Crossmodal congruency effect scores decrease with repeat test exposure

Daniel Blustein[1*], Satinder Gill[2*], Adam Wilson[3], Jon Sensinger[2,3]

[1]Department of Psychology, Neuroscience Program, Rhodes College, Memphis, TN, USA
[2]Institute of Biomedical Engineering, University of New Brunswick, Fredericton, NB, Canada
[3]Department of Electrical and Computer Engineering, University of New Brunswick, Fredericton, NB, Canada

*These authors contributed equally to this work.

Corresponding author**:**
Daniel Blustein[1]
Email: blusteind@rhodes.edu

Abstract

The incorporation of feedback into a person's body schema is well established. The crossmodal congruency task (CCT) is used to objectively quantify incorporation without being susceptible to experimenter biases. This visual-tactile interference task is used to calculate the crossmodal congruency effect (CCE) score as a difference in response time between incongruent and congruent trials. Here we show that this metric is susceptible to a learning effect that causes attenuation of the CCE score due to repeated task exposure sessions. We demonstrate that this learning effect is persistent, even after a 6 month hiatus in testing. Two mitigation strategies are proposed: 1. Only use CCE scores that are taken after learning has stabilized, or 2. Use a modified CCT protocol that decreases the task exposure time. We show that the modified and shortened CCT protocol, which may be required to meet time or logistical constraints in laboratory or clinical settings, reduced the impact of the learning effect on CCT results. Importantly, the CCE scores from the modified protocol were not significantly more variable than results obtained with the original protocol. This study highlights the importance of considering exposure time to the CCT when designing experiments and suggests two mitigation strategies to improve the utility of this psychophysical assessment.

Keywords: Crossmodal congruency; Learning effect; multisensory integration; reaction time

Introduction

There is an increasing interest concerning how the human brain represents the space surrounding its body due to converging findings from a number of different disciplines. The crossmodal congruency task (CCT) is a visual-tactile interference task that has been used to investigate multisensory representation of space in humans (Spence, Pavani & Driver, 1998; Spence, Pavani & Driver, 2004; Spence et al., 2004), including those with brain damage (Spence et al., 2001a). Investigation of crossmodal selective attention has been used to demonstrate the detrimental effects of age on the ability to ignore irrelevant sensory information when attending to relevant sensory information (Poliakoff et al., 2006). Other studies have investigated changes in the representation of peripersonal space that are elicited by the prolonged use of hand-held tools (Maravita et al., 2002; Holmes & Spence, 2004; Holmes, Calvert & Spence, 2007). Further studies used this task to extend findings regarding physical and pointing tools to virtual robotic tools using techniques from haptics and virtual reality (Sengül et al., 2012; Sengül et al., 2013). Recently it has been used in conjunction with the rubber hand illusion paradigm to investigate the degree of incorporation of a rubber hand into a person's body schema (i.e. how strongly the rubber hand is experienced as one's own hand) (Zopf, Savage & Williams, 2010; Zopf, Savage & Williams, 2013).

While the CCT assessment has become a well-established psychophysics method used to quantify the degree of feedback incorporation in a variety of contexts, in this study we focus on visuo-tactile integration, motivated by its application in the field of neuroprosthetics (Spence, 2015). In such investigations, a CCT typically consists of a participant holding two foam blocks in either hand with vibrotactile targets and visual distractors embedded in the top and bottom of each foam block. A trial consists of a random and independent presentation of a single vibrotactile target paired with a single visual distractor from any of the four possible pairs of locations. Participants are instructed to make a speeded response regarding the elevation of the vibrotactile target (i.e. "up", at the index finger verses "down", at the thumb), while simultaneously ignoring visual distractors (Spence et al., 2001a).

The CCT is a simple stereotypical behavioral task that provides a robust performance metric of feedback incorporation. Participants are typically slower at discriminating the elevation of vibrotactile targets when visual distractors are presented from an incongruent elevation (i.e. when vibrotactile target and visual distractors are presented from different elevations) as compared to a congruent elevation. The crossmodal congruency effect (CCE) score is calculated as a difference in the reaction time between incongruent and congruent trials and is used as a quantitative performance metric for multisensory representation of space. It has been demonstrated that CCE scores are typically higher when the spatial separation between visual distractor and vibrotactile

Deleted: T
Deleted: crossmodal congruency task (
Deleted:)
Deleted: . I
Deleted: ¶
Deleted: Typically
Moved (insertion) [1]

targets is low (e.g. both locations on same block vs. locations on two different blocks held in different hands) (Maravita, Spence & Driver, 2003; Spence, Pavani & Driver, 2004).

Moved up [1]: The CCT is a simple stereotypical behavioral task that provides a robust performance metric of feedback incorporation.

Despite the widespread use of the crossmodal congruency task in a variety of experimental paradigms, important questions related to task exposure times remain unanswered. For example, in a majority of the studies the experimental protocol consists of multiple blocks of trials with experimental sessions lasting up to approximately 60 minutes. These hour-long time-intensive experimental sessions might result in extended learning of the task and adversely affect a participant's CCE score. There is also no specific mention of participant selection criterion based on previous knowledge of the CCT. To our knowledge, there exist no studies to date that have investigated modulation of CCE score due to repeated task exposures.

Deleted: a

Deleted: crossmodal congruency task

Various neuropsychological and cognitive assessments have shown profound learning effects with repeat exposures, including various reaction time tasks (Collie et al., 2003), cognitive function tasks, such as the Paced Auditory Serial Addition Test (Beglinger et al., 2005), and other interference tasks, such as the Stroop Color and Word Test (Davidson, Zacks & Williams, 2003; Beglinger et al., 2005). If a task learning or practice effect exists for the CCT, results from previous studies that overlooked such an effect may be inaccurate. Participants tested with previous CCT experience would be expected to have lower CCE scores than participants naïve to the assessment.

Any study that presents more than one CCT exposure to a participant may lead to results affected by a learning effect. One previous study repeatedly ran the CCT with a single subject, Patient J.W. (Spence et al., 2001b). Spence et al. only directly compared performance within counterbalanced sessions in order to reduce the effect of differential motivation or fatigue (2001b). However, there is no discussion of a task learning effect, the results from five different sessions across two days are presented side-by-side (Spence et al., 2001b), and the same patient J.W.

appears in another CCT study (Spence et al., 2001a). Additionally, studies utilizing a counterbalanced within-subjects design may average out the learning effect, such as Holmes, Calvert & Spence (2007) in which each participant completed six blocks of 96 trials each across three different counterbalanced conditions. A learning effect, if present, would conflate the research findings that form the basis of our current understanding of multisensory space representation in humans. In the present study we show evidence that CCE scores decrease with repeated task exposure sessions and that this learning effect is persistent over time. We propose two mitigation strategies, including a shortened CCT protocol that we compare to the established protocol. Best practices for future CCE score use are suggested.

Materials and Methods

CCT implementation

During CCT testing, participants sat comfortably on a chair in front of a table in a dimly illuminated room. They wore over-ear noise-canceling headphones playing Brownian noise to mask background noise. Identical vibrotactile stimulation motors were placed on the thumb and index fingertips of the participant's right hand (see the Materials and Apparatus section for hardware details). Distractor light emitting diodes (LEDs) were placed on the thumb and index finger of the participant's right hand, while a fixation LED was centered between the distractor LEDs using a plastic mounting strip. Figure 1 shows the system setup used for this experiment.

At the beginning of the CCT assessment, participants were instructed to make a speeded response to vibrotactile targets that were presented randomly to either the thumb or index fingertip. Participants responded by pressing the left or right foot pedals to indicate stimulation of the thumb or index finger, respectively. Participants were explicitly instructed to respond as fast as possible, while making as few errors as possible. Participants were also informed that visual distractors would be presented simultaneously with the vibrotactile targets but that they were completely irrelevant to the vibrotactile target discrimination task. They were specifically instructed to ignore visual distractors by keeping their eyes open and fixating their vision on the central fixation LED. The fixation LED was presented 1000ms before the vibrotactile targets and visual distractors were simultaneously presented for 250ms.

The CCT test session was preceded by a practice session consisting of three blocks of 10 trials each. The first block of the practice session consisted of presenting the fixation LED and vibrotactile targets only (i.e. no visual distractors were presented) so that participants could familiarize themselves with the vibrotactile target discrimination. The visual distractors were presented along with vibrotactile targets in the following two blocks of the practice session. This was followed by the experimental session during which 64 trials were presented to each participant during each block of trials. Either four or eight blocks of trials were presented depending upon the particular experiment. Each trial ended when the participant responded by pressing one of the foot pedals or no response was made within 1500ms of target onset.

Participants

A total of 30 healthy participants were recruited from the local community (23 male, 7 female; aged 19 – 57 years, mean ± SD = 27.5 ± 7.8 years; 26 right hand dominant, 1 left hand dominant, 3 reported no particular hand dominance). All participants had normal or corrected to normal vision, no disorder of touch, were able to use both foot pedals, and were naïve to the CCT without previous exposure to or knowledge of the task. Participants were informed about the general purpose of the research and were given the opportunity to ask questions but they were not

Commented [LMH1]: Were they compensated in any way?

aware of the specific goals of the research. Experiments were conducted under human ethics approval of the U.S. Department of the Navy's Human Research Protection Program and the University of New Brunswick (Fredericton, NB, Canada) Research Ethics Board (Protocol #2016-032). Written informed consent was obtained from each participant before conducting experiments and no compensation was provided to participants.

Materials and apparatus

The experimental test platform was designed around the National Instruments (NI) myRIO embedded hardware system to achieve millisecond-timing accuracy. The three digital outputs of the myRIO system were used to drive one 3mm green LED, termed the fixation LED, and two 3mm green LEDs named distractor LEDs. Two 308-107 Pico Vibe 8mm vibratory motors from Precision Microdrives were used as the vibrotactile targets for the thumb and index finger. The analog outputs of the myRIO system were used to drive STMicroelectronics L272 power operational amplifiers that drove the vibratory motors. The amplifiers were necessary to provide sufficient driving current to the vibratory motors. Two OnStage KSP100 keyboard sustain pedals were interfaced to the myRIO system through digital inputs to measure a participant's speeded responses to the vibrotactile targets. NI LabVIEW was used to develop the firmware for the myRIO system to randomly activate the visual distractors and vibrotactile targets at a fixed interval and measure the speeded response time. A desktop NI LabVIEW graphical user interface (GUI) was designed to interact with the myRIO embedded system, allowing the experimenter to set various experimental parameters such as vibratory stimulus amplitude, the number of trials and to specify the filename to record the timing results of speeded responses.

Experiment design

Three experiments were run to examine CCT learning effects.

In Experiment 1 we assessed the learning effect over five sessions of the CCT. Twelve subjects each completed five consecutive days of CCT testing using the standard 8-block protocol (Spence, Pavani & Driver, 2004). The number of sessions and subjects was determined using a power analysis using preliminary CCT results from eight pilot subjects (see Appendix). Subjects were naïve to the CCT and testing was run at about the same time each day. Before each test session, generalized reaction time was measured using a ruler grasp task. Subjects held their thumb and first finger 4 cm apart with a vertically-aligned ruler held by the experimenter set so the bottom of the ruler was just within the grasp of the subject. The subject was asked to grasp the ruler as quickly as possible after they noticed it falling. The researcher dropped the ruler at an unspecified and variable time and recorded the distance fallen as a proxy for reaction time. The 'reaction distance' was calculated as the average result of three ruler grasp trials.

In Experiment 2, the persistence of the CCT learning effect was assessed using a within-subjects design comparing the responses of subjects tested in experiment 1 with their responses tested following the same protocol at a later time. Eight of the 12 subjects from Experiment 1 returned for an additional CCT test session after more than six months with no CCT exposure. For both Experiments 1 and 2, all CCT sessions consisted of eight blocks of 64 trials each. Each block lasted for approximately four minutes and a break period of two minutes was provided between consecutive blocks.

In Experiment 3, using a between-subjects design, we examined the effect of reducing the number of CCT blocks per session from eight to four. We expected fewer CCT trials to result in a reduced learning effect, and we sought to determine if CCE scores would maintain their

Commented [LMH2]: Here, please provide the demographic information for these subjects

Commented [LMH3]: What time of day was this?

Commented [LMH4]: Did you do this for experiments 2 and 3 as well? If so, please provide this information.

Commented [LMH5]: Here, please provide the demographic information for these subjects.

Commented [LMH6]: What was the average delay in testing between experiment 1 and experiment 2 for these eight subjects?

Commented [LMH7]: The description of this is still slightly unclear to me. Did one group of 9 subjects complete two blocks on one day and two blocks on a second day while the second group of 9 subjects completed four blocks on one day and four blocks on a second day? Please clarify this.

Additionally, please confirm the number of trials per block (I presume 64 as for experiments 1 and 2 but I think it bears repeating).

consistency with the shortened protocol. 18 subjects were tested across two sessions on consecutive days; half of the subjects (n=9) were randomly assigned to test with four blocks and the other nine subjects tested with eight blocks for each session.

Commented [LMH8]: Was testing run at the same time each day? What time was testing run?

Commented [LMH9]: Please provide the demographic information for each of the two groups of nine participants.

Analysis

Commented [LMH10]: Please provide information as to what software you used to run your analyses

Practice session trials were not analyzed. To calculate the CCE score from the test trials, we first discarded trials with an incorrect response, trials with a premature response (i.e. reaction time less than 200ms), and trials with a delayed response (i.e. reaction time greater than 1500ms) as these conditions most likely occurred due to lapses in attention (Spence, Pavani & Driver, 2004; Sengül et al., 2012). The remaining trials in each block were used to calculate the mean congruent and mean incongruent reaction times. The CCE score for each block was calculated by taking the difference between mean incongruent and congruent reaction times. The mean of the block CCE scores was used to calculate the CCE score for each participant for a particular exposure session. Selection error rates were used as a separate metric of analysis.

In Experiment 1, attenuation across sessions was evaluated using a one-way repeated-measures ANOVA. For Experiment 2, we ran an intraclass correlation coefficient analysis [two-way mixed model with single measurements] and calculated the standard error of measurement [SEM = SD√(1-ICC)] to assess the test-retest reliability between the 6+ month follow-up visit and the Day 5 visit (Weir, 2005). Experiment 3 data were analyzed using paired t-tests within the data collected using each protocol. All data are available on Dryad at doi:10.5061/dryad.150v8g3.

Commented [LMH11]: In some places you use the active voice (as here) whereas in other you use the passive voice (e.g., the previous sentence). I encourage you to be consistent (and I recommend the active voice).

Results

Experiment 1: CCE score decreases over repeated exposures

We measured the CCE score of 12 subjects across five sessions using the conventional CCT method of 8-blocks. We observed a significant effect of exposure number on CCE score determined with a repeated-measures ANOVA with Greenhouse-Geisser correction (Fig. 2) ($F(2.00,21.95)=6.93$, $p = .005$, partial-eta=.39). Tests of within-subjects polynomial contrasts over session number indicated a significant linear trend ($F(1,11) = 8.63$, $p=.013$), a significant 4th order trend ($F(1,11) = 5.14$, $p=.044$), and a quadratic trend ($F(1,11) = 4.83$, $p=.05$). Post-hoc comparisons with Bonferroni adjustments showed no significant pair-wise differences. The statistically-supported trends match the observed initial decrease in CCE score and subsequent stabilization, indicating a task learning effect (Fig. 2).

To verify that the change in CCE score was due to a learning effect, and not attributed to other interacting factors such as variability in motivation or baseline reactivity, we analyzed the CCT selection error rates, generalized reaction times measured with a ruler grasp task, and overall reaction times for congruent and incongruent stimuli. We found no significant effect of exposure number on CCT correct trial rate (Fig. 3A), as indicated by a repeated-measures ANOVA with Greenhouse-Geisser correction ($F(1.15,12.66)=0.53$, $p = .51$). The correct trial rates on congruent and incongruent trials independently did not show consistent trends (Fig. S1). Repeated-measures ANOVA with Greenhouse-Geisser correction indicated a significant effect of exposure on both incongruent reaction time ($F(2.16,23.78)=13.57$, $p <.01$) and congruent reaction time ($F(1.46,16.07)=8.89$, $p<.01$) (Fig. 3B).

To summarize, the only metrics presenting similar exposure-dependent decreases as the CCE scores (Fig. 2) were the congruent and incongruent reaction times (Fig. 3B). We found no significant effect of exposure number on generalized reaction time as measured by the fall-to-grasp distance of a ruler, as indicated by a repeated-measures ANOVA with Greenhouse-Geisser correction ($F(1.463,14.63)=1.269$, $p = .29$).

Experiment 2: CCT learning effect persists over time

After observing a statistically significant CCT learning effect in both incongruent and congruent reaction times, but not in error rates and generalized reaction times, we wanted to see if the learning effect persisted over longer time periods. Eight subjects who completed Experiment 1 were re-tested between 6 and 7 months after their initial testing. Re-visit CCE scores were similar, and even slightly lower, than the CCE scores measured on Session 5, about 6 months prior (Fig. 4). Similarity between Session 5 and the re-visit session was indicated by a high intraclass correlation coefficient (ICC(3,1) = .71) and a low standard error of measurement (SEM = 21.6ms) compared to baseline reliability data calculated by comparing Session 1 and re-visit results (ICC(3,1) = .46; SEM = 45.7ms).

Experiment 3: Modified CCT protocol reduces learning effect

We next sought to explore mitigation strategies to diminish the impact of the CCT learning effect on research results. The persistence of the learning effect suggests that CCE scores should only be considered after task learning has stabilized. However, in certain research or clinical settings with subject access or time constraints, extended testing may be impractical. In Experiment 3 we tested a modified CCT protocol designed to reduce task exposure to determine if the shortened testing could: 1. Produce valid CCT results; and 2. Mitigate the learning effect.

Commented [LMH12]: More accurate reporting required.

The modified protocol reduced the duration of task exposure in each session from eight blocks to four blocks. The 4-block protocol reduced the drop in CCE score from first exposure to second exposure (Fig. 5). For the 8-block protocol, mean CCE score dropped by 14.7ms from the first exposure to the second, compared to a 7.6ms drop for the 4-block protocol. Variability in the second session was only slightly higher for the 4-block protocol (SD = 43.8ms) compared to the 8-block protocol (SD = 39.0ms). The variability difference was more pronounced when analyzing the first session results [4-block SD =64.9ms; 8-block SD = 47.5ms]. None of these differences were statistically significant (paired t-tests, $p > .05$), however.

Discussion

The CCT is a well-established quantitative measure of feedback incorporation that has important implications for objective human sensorimotor assessments (Spence, Pavani & Driver, 1998; Spence, Pavani & Driver, 2004; Spence et al., 2004; Spence, 2015). The resulting scores have been used to quantify the rubber hand illusion (Zopf, Savage & Williams, 2010) and show promise for assessing neuroprosthetic devices (Spence, 2015). Many applications of the CCT require repeated testing – for example, testing the same subject under different sensory feedback conditions to see which one enables the highest degree of feedback incorporation. Oftentimes in research or clinical settings time allotted for sensorimotor assessment can be limited. This study sought to characterize CCT performance through repeat testing to inform its future implementation.

In Experiment 1, we found that with repeat testing, CCE scores decreased over the initial exposure sessions. The results are not explained by a fatigue effect as the CCE score trend was not matched by the overall reaction times measured with a ruler grasp task. Different motivation levels

Commented [LMH13]: This information should be moved to your methods section where you describe experiment (i.e. your rationale for the experiment and your predictions) rather than here in the Results section.

Commented [LMH14]: As you do when describing the 8-block protocol, I think it would be helpful to provide the mean CCE scores here.

Commented [LMH15]: This sentence is unclear. Do you mean:

"In the 4-block protocol, the CCE score dropped from first exposure to second exposure"

Or

"The 4-block protocol reduced the drop in CCE score from first exposure to second exposure as compared to the 8-block protocol" or strategies fail to explain the CCE score trend as CCT error rates did not show a trend matching the decreasing scores. Thus, we consider the decrease in CCE score to be evidence of a task learning effect. Interestingly, we observed no significant pair-wise differences between CCE scores on different days, even though our statistical tests showed overall significant effects. This was likely due to the large number of pairwise comparisons and the conservative Bonferroni correction we applied.

Although other psychophysical paradigms similarly show significant practice effects (Collie et al., 2003; Davidson, Zacks & Williams, 2003; Beglinger et al., 2005), for the case of the CCT, scores tend to increase with increased training durations (Blustein, Wilson & Sensinger 2018). The observation of a learning effect in light of these counteracting processes suggests that either an increase in CCE score does not occur with repeat assessment (rather than generalized training), or the learning effect has a much greater impact in the resulting CCE score.

In Experiment 2, we demonstrated that the learning effect is persistent as CCE scores after six months remained near the stabilized scores observed at the end of five days of consecutive testing. The persistence of the learning effect provides an opportunity for researchers to only use CCE scores collected after learning has stabilized.

In certain circumstances, additional testing to arrive at stable CCE scores may not be feasible. In Experiment 3, we showed that a shortened CCT protocol with four testing blocks instead of eight reduced the impact of the learning effect across two testing sessions. Experiment 3 was underpowered statistically and thus these findings can only be considered preliminary trends. In certain circumstances where repeat testing is required but time constraints exclude the possibility of testing until learning asymptotes, the reduced exposure protocol may be necessary.

Importantly, CCE score variability did not substantially increase with the 4-block protocol, although the statistical caveats apply here as well.

Reducing the number of blocks per session has the added benefit of reducing fatigue, improving the ability of subjects to concentrate over the duration of the study, and shortening the length of the session to approximately 30 minutes (which in turn makes it viable to test a greater number of subjects or for use in a clinical setting). Using a modified version of the CCT with only four blocks of 64 trials has many potential benefits, with the only drawback identified being a slight increase in the variability of results.

Through this characterization of the learning effect associated with the CCT, we have provided support for two distinct mitigation strategies. If extended testing is feasible, researchers can provide enough practice so that CCE scores are only measured and compared after task learning has stabilized. This would be of particular use when tracking an individual's progress over time with repeat testing. In situations where reduced testing is required, researchers can use the 4-block CCT protocol to reduce overall task exposure and subsequently reduce the impact of the learning effect on the results. It is important that these strategies not be mixed; CCE scores should only be compared to other scores that were collected in the same manner. Randomