, upward responses were faster compared to downward, when the numbers were most negative but as the pairs increase in temperature, downward responses became faster compared to upward. Also shown is the nonsignificant effect for the intermixed condition, $p=.31$. The right panel of Figure 25 displays the nonsignificant SNARC effect for positive numbers in the blocked, $p=.13$ and intermixed conditions, $p=.66$ with a thermometer present.

Figure 26. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, with a thermometer present, in blocked and intermixed conditions.

Figure 26 shows SNARC data regressed across the entire range of temperatures. As can be seen, there were significant SNARC effects in both the blocked condition, $M = .195$, $SD = .348$, $t(31) = 3.172$, $R^2 = .213$ as well as the intermixed condition, $M = .121$, $SD = .304$, $t(31) = 2.248$, $R^2 = .129$. Upward responses were faster compared to left for the numerically smaller pairs but downward responses were faster than upward for larger pairs.

Figure 27. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

Figure 27 displays the nonsignificant results that occurred without a thermometer for the instruction “colder” in the negative blocked condition, $p = .61$, negative intermixed condition, $p = .16$, positive blocked condition, $p = .19$ and the positive intermixed condition, $p = .27$.

Figure 28. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “colder”, without a thermometer present, in blocked and intermixed conditions.

When SNARC data was regressed across both negative and positive temperatures there was no significant effect for either the blocked condition, $M = .058$, $SD = .316$, $t(31) = 1.045$, $R^2 = .095$ or the intermixed condition, $M = .039$, $SD = .303$, $t(31) = .720$, $R^2 = .024$. As seen in Figure 28, upward responses were faster than downward responses for the numerically smaller pairs but downward responses were faster than upward responses for numerically larger pairs.

For the instruction “warmer”, SNARC was significant for positive numbers with a thermometer present, in the blocked condition, $M = -.239$, $SD = .401$, $t(31) = -3.368$, $R^2 = .453$, as well as the intermixed condition, $M = -.171$, $SD = .447$, $t(31) = -2.166$, $R^2 = .503$. As seen in the right panel of Figure 29, the direction of SNARC is consistent with previous results; upward responses were faster compared to downward, for the relatively higher temperature pairs but downward responses were faster than upward for the relatively lower temperature pairs. Also shown in the left panel of Figure 29 are the nonsignificant results for the negative pairs in the blocked condition as well as the intermixed condition.

Figure 29. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

Figure 30. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, with a thermometer present, in blocked and intermixed conditions.

When SNARC data was regressed across the entire range of temperatures there was a significant effect for both the blocked condition, $M = -.201$, $SD = .323$, $t(31) = -3.508$, $R^2 = .405$, as well as for the intermixed condition, $M = -.163$, $SD = .378$, $t(31) = -2.446$, $R^2 = .652$. As seen in Figure 30, downward responses were faster than upward responses for numerically smaller pairs but upward responses were faster than downward responses for relatively larger pairs.

Displayed in the panels of Figure 31 are the nonsignificant results for the instruction “warmer” without a thermometer for negative numbers in the blocked $p = .17$, and intermixed conditions, $p = .82$, as well as for positive numbers in the blocked, $p = .35$, and intermixed conditions, $p = .90$.

Figure 31. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

Figure 32. Mean downward response time subtracted from median upward response times as a function of each stimulus pair, for the instruction “warmer”, without a thermometer present, in blocked and intermixed conditions.

When SNARC data was regressed across the entire range of temperatures there was a significant effect for both the blocked condition, $M = -.207$, $SD = .323$, $t(31) = -3.627$, $R^2 = .381$, as well as for the intermixed condition, $M = -.221$, $SD = .223$, $t(31) = -5.603$, $R^2 = .785$. As seen in Figure 32, downward responses were faster than upward responses for numerically smaller pairs but upward responses were faster than downward responses for relatively larger pairs.

Error Rates

The overall error rate was low at 4.21 %. The correlation between response time and mean error rate was $r = .333$, $p < .001$. Therefore, lower accuracy was associated with longer response times, so there was no speed-accuracy trade-off.

An ANOVA was conducted with mean error rate as the dependent variable and instruction, polarity, condition, magnitude, and button as the independent variables. There was a main effect of condition, $F(1, 31) = 10.122$, $MSE = .039$. As expected there were fewer errors in the blocked condition (2.90 %) compared to the intermixed condition (4.49 %). There was a main effect of temperature pair, $F(5, 155) = 5.034$, $MSE = .009$. Error rates seemed to vary in an idiosyncratic manner.

The interaction between polarity and magnitude was significant, $F(5, 155) = 23.276$, $MSE = .008$. Like in Experiment 4, for negative pairs, errors decreased as magnitude increased towards zero, but for positive pairs, errors increased as magnitude increased. The interaction between condition and magnitude was significant, $F(5, 155) = 4.421$, $MSE = .006$. Error rates seemed to vary in an idiosyncratic manner across magnitude for both blocked and intermixed conditions. There was an interaction between condition and response button, $F(1, 31) = 4.925$, $MSE = .005$. In the blocked condition there was a slightly higher error rate for “down”, but for the intermixed condition there was a slightly higher error rate for “up”.

There was a three-way interaction for instruction, polarity and condition, $F(1, 31) = 12.966$, $MSE = .019$. Although the intermixed condition always had more errors than the blocked condition, the error rate between the blocked and intermixed condition was more

comparable for negative pairs, when the instruction was “warmer” and for positive pairs when the instruction was “colder”, compared to other pairs.

There was a three-way interaction between instruction, numerical magnitude and response button, $F(1, 155) = 3.132$, $MSE = .005$. There were slight variations between up and down error rates for each condition. There was also an interaction for polarity, condition, and magnitude, $F(5, 155) = 3.261$, $MSE = .006$. Error rates seemed to vary in an idiosyncratic manner for each condition.

Discussion

This experiment was a replication of Experiment 4 except that the number pairs, which were coloured as red or blue in Experiment 4, now appeared as black, and the instruction was printed on the screen during each trial. The results of this study generally replicate those of Experiment 4. I found SNARC for positive numbers with the instruction “warmer” and for negative numbers with the instruction “colder”. However, SNARC only approached significance for positive numbers with the instruction “colder” and for negative numbers with the instruction “warmer”. The direction of SNARC once again could be reliably predicted based on instruction and polarity. When the instruction was “colder”, upward responses were faster compared to downward for lower temperature pairs but downward responses were faster than upward responses for higher temperature pairs. When the instruction was “warmer”, the opposite was true; upward responses were faster for higher temperature pairs and downward responses were faster for lower temperature pairs.

Once again my results are consistent with the hypothesis that the mental number line exists in the vertical dimension and includes negative numbers and that the order of the

line can be reversed according to instruction. An important question arises however, which is, “What advantage would there be to reorder the line for one instruction or the other?” The most obvious answer might be that participants have learned over time from day to day life that generally, “more is always on top” so that ordering stimuli in this fashion may allow for the most efficient processing. However, in order to make this argument, I would also have to make the argument that participants map numbers onto dimensions of “cold” and “warm” separately, which is possible.

The results in this experiment are consistent with the results from Experiments 3 and 4. SNARC seems to be dictated by instruction and polarity and this does not seem to change between blocked and intermixed conditions. Additionally, it seems as though participants are processing numbers according to their numerical value on a full mental number line. However, the order of numbers can shift based on instruction such that the least is on the bottom and the most is on the top. Also, like in Experiments 3 and 4, I found a semantic congruity effect that is reflective of numbers processed along a full mental number line that includes negative numbers, suggesting a full mental number line.

It is still not clear why the context of temperature would induce an ordering of numbers with “the least” on the left or bottom, and “the most” on the right or top. One hypothesis is that the SNARC effect in general, is *always* reflective of an ordering from least on the left or down, to most on the right or up, *as interpreted by the participant*, no matter what numbers represent these semantic properties.

Therefore, it could be that the actual magnitudes of the numbers used in the experiment are not that important for generating SNARC. It could be that it is the semantic value of stimuli from least to most on a dimension that is dictated by the

instruction that is responsible for the direction of SNARC. For instance, in Experiment 4 there was SNARC for positive numbers with the instruction “colder”. The direction of SNARC suggested that participants had mentally ordered numbers with the largest pairs on the bottom (9, 8) and the smallest pairs on top (1, 2) because that represents the least cold to most cold.

An additional facet of my results that is worth discussion is that, my results were more robust for conditions when a thermometer was present compared to when it was not. Interestingly, when regression analyses were conducted for positive and negative pairs separately, the SNARC effect was only significant for one condition without the presence of a thermometer. I had previously hypothesized that simply by adding a thermometer to the screen might emphasize, diminish, or even reverse the SNARC effect. The thermometer seems to have enhanced my results rather than to have reversed them. The one condition in which a nonthermometer trial was significant, showed that the direction of SNARC was the same as the corresponding thermometer condition that just failed to reach significance ($p=.06$). This finding shows that the context in which numbers are presented is important not only in determining the direction of SNARC but also the strength of SNARC.

Hypothetically, I could have expected the SNARC effect to diminish with a thermometer present because it could prime a mental number line ordered with the largest numbers on top and the most negative numbers at the bottom, like the ordering of a thermometer. In fact though, the thermometer seemed to enhance SNARC even when the direction suggested that participants had ordered numbers with the most negative on top and most positive on the bottom. This, along with my SNARC results from the colder-

positive and warmer-negative conditions, that seem to contradict the ordering of a mental number line, suggest that perhaps, the idea that a mental number line that is reversible is not the best explanation for SNARC.

Recently Proctor and Cho (2006) suggested that a previous model from Seymour (1973) is the best model to account for a variety of effects involving binary choice tasks including the SNARC effect. Proctor and Cho (2006) recently brought to light, the pervasiveness of the model's applications for binary choice tasks, including SNARC experiments. The premise of Seymour's model is that relevant and irrelevant features of stimuli as well as the possible dichotomous responses are inherently coded as either positive (+) or negative (-), based on their positions on the underlying dimensions they represent. Therefore, both stimuli and possible responses are coded as positive (+) or negative (-) to produce compatible, (+, +), (-, -) or incompatible (+, -) (-, +) combinations. Compatible combinations result in quick responses, but incompatible pairs result in longer responses because the instruction must be recoded to match. Pertaining to many SNARC experiments, responding left (-) to small (-) numbers would produce quick responses as well as right (+) to large (+) numbers. Slow responses however, would occur for left responses to large numbers (-, +) and right responses to small numbers (+, -).

Generally, Seymour's model can account for SNARC effects including ours, just as well as a number line hypothesis; however, it would need to allow for recoding of stimuli based on relative size. For instance, numbers 1- 4 could not be inherently coded as small (-) but rather sometimes coded as small (-) or large (+) depending on the context (relative size within subset or perhaps instruction). To account for the results of my study I would have to make the assumption that participants code numbers on dimensions of cold and

warm separately with “relatively not cold” coded as (-), and “relatively cold” coded as (+). Additionally, in order to account for a graded SNARC effect (progressive from 1-9), as opposed to a simple categorical SNARC effect (difference between Small and Large numbers) the codes would need to be weighted such that numbers closest to the poles on each dimension are given the most weight.

Generally my results thus far, show that SNARC occurs in the horizontal and vertical direction for temperatures represented as positive and negative numbers. Participants seem to process numbers according to their numerical value representing a mental number line that includes negative numbers, which is reversible depending on instruction. Alternatively, stimuli and responses may be coded as (-) and (+) which correspond or conflict with each other to produce varying response times resulting in the SNARC effects I found.

I have hypothesized that what determines the direction of SNARC is the semantic value of numbers from least to most, even when this contradicts their numerical value. Therefore if participants view the smallest numbers as having the most value, and the largest numbers as having the least value then SNARC should reflect the semantic value, rather than numerical value. Experiment 6 replicates Experiment 3 with some modifications, in order to test this.

Experiment 6

Up to this point there is evidence from all of my experiments to suggest that the SNARC effects that I have found have been caused not so much by the magnitude of numbers but rather, an ordering of magnitude from least on the left or down and the most on the right or up. The following experiment will expand on the results of Experiments 3-

5. In my previous experiments, instruction has determined the direction of SNARC. It is possible that the origin of this SNARC effect is a full mental number line that reorients itself to adapt to a particularly task, or there could be another explanation.

In the following experiment I will test if participants will arrange numbers according to their underlying semantic dimension with the least on the left and the most on the right, as they did before. However, in this experiment participants will be told that negative numbers and positive numbers represent the same magnitude. Each number from 1 to 9 will represent a runner with 1 being the fastest and 9 being the slowest. Participants will be told that the negative forms of these numbers (-1 to -9) will represent the exact same runners, but with a different name, so 1 and -1 is a single runner and is always the fastest, and 9 and -9 is also a single runner and always the slowest. Participants were required to choose the “faster” or “slower” runner, depending on the given instruction.

I continued to present the trials in blocked and intermixed conditions. If the semantic dimension is what is important for determining SNARC, then SNARC should always be in the same direction for positive and negative numbers. I hypothesize that when the instruction is “slower”, SNARC will reflect that participants ordered numbers from 1 and -1 on the left, progressing to 9 and -9 on the right. For the instruction “faster”, I predict that SNARC will reflect an ordering of numbers from the least fastest (9, -9) on the left to the fastest on the right (1, -1).

Method

Participants

Twenty-one Carleton University students participated in a single session lasting approximately 45 minutes, in exchange for course credit.

Procedure

The stimuli, design, and procedure were exactly the same for Experiment 6 as for Experiment 3 except that I exchanged the instruction “colder” with “slower” and “warmer” with “faster”. Additionally, it was explained to participants that 1 and -1 represented a runner who was the fastest and 9 and -9 represented a runner who was the slowest, as previously mentioned. On each trial, participants were required to choose the “faster” or “slower” runner from the presented pair.

Results

The data used in the analysis were from 20 participants. Data from one participant was not included in the analysis because accuracy was low and there were too many responses that were below 200 ms. Separate analyses will be presented for response time data, followed by an analysis for accuracy data.

Response time

After combining the data over the two presentation orders, medians for each cell, from each participant were averaged for each combination of instruction and polarity. Cells that exceeded the mean plus three standard deviations were replaced with the mean plus three standard deviations for that condition. A total of 50 medians were replaced.

An ANOVA was then conducted with mean correct RTs as the dependent variable and the six magnitude pairs, two polarities of the number pair (negative, positive), the

two instructions (slower, faster), and the direction of response (left, right) as the within participant factors. Significance levels for analysis of variance (ANOVA) were based on the Huyn-Feldt adjusted degrees of freedom, although the degrees of freedom and error mean squares provided are those specified by the design. The level of significance used for all tests was $\alpha = .05$.

Semantic Congruity Effects

Figure 33. Mean response times, for each instruction and condition, as a function of each stimulus pair. The x-axis represents pairs of “runners” with -6 representing the most negative pair (-9, -8) progressing to 6 representing the most positive (9, 8).

None of the interactions that would indicate a SNARC effect were significant from the ANOVA. I explore the SNARC effect further in the next section, “Individual Participant Analyses”. The ANOVA showed that, interestingly, participants responded faster when the instruction was “slower” ($M = 560.110$ ms), compared to “faster” ($M = 583.975$ ms), $F(1, 19) = 4.414$, $MSE = 61932.704$. Participants also responded faster to positive pairs ($M = 562.239$ ms), compared to negative pairs, ($M = 581.856$ ms), $F(1, 19) = 15.906$, $MSE = 11625.375$.

Generally, response times differed significantly depending on the numerical magnitude, $F(5, 95) = 9.931$, $MSE = 6068.149$. More specifically, as can be seen in Figure 33, response times were fastest for pairs closest to zero but generally slowed as absolute magnitude increased for both positive and negative numbers. This magnitude effect was confirmed by a significant interaction of polarity (negative, positive) and magnitude of temperature, $F(5, 95) = 45.092$, $MSE = 8554.800$.

A semantic congruity effect was indicated by a significant interaction of instruction, polarity and magnitude, $F(5, 95) = 3.037$, $MSE = 4135.005$. As seen in Figure 33, responses for the most negative pairs were faster when the instruction was “slower” ($M=597.31$ ms) compared to “faster” ($M=625.42$ ms). However, as the speed represented by the pair increased towards -1, there was little difference in response time for pairs between the instructions “faster” ($M=553.49$ ms) and “slower” ($M=551.21$ ms). However, it is unclear from the figure if there is a semantic congruity effect for positive pairs as well. The means however indicate that a similar pattern may be developing. Responses to the most positive (slowest) pairs were faster when the instruction was “slower” ($M=576.352$ ms) compared to “faster” ($M=615.08$ ms). For smaller (faster) pairs even

though there was still a speed advantage when the instruction is “slower” ($M=515.57$ ms) compared to the instruction “faster” ($M=541$ ms), it had decreased, indicative of a semantic congruity effect.

The interaction of polarity and instruction was significant, $F(1, 19) = 4.978$, $MSE = 7247.908$. The speed advantage of positive pairs over negative pairs decreased from the instruction “slower” ($MD = 28.30$ ms), to the instruction “faster” ($MD = 10.96$ ms).

There was a significant interactions for instruction and magnitude of pair $F(5, 95) = 3.644$, $MSE = 2840.358$. Responses were always faster for the instruction “slower” compared to “faster” but this difference was generally greater for pairs of smaller numerical magnitude compared to larger numerical magnitude.

There was a significant interaction for instruction, polarity and response button, $F(1, 19) = 4.978$, $MSE = 7247.908$. For negative pairs, left responses ($M = 573.00$ ms) were faster than right responses ($M = 575.52$ ms) when the instruction was “slower”, but when the instruction was “faster”, right responses ($M = 585.96$ ms) were faster than left responses ($M = 592.95$ ms). For positive pairs, the opposite pattern occurred. When the instruction was “slower” right responses ($M = 540.97$ ms) were faster than left ($M = 550.96$ ms) but when the instruction was “faster” left responses ($M = 574.89$ ms) were faster than right responses ($M = 582.10$ ms).

Finally, there was an interaction for polarity, magnitude of pair, and response button $F(5, 95) = 3.277$, $MSE = 3541.563$. For positive and negative pairs, there were slight variations of the differences between left and right responses across magnitude.

Individual participant analyses

To further investigate the possibility of a SNARC effect, I conducted repeated measures regression analyses as outlined in Lorch and Myers (1990), separately with each instruction (slower, faster), each polarity (positive, negative), and condition (blocked, intermixed). After combining the data over order, I first obtained standardized beta regression coefficients for each participant with the SNARC index (Right hand RT – Left-hand RT) as the dependent variable and magnitude (speed) of stimulus pair as the predictor. I then tested the hypothesis (two tailed) that the mean of these weights differed from zero using the t-distribution. R^2 values were calculated using group means rather than individual means.

Significant effects were only attained with the instruction “slower” but were consistent with the hypothesis that stimuli would be ordered from “least” to “most” according to semantic magnitude. However, in this condition, semantic magnitude and absolute magnitude were compatible.

Figure 34. Mean left-hand response time subtracted from median right hand response times as a function of each stimulus pair, for the instruction “slower”, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

For the instruction “slower”, SNARC was significant in the blocked condition for negative numbers, $M=.287$, $SD = .325$, $t(19) = 3.946$, $R^2 = .677$. As seen in the left panel of Figure 34, in the blocked condition, right hand responses were faster compared to left-hand responses when the numbers were most negative but as the pairs increased in “speed” or become less negative, left-hand responses became faster compared to right-hand. Also shown is the nonsignificant effect for the intermixed condition, $M=.183$, $SD=.429$, $t(19)=1.913$, $p=.07$, $R^2 = .711$.

As seen in the right panel of Figure 34, SNARC was also significant for positive numbers in the blocked condition for the instruction, “slower” $M= -.252$, $SD = .445$, $t(19) = -2.53$, $R^2 = .864$. Again, for numbers of small magnitude, left-hand responses were faster compared to right, but for larger numbers, right hand responses are faster than left.

Also shown are the nonsignificant results for the intermixed condition, $M = -.154$, $SD = .435$, $t(19) = -1.577$, $p = .13$, $R^2 = .450$.

Figure 35. Mean left-hand response time subtracted from median right hand response times as a function of each stimulus pair, for the instruction “faster”, in blocked and intermixed conditions. The left panel represents negative pairs and right panel represents the positive pairs.

For the instruction “faster”, SNARC was not significant in any condition. The left panel of Figure 35 shows the blocked condition for negative numbers, $M = .123$, $SD = .599$, $t(19) = .915$, $p = .37$, $R^2 = .398$. Also seen in the left panel of Figure 35 is the intermixed condition for negative numbers which actually approached significance $M = .172$, $SD = .459$, $t(19) = 1.67$, $p = .11$, $R^2 = .492$. Again, it seemed like the possible SNARC was determined by the absolute magnitude of numbers. When numbers were least negative, or “fastest” (e.g., -1, -2), left-hand responses were faster compared to right hand responses, and when numbers were most negative, or “slowest”, right hand

responses were faster compared to left-hand. This effect, although not significant is in contrast to the SNARC effects found for negative numbers with the instruction “warmer”, and for positive numbers with the instruction “colder” in Experiment 4. SNARC in these conditions, like all the other conditions, was determined by the order of semantic magnitude even though it was not compatible with the absolute magnitude of the pairs.

The right panel of Figure 35 shows the nonsignificant results for the blocked condition with positive numbers and the instruction “faster”, $M=.112$, $SD = .459$, $t(19) = 1.089$, $p = .29$, $R^2 = .689$, as well as the mixed condition, significance $M=.055$, $SD = .523$, $t(19) = 0.469$, $p = .644$, $R^2 = .122$.

Error Rates

The overall error rate was low at 2.51 %. The correlation between response time and mean error rate was $r = -.077$, $p < .001$. Therefore, there was a weak but reliable relationship between faster response times and lower accuracy, showing a speed-accuracy trade-off.

An ANOVA was conducted with mean error rate as the dependent variable and instruction, polarity, condition, magnitude, and button as the independent variables. There was a main effect of instruction, $F(1, 19) = 10.137$, $MSE = .004$. The error rate was 2.03 % for the instruction, “slower” compared to 3.00 % for the instruction “faster”. There was a main effect of magnitude $F(5, 95) = 8.031$, $MSE = .003$. Errors were lowest for pairs next to zero and at the extremes, but otherwise seemed to differ in an idiosyncratic manner.

The interaction between polarity and magnitude was significant, $F(5, 95) = 2.784$, $MSE = .089$. This results from the fact that as magnitude increased for negative numbers

(approached zero) errors decreased and as magnitude of positive numbers increased (grew further away from zero) errors increased.

There was a three-way interaction for instruction, polarity and condition, $F(1, 1824) = 9.055$, $MSE = .002$. There was also an interaction for instruction, polarity and magnitude, $F(5, 95) = 6.272$, $MSE = .003$. Error rates seemed to vary in an idiosyncratic manner for each condition.

The four-way interaction between instruction, polarity, condition and magnitude $F(5, 95) = 2.418$, $MSE = .002$, shows that there was an increase in accuracy as the absolute magnitude of pairs decreased, consistent with Weber's law. However, this increase was more pronounced when the instruction was "faster" compared to "slower".

Discussion

I hypothesized that when participants were asked to choose the faster or slower runner, they would mentally arrange numbers from least on the left to most on the right, despite the numerical value of the stimuli, therefore generating opposing directions of SNARC for each instruction. In fact I only partially found this. When the instruction was "slower", the direction of SNARC occurred as I had anticipated. SNARC occurred in the same direction for both positive and negative numbers, ordered with the "fastest" (1, -1) on the left and the slowest (9, -9) on the right. This is consistent with my previous experiments that suggest that "the most" is ordered on the right and "the least" on the left, no matter what the numeric value of the numbers are.

Perhaps the more interesting part of this experiment however, was for the conditions with the instruction "faster" because "the most", in this condition, was defined by numbers with a small absolute magnitude (1, -1) which, intuitively, might be expected to

conflict with the concept of “the most”. In this condition, I hypothesized that SNARC would reflect an ordering of numbers from the “least fast” on the left (9, -9) and the “most fast” on the right (1, -1). In fact I did not find this. Due to a lack of power I failed to find significance for either positive or negative numbers. However, for negative numbers in the intermixed condition, SNARC approached significance and it was in the same direction as the SNARC I found for the instruction “slower”. In this condition, participants seemed to be processing numbers according to their absolute magnitude, not according to the “least” or “most” as I anticipated. Unfortunately I was unable to test more people for this experiment and unable to complete an additional experiment, using the reverse ordering of stimuli (i.e., 1, -1 are the slowest and 9, -9 are the fastest). My results are unclear from this experiment however I hypothesize that the absolute magnitude of numbers as well as the semantic magnitude may both play a role in determining the direction of SNARC.

General Discussion

Results from SNARC experiments prior to my experiments seem to suggest that the direction of SNARC was caused by a mental number line with small numbers on the left and larger numbers on the right. It was hypothesized that this occurs as a result of overlearning numbers in a certain direction in day-to-day life or by reading and writing habits in general (Dehaene et al., 1993). The direction of SNARC seems to reverse in cultures, in which reading and or writing occurs in the opposite direction of the Western world, supporting the notion that there is a mental number line with a learned order, which generates the SNARC effect.

In all of my experiments, the direction of SNARC depended on instruction, such that participants seemed to order the pairs representing the “least” semantic quantity on the left or downwards, and pairs representing “the most” semantic quantity on the right or upwards. My animal intelligence experiment and temperature experiments showed that SNARC could occur in opposing directions for a single stimulus set, when the semantic quantity assigned to it changed, because of the instruction. For instance, -9 was “large” when looking for something cold but “small” when looking for something warm and the SNARC effect reflected this.

I hypothesized that perhaps the numerical value of number didn’t actually matter at all, but rather that the direction of SNARC was determined by the semantic properties of numbers. “Normal” SNARC would only occur therefore when smaller numbers represented smaller semantic quantities and larger number represented larger semantic quantities. If positive and negative numbers held the same semantic properties, then SNARC should occur in the same direction for positive and negative numbers for each instruction. Additionally the absolute magnitude of the numbers shouldn’t influence the direction of SNARC. When I tested this in Experiment 6, my preliminary results suggested that positive and negative numbers could hold the same semantic properties (e.g. 9 and -9 were recognized as slowest), although generating a significant SNARC effect was difficult when the absolute magnitude of the pair was not compatible with the semantic magnitude of the pair.

If I consider Experiment 6, as well as all of the temperature experiments, I can see that generally, SNARC was easily generated when the task required participants to order numbers from left to right according to their absolute magnitude (1 to 9 or -1 to -9), when

this was compatible with the instruction-dependent semantic magnitude of least on the left to most on the right. For instance SNARC was consistently significant for negative pairs with the instruction “colder”, and with positive pairs and the instruction “warmer”, and also for positive and negative pairs with the instruction “slower”, in Experiment 6.

However when participants were given a task that required them to order the pairs from least to most but the absolute magnitude was large for the pair that represented “the least” and the absolute magnitude was small for the pair that represented “the most”, it became more difficult. When positive and negative pairs were analyzed separately, I found only a single significant condition for positive numbers with the instruction “colder”, and for negative numbers with the instruction “warmer”. Additionally, in Experiment 6, I failed to find a significant SNARC effect when the instruction was “faster” and the fastest pairs were also the pairs with the smallest absolute magnitude. This suggests that when the instruction was “faster” the effect may have weakened due to a conflict between the tendency to order numbers according to their absolute value, but also according to their semantic magnitude, “pulling” the ordering of numbers in the opposite direction, resulting in a nonsignificant SNARC effect.

I began a second experiment in which I reversed the meaning of the numbers such that the slowest runner was represented by (1, -1) and the fastest by (9, -9). I fully expect that in this experiment I will see a strong SNARC effect this time for the instruction “faster” as it is compatible with the ordering of the absolute value of the numbers, but also a weaker and perhaps reversed SNARC effect for the instruction “slower”.

I suggest that in my experiments, there are at least two factors that contribute to the direction of SNARC. One factor is indeed the tendency to order stimuli according to their

semantic quantity, with the perceived least on the left or downwards, and the perceived most on the right or upwards. The other contributing factor is the tendency to order numbers according to their absolute value with the smallest on the left or downwards and the largest on the right or upwards, at least in Western society. I would suggest that the instruction-dependent semantic magnitude was more potent because whenever, a significant SNARC effect occurred, it dictated the order. However I also believe that the absolute magnitude was strong enough to cancel the SNARC effect when it was not compatible with the order of semantic magnitude. When the absolute value of the numbers and the semantic quantity are compatible, SNARC will be robust. However, when the semantic magnitude and absolute magnitude are incompatible SNARC will be weakened or diminished.

One condition from my experiments that may be an exception to this rule is when the instruction was “faster” in Experiment 6. SNARC was not significant, although it was close for negative numbers in the intermixed condition. It suggested that the direction of SNARC was dictated by the absolute value of the numbers, the opposite direction of the semantic value. It may be that, as suggested, the direction of SNARC may not be “all or nothing” but rather an additive effect between a direction of SNARC suggested by the absolute value of numbers as well as that of the semantic value. As these components are strengthened or weakened, the direction of SNARC will may reflect this.

A problem for a hypothesis though, is that SNARC experiments that use the instructions “smaller” and “larger” with numbers should follow the same rule. When the instruction is “smaller” participants should have the tendency to order the “most small” on the right and the “least small” on the left. This should create a conflict resulting in a

weaker SNARC effect for the instruction “smaller” compared to “larger”. This does not necessarily occur though (Shaki and Petrusic 2005a, 2005b).

One possibility why number experiments in the past generally haven’t shown an instruction dependent SNARC effect might be that the numbers need some kind of underlying semantic dimension in order to be affected by instruction. In all of my experiments, the numbers represented some specific property like temperature or speed, compared to other SNARC experiments using numbers that are not specific. It is not clear why this would be but perhaps it could be that numbers are overlearned in a left to right sequence, such that participants can respond most efficiently using the “default” ordering, unless the numbers represent more specific semantic properties, in which case, reordering them from least to most becomes more efficient, for some reason.

One way to test this would be to test for SNARC using the instructions “smaller” and “larger” with non-specific positive numbers in one group and numbers representing the size of buildings in another. It is possible that for non-specific numbers, the direction of SNARC would be the same for both instructions but for building size they will be different.

Semantic Congruity Effects

The semantic congruity effects that I found were typical in the sense that response times were facilitated when the instruction corresponded to the semantic quantity of the stimuli and were slower when they were not. Again, semantic congruity effects have nothing to do with left or right responses, but help to verify that participants associated the stimuli along the dimension described by the instruction. When a scaled or gradual semantic congruity effect is found, it demonstrates that stimuli have a linear mental

ordering. Some of the semantic congruity effects I found however were statistically categorical, rather than scaled, but demonstrate a distinction between processing of positive numbers versus negative numbers. Importantly, what the semantic congruity effect showed, for the temperature experiments I conducted, is that participants seem to process negative numbers as smaller compared to positive numbers. This indicates that at some stage in processing, participants processed numbers according to their full numerical value rather than just their absolute value.

The Importance of the Thermometer

In the temperature experiments, conditions with a thermometer tended to generate more robust SNARC effects compared to conditions without a thermometer. Interestingly even when the apparent mental ordering of numbers was in the opposite direction of that of a thermometer, the SNARC effect was more robust. This suggests that perhaps the thermometer was important not because it suggested a particular ordering of numbers, but rather because it reinforced the semantic properties of the numbers.

Therefore, it seemed as though, the more that the numbers used in my experiment were thought of as temperatures, the stronger the SNARC effect was because hypothetically, as the semantic properties increased, so did the tendency to arrange numbers from “the least” to “the most” which is then reflected in the SNARC effect.

One way to test this would be to run a sequence of conditions with positive and negative numbers. In one condition participants would simply have to choose the *non-specific number* of the pair with the instructions “less” or “more”. Then in a second condition they would choose the *temperature* that is “less” or “more”. In a third condition they would choose the “*colder*” or “*warmer*” temperature. In a fourth condition they

could choose the “colder” (written in blue) or “warmer” (written in red) temperature. Finally, the fourth condition could be repeated with a thermometer on the screen. It is possible that in the first condition SNARC would reflect the absolute value of numbers, because they do not have a specific semantic representation. However, for some conditions, such as negative pairs with the instruction “warmer”, a diminished or reversed SNARC effect may occur as the semantic properties of the numbers become more salient across conditions.

Blocked Versus Intermixed

One of the goals of my experiment was to determine whether or not SNARC for temperatures would mimic SNARC of “normal” numbers. Shaki and Petrusic found that when negative pairs were compared exclusively, SNARC was dictated by absolute magnitude. However, when negative pairs were intermixed with positive pairs, SNARC for negative numbers seemed to be dictated by numerical value.

In comparison, in Experiments 3 through 6, there was no difference in the direction of SNARC for positive or negative numbers between blocked and intermixed conditions. I do not have any new insights to explain the difference Shaki and Petrusic (2005a) found between blocked and intermixed conditions in their experiment. However, this difference is consistent with the hypothesis participants may respond differently to numbers that do not have specific semantic properties attached to them, compared to numbers that do, such as temperatures.

SNARC Model

The model to account for SNARC proposed by Gevers et al. (2006) has some difficulty accounting for the results of my experiments for several reasons. Their model does not make any specifications for negative numbers; for pairs of numbers rather than single digits; for the influence of instruction; or for vertical responses. Additionally, I believe the model needs an additional automatic field, the semantic magnitude field, in order to account for my results.

Therefore I need to make some alterations to their model based on some of my assumptions, in order to make it viable. First, I will assume that negative numbers in the magnitude field of the model are processed according to their absolute value. Second, when the task involves a pair of digits for comparison, the average of the digits will be used and compared against the comparative standard, as described in their model. Third, depending on the requirements of the experiment, the direction of left and right must be interchangeable with downward and upward respectively. Last of all, in addition to the automatic magnitude field and the task driven standard field described in their model, I believe it needs an additional automatic path for coding the semantic magnitude of the pair. However I would also propose that the semantic field has more influence (at least sometimes) on the final decision, than the absolute magnitude field.

According to their model, on the automatic pathway, numbers that are smaller than the standard will always facilitate a left response and numbers larger than the standard will always facilitate a right response. When the response to the task is compatible with the automatic pathway, a quick response will occur, but when they are incompatible, a slower response will occur, thus generating SNARC. The additional, more influential semantic

magnitude field, would potentially enhance activation of the same node as the absolute magnitude or it may activate the other response node, in effect cancelling the activation of the first activated node, and possibly resulting in a weaker activation in the new node. This could explain why SNARC occurred in the direction that it did in all of my conditions but may also explain why SNARC was rarely detected when the semantic magnitude and absolute magnitude were not compatible.

As an example, if I consider their model with the pair (-8, -9) or (-9, -8) and the instruction “colder”, a right response will be facilitated by the magnitude field for numbers of relatively large absolute magnitude and the semantic magnitude field will also facilitate a right hand response because it is high on the “coldness” scale. Therefore right hand responses will be faster than left for cold pairs and left responses will be faster than right for warmer pairs. This would result in SNARC in the same direction I found for all of my colder-negative conditions.

If I consider the pair (1, 2) or (2, 1) and the instruction “colder”, the semantic field will activate the “right response” node because the pair is relatively cold, but the magnitude field will activate the “left response” node because it is relatively small. In the end, right responses will be faster than left because the semantic field has more influence.

At this point I don't know if this could a viable model or not, but I have made crude adjustments to their model in an attempt to make it fit with my data. Their model, as they describe it, is unable to account for the dominance of the semantic properties of the stimuli used in my experiments, over absolute value.

Seymour's Model

The second model I have to consider is the “response availability” model of Seymour (1973). Although he did not specifically make a SNARC model, the principles of his model can be applied to my experiments. Seymour’s model works on the premise that features of a task (e.g. instructions, dimensions, responses, etc.) are coded as (-) or (+) based on their position on an underlying dimension. The response time for a task is determined by the ease or difficulty of translating a stimulus display into a response.

In his experiment (Seymour, 1969) on each trial the word “above” or “below” appeared on the screen. Additionally there was a dot that appeared above or below the word. Participants were required to respond “yes” or “no” depending on whether the position of the dot relative to the word, matched the meaning of the word. Participants were faster to respond to the dot above the word “above” than they were to the dot below the word “below”.

Seymour argued that each prevalent feature of the trial represented a position on a particular dimension. The responses “yes” and “no” lay on the affirmation dimension, and the positions and words “above” and “below” lay on the verticality dimension. The idea is that the degree that the summed stimulus codes match the response polarity code will determine the response time. Therefore a series of only (+), or only (-), will result in the fastest response times but response times will slow to the degree that the codes are mixed. The polar codes are not referring to positive or negative numbers but rather are simply a means of binary coding like 1s and 0s.

When he repeated the experiment (1973) he told participants to reply “no” to matches and “yes” to mismatches. Now, the advantage for the “above/above” trials was gone and

there was instead an advantage for the “below/below” trials. Seymour argued that the word and position “above”, as well as the response “yes” were coded as (+) and the word and position “below” as well as the response “no” were coded as (-). When the word, position and response all match (+, +, +) his model correctly predicted that responses would be faster than any other condition, including that for a correct affirmation of “below/below” which would be coded as (-, -, +). However, using the negative response “no” to affirm the negatively coded “below/below” trials, would result in a compatible combination of codes (-, -, -) compared to the now mixed combination of codes for the above/above trials (+, +, -).

Semantic magnitude, defined by instruction, can correctly predict the direction of SNARC in every significant condition from my experiments. In an attempt to attain a set of codes that can be generalized to other experiments and make predictions in the future, I have applied Seymour’s concept. The simplest model is to simply code semantic magnitude as small (-) or large (+) and combine them with the codes for left (-) or right (+). I assume that semantic magnitude is considered to be small or large, relative to the semantic midpoint of other stimuli in the set used in a condition. For instance in the blocked condition, when the instruction is “colder” and the pair is (-1, -2), it will be considered relatively warm or in this case, semantically small, because the instruction is “colder”, and is therefore coded as (-). However, in the intermixed condition, the pair would be considered to be relatively cold compared to positive numbers and therefore coded as (+). Using codes for just the semantic magnitude generally makes correct predictions. For instance, in the blocked-colder-negative condition, if I consider the pair (-1,-2) or (-2, -1), I would code it as semantically small (-) and therefore left responses (-,

-) are predicted to be faster than right responses (-, +). If I consider a semantically large pair like (-9, -8), the model now predicts a faster right response (+, +) over a left response (+, -). For the intermixed condition, it is not as clear because now both pairs are semantically large and therefore a faster right response is predicted for both conditions. However, this is not what I found. But, the model also predicts faster right responses for semantically small and large pairs in the warmer-intermixed-positive condition, and this is consistent with what I found when the mean response times for semantically small and large numbers are averaged. In fact if I average the mean response times for relatively small and large pairs in every experiment, this model correctly predicts a response time advantage of one hand over the other 20 out of 24 times.

A more elaborate coding system that predicted a right or left-hand advantage 21 out of 24 times, included codes for absolute magnitude (small -, or large +), codes for whether the numbers were positive (+) or negative (-) in addition to codes for the semantic magnitude (small -, or large +). Included in Tables 2-5 are the predicted advantages for either the left or right hand, and the actual response time means I attained for each condition where a significant SNARC effect was found.

Table 2.

Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 3.

<i>Condition</i>	<i>Instruction</i>	<i>Numbers</i>	<i>ABS MAG</i>	<i>SEM MAG</i>	<i>POS NEG</i>	<i>Left Predict</i>	<i>Right Predict</i>	<i>Left</i>	<i>Right</i>
Blocked	Colder	-1,-2,-3,-4	-	-	-	- - - -	- - - +	566	570
Blocked	Colder	-6,-7,-8,9	+	+	-	+ + - -	+ + - +	638	608
Intermixed	Colder	-1,-2,-3,-4	-	+	-	- + - -	- + - +	972	987
Intermixed	Colder	-6,-7,-8,9	+	+	-	+ + - -	+ + - +	997	964
Blocked	Warmer	1,2,3,4	-	-	+	- - + -	- - + +	524	533
Blocked	Warmer	6,7,8,9	+	+	+	+ + + -	+ + + +	606	550
Intermixed	Warmer	1,2,3,4	-	+	+	- + + -	- + + +	772	750
Intermixed	Warmer	6,7,8,9	+	+	+	+ + + -	+ + + +	827	777

Table 3.

Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 4. The X shows an incorrect prediction.

<i>Condition</i>	<i>Instruct</i>	<i>Numbers</i>	<i>ABS MAG</i>	<i>SEM MAG</i>	<i>POS NEG</i>	<i>Left Predict</i>	<i>Right Predict</i>	<i>Left</i>	<i>Right</i>
Blocked	Colder	1,2,3,4	-	+	+	- + + -	- + + +	748	734
Blocked	Colder	6,7,8,9	+	-	+	+ - + -	+ - + +	837 X	838
Blocked	Colder	-1,-2,-3,-4	-	-	-	- - - -	- - - +	750	764
Blocked	Colder	-6,-7,-8,9	+	+	-	+ + - -	+ + - +	812	783
Blocked	Warmer	1,2,3,4	-	-	+	- - + -	- - + +	731	739
Blocked	Warmer	6,7,8,9	+	+	+	+ + + -	+ + + +	814	791
Blocked	Warmer	-1,-2,-3,-4	-	+	-	- + - -	- + - +	786	812
Blocked	Warmer	-6,-7,-8,9	+	-	-	+ - - -	+ - - +	880	910

Table 4.

Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 5. The Xs show predictions that were wrong.

<i>Condition</i>	<i>Instruct</i>	<i>Numbers</i>	<i>ABS MAG</i>	<i>SEM MAG</i>	<i>POS NEG</i>	<i>Left Predict</i>	<i>Right Predict</i>	<i>Left</i>	<i>Right</i>
Blocked	Colder	-1,-2,-3,-4	-	-	-	- - - -	- - - +	612	580 X
Blocked	Colder	-6,-7,-8,9	+	+	-	+ + - -	+ + - +	683	660
Blocked	Warmer	1,2,3,4	-	-	+	- - + -	- - + +	558	547 X
Blocked	Warmer	6,7,8,9	+	+	+	+ + + -	+ + + +	644	616

Table 5.

Presented are predictions for an advantage of left or right response time based on compatibility of codes. The polar codes are based on absolute magnitude (ABS), semantic magnitude (SEM), positive or negative pairs (POS NEG), and left and right responses. The midpoint is 5 or –5 in the blocked condition and 0 in the intermixed condition. The “numbers” column represents all the pairs made using from those adjacent numbers. Also included are the actual averages of mean response times in milliseconds, for small pairs and large pairs for conditions that generated a significant SNARC effect in Experiment 6.

<i>Condition</i>	<i>Instruct</i>	<i>Numbers</i>	<i>ABS MAG</i>	<i>SEM MAG</i>	<i>POS NEG</i>	<i>Left Predict</i>	<i>Right Predict</i>	<i>Left</i>	<i>Right</i>
Blocked	Slower	1,2,3,4	-	-	+	- - + -	- - + +	508	517
Blocked	Slower	6,7,8,9	+	+	+	+ + + -	+ + + +	597	566
Blocked	Slower	-1,-2,-3,-4	-	-	-	- - - -	- - - +	541	575
Blocked	Slower	-6,-7,-8,9	+	+	-	+ + - -	+ + - +	609	600

The application of Seymour’s model in this fashion is one way of interpreting the significant results I attained from my experiments. It is not clear from this model however, why I rarely attained SNARC in intermixed conditions or when the semantic magnitude and absolute magnitude were less compatible, such as when the pairs were positive and the instruction was “colder”.

One advantage this model has over Gevers's model is that it also accounts for the direction of semantic congruity effects, which importantly, seems to resolve the contradicting evidence regarding the mental number line. My experiments, as well as past experiments, often show conflicting results between the SNARC effect and the semantic congruity effect for the same data. For instance, the direction of SNARC for negative numbers in the blocked condition from Shaki and Petrusic (2005a) supports the idea of a mental number line that does not extend left past zero to include negative numbers, however, the semantic congruity effects do support this idea. If I apply a code for absolute magnitude (small -, or large +) and a code for semantic magnitude (small -, or large +), I can make a crude estimate that correctly accounts for all of the semantic congruity effect I found. For instance, if I consider the pair (-9, -8), when the instruction is "colder", the pair will be coded as large for absolute magnitude (+) and large for semantic magnitude (+). The compatibility of codes with the instruction "colder" gives it a response time advantage compared to when the instruction is "warmer" for the same pair, which would be coded as large (+) for absolute magnitude but small (-) for semantic magnitude. When a positive pair is considered, (8, 9) the codes are reversed. Responses will be faster for the instruction "warmer" coded as (+) for relatively large absolute magnitude and (+) for relatively large semantic magnitude, but when the instruction is "colder" the response time will be slower because it will be a mixed set of codes.

The application of Seymour's model to explain semantic congruity effects is in effect, the same as that presented by Banks, Clark and Lucy (1975) and Banks, Fujii and Kayra-Stuart and (1976), as described previously. The "semantic congruity" model describes two discrete stages of processing. The first is described as the "encoding stage"

where stimuli are coded according to their physical intensity or magnitude (akin to absolute magnitude). For instance, two relatively large numbers might be coded as L and L+. The second stage, the “comparison stage” is the locus of the semantic congruity effect. Here the coded stimuli must be matched to an appropriate response. If the codes match the instruction (e.g., Choose the “larger”) the “L+” stimulus is easily matched instead of “L” and a quick response is given. However, if the instruction does not match a coded stimulus (e.g., Choose the “smaller”) the stimuli must be recoded as S+ and S before the response is made. The additional processing takes time and therefore results in a semantic congruity effect. The comparison stage is akin to the semantic code in Seymour’s model, except that I incorporated the influence of instruction into the numbers as large or small on the semantic scale, rather than using a (+) or (-) code for the instruction itself. I could change Banks’s model such that there is a (+) code for when the instruction matches the stimulus or a (-) code when it does not.

At this point it seems as though the application of Seymour’s model, using the codes I have suggested may not be able to account for all SNARC experiments. One immediate problem with the model is that it is too easy for a researcher to decide what codes are important to be included based on the task, but as a starting point, this is necessary.

If I consider parity experiments, typically, responses for relatively small numbers (e.g. 1- 4) are faster with the left-hand compared to the right, but responses for relatively large numbers (e.g. 6- 9) are faster with the right hand compared to the left. If I consider the number 1 and apply a code for small (-) and for odd (-), the model correctly predicts a faster left-hand response compared to right. However, if the number 2 is coded as

“small” (-) and “even” (+), it predicts the same response time for the left and the right. If I consider all of the codes generated by the model as I have described it, I attain the following: 1 (-, -), 2 (-, +), 3 (-, -), 4 (-, +), 6 (+, +), 7 (+, -), 8 (+, +) and 9 (+, -). By comparing these codes with the codes for left (-) and right (+) responses, it becomes evident that the numbers 1, and 3 will have a left-hand advantage and the numbers 6 and 8 will have a right hand advantage and the rest of the numbers will show neither a left or right advantage. It is not impossible however that the differences in response time between left and right, for some but not all of the numbers, may be a powerful enough effect overall, to generate a downward SNARC slope across the numbers 1-9. This would need to be tested in the future.

Mental Number Line?

Many researchers have assumed that a mental number line is the origin for SNARC. Somehow, a mental representation of numbers with small numbers on the left and larger on the right, facilitates motor responses in the corresponding direction. A debate has emerged as to whether or not the mental number line includes negative numbers or if only absolute magnitude is represented.

My experiments demonstrate that negative numbers are indeed recognized according to their numerical value in a linear order. In the temperature experiments the direction of SNARC shifted between conditions with positive and negative pairs indicating the numerical value of pairs was recognized. If only the absolute value of pairs was recognized then the direction of SNARC would be the same for positive and negative numbers. Therefore the semantic value of the pairs, defined by instruction, including when they were negative was recognized and influenced SNARC.

However, I have also argued that the absolute magnitude of the pairs was a contributing factor to SNARC. Therefore, the notion that a single mental number line exists and is fixed with the negative numbers on the left or downward and positive numbers on the right or upward, is not supported by my experiments. The idea of a mental number line seems to be incompatible with the idea that there are two factors determining the direction of SNARC, the “least/most” representation and the order of absolute magnitude. It seems that this hypothesis would require two mental number lines, one with a full representation of numbers and one for absolute magnitude only, combating for dominance perhaps in a manner described by the model of Gevers et al (2006). This, however, is not the most parsimonious explanation of my results.

Seymour’s response availability model not only accounts for all of my results, including semantic congruity effects, just as well as a mental number line hypothesis, but is also consistent with the notion that both absolute magnitude and semantic magnitude can contribute to SNARC simultaneously. Therefore, it is my opinion that if a mental number line exists, it is probably not a contributing factor to the SNARC effect. Rather, SNARC has its origin in the degree of compatibility of dimensional codes between the absolute magnitude, semantic magnitude, polarity and response codes.

Although I was generally able to account for the direction of the SNARC effects I found, I was unable to account for all SNARC effects in a similar way. My application of Seymour’s model is a starting point to expand and refine the application of relevant codes appropriately, which may in fact be able to account for most or all of the previous SNARC results as well as make clear predictions for future studies. This is a goal for the future.

References

- Audley, R. J. & Wallis, C. P. (1964). Response instructions and the speed of relative judgments. I. Some experiments on brightness discrimination. *British Journal of Psychology*, 76, 387-404.
- Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus-response compatibility in representational space. *Neuropsychologia*, 36 (8), 731-735.
- Banks, W. P., Clark, H. H., & Lucy, P. (1975). The locus of the semantic congruity effect in comparative judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 104 (1), 626-641.
- Banks, W. P. & Flora, J. (1977). Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology: Human Perception and Performance*, 3 (2), 278-290.
- Banks, W., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance*, 2 (3), 435-447.
- Banks, W. P., Mermelstein, R., & Yu, K. H. (1982). Discriminations among perceptual and symbolic stimuli. *Memory & Cognition*, 10 (3), 265-278.
- Berch, D. B., Foley, E. J., Hill, R. J., and Ryan, P. M. (1999). Extracting parity and magnitude from Arabic numerals: Developmental changes in number processing and mental representation. *Journal of Experimental Child Psychology*, 74, 286-308.
- Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, 103, 1131-1136.

- Calabria, M. and Rossetti, Y. (2005). Interference between number processing and line bisection: a methodology. *Neuropsychologia*, 43, 779-783.
- Dehaene, S., and Akhavién, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 21 (2), 314-326.
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, 122, 371-396.
- Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 626-641.
- Ekstrand, B. R. (1967). The effect of sleep on memory. *Journal of Cognitive Neuroscience*, 75 (1), 64-72.
- Ellis, S. (1972). Interaction of encoding and retrieval in relative age judgments: An extension of the “crossover” effect. *Journal of Experimental Psychology*, 94, 291-294.
- Fischer, M. H. (2003a). Spatial representations in number processing – evidence from a pointing task. *Visual Cognition*, 10 (4), 493-508.
- Fischer, M. H. (2003b). Cognitive representation of negative numbers. *Psychological Science*, 14, 278-282.
- Fischer, M. H. & Rottmann, J. (2005). Do negative numbers have a place on the mental number line? *Psychology Science*, 47, 22-32.
- Fias, W., Lauwereyns, J., & Lammertyn, J. (2001). Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. *Cognitive Brain Research*, 12, 415-423.

- Ficca, G., Lombardo, P., Rossi, L. & Salzarulo, P. (2000). Morning recall of verbal material depends on prior sleep organization. *Behavioural Brain Research*, 112, 159-163
- Friend, K. E. (1973). Perceptual encoding in comparative judgments of race. *Memory & Cognition*, 1, 80-84.
- Galistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*, 4 (2), 59-65.
- Ganor-Stern, D., Tzelgov, J. (2008). Negative numbers are generated in the mind. *Experimental Psychology*, 55 (3), 157-163.
- Gevers, W., Reynvoet, B., and Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition*, 87, B87-B95.
- Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., and Fias, F. (2006). Numbers and space: A Computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, 32 (1), 32-44.
- Göbel, S., Walsh, V. and Rushworth, M. F. S. (2001). The mental number line and the human angular gyrus. *Neuroimage*, 14, 1278-1289.
- Henmon, V. A. C. (1906). The time of perception as a measure of differences in sensations. *Archives of Philosophy, Psychology, and Scientific Method*, 8, (1-75).
- Holyoak, K. J. & Walker, J. H. (1976). Subjective magnitude information in semantic orderings. *Journal of Verbal Learning and Verbal Behavior*, 15, (287-299).
- Holyoak, K. J. and Mah, W. A. (1981). Congruity at the instruction level. *Memory & Cognition*, 9, 197-204.

- Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, 6, 435-448.
- Ito, Y. & Hatta, T. (2004). Spatial structure of quantitative representation of numbers: Evidence from the SNARC effect. *Memory & Cognition*, 52 (4), 662-673.
- Jamieson, D. G., & Petrusic, W. M. (1975). Relational judgments with remembered stimuli. *Perception & Psychophysics*, 18 (6), 373-378.
- Kerst, S. M. & Howard, J. H. Jr. (1977). Mental comparisons for ordered information on abstract and concrete dimensions. *Memory & Cognition*, 5, 227-234.
- Keus, I. M. & Schwarz, S. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition*, 33 (4), 681-695.
- Keus, I. M., Jenks, K. M., & Schwarz, S. (2005). Psychophysiological evidence that the SNARC effect has its functional locus in a response selection stage. *Cognitive Brain Research*, 24, 48-56.
- Lorch, R. F. and Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 16, 149-157.
- Lucas, J. & Petrusic, W. M. (2006). SNARC for irrelevant digits during facial recognition. Canadian Society for Brain, Behaviour and Science 16th Annual Meeting.
- Maquet, P., Smith, C., & Stickgold, R. (Eds.) Sleep and Synaptic Plasticity. Oxford University Press: Oxford, 2003.
- Marschark, M. & Paivio, A. (1979). Semantic congruity and lexical marking in symbolic comparisons: An expectancy hypothesis. *Memory & Cognition*, 7 (3), 175-184.

- Marschark, M. & Paivio, A. (1981). Congruity and the perceptual comparison task. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 290-308.
- Marks, D. F. (1972). Relative judgment: A phenomenon and a theory. *Perception & Psychophysics*, 11 (2), 156-160.
- Moyer, R. S. & Landauer, T. K. (1967). Time required for judgements of numeral inequality. *Nature*, 215 (109), 1519-20.
- Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, 13, 180-184.
- Petrusic, W. M. & Baranski, J. V. (1989). Semantic congruity effects in perceptual comparisons. *Perception & Psychophysics*, 45 (4), 439-452.
- Pilhal, W. & Born, J. (1997). Effects and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience*, 9 (4), 534-547.
- Schwarz, W. and Keus, I. (2004). Moving the eyes along the mental number line: Comparing SNARC effects with saccadic and manual responses. *Perception & Psychophysics*, 66 (4), 651-664.
- Seron, X., Pesenti, M., Noël, M., Deloche, G. and Cornet, J. (1992). Images of numbers or “When 98 is upper left and 6 sky blue”*, *Cognition*, 44, 159-196.
- Seymour, P. H. K. (1969). Response latencies in judgements of spatial location. *British Journal of Psychology*, 60, 31-39
- Seymour, P. H. K. (1973). Judgments of verticality and response availability. *Bulletin of the Psychonomic Society*, 1 (3), 196-198.

- Shaki, S. and Petrusic, W. M. (2005a). On the mental representation of negative numbers: Context dependent SNARC effects with comparative judgments. *Psychonomic Bulletin & Review*, 12 (5), 931-937.
- Shaki, S. and Petrusic, W. M. (2005b). Extending the SNARC effect: Spatial components in the mental representations of numbers, symbols, and percepts. In J.S. Monahan, S. M. Sheffert & J. T. Townsend (Eds.) (2005). *Fechner Day 2005. Proceeding of the Twenty-first Annual Meeting of the International Society for Psychophysics*. Mt. Pleasant, MI: The International Society for Psychophysics.
- Shaki, S., Fischer, M. & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review*, 16 (2), 328-331.
- Shipley, Coffin & Hadsell (1945). Reaction time in judgment: A phenomenon and a theory. *Perception of Experimental Psychology*, 35, 206-215.
- Shipley Norris & Roberts (1946). The effect of changed polarity of set on decision time of affective judgments. *Journal of Experimental Psychology*, 69, 237-243.
- Stickgold, R. Whidbee, D., Schirmer, B., Patel, V., & Hobson, J. A. (2000). Visual discrimination task improvement: A multi-step process occurring during sleep. *Journal of Cognitive Neuroscience*, 12 (2), 246-254.
- van Garderen, D. & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research and Practice*, 18 (4), 246-254.
- Wagner, U., Gais, S., Haider, H., & Born, J. (2004). Sleep inspires insight. *Nature*, 427, 352-355.

- Wallis, C. P. & Audley, R. J. (1964). Response instructions and the speed of relative judgements. II. Pitch discrimination. *British Journal of Psychology*, 55 (2), 121-132.
- Zorzi, M., Priftis, K. & Umiltà, C. (2002). Neglect disrupts the mental number line. *Nature*, 417, 138-139.

Appendix A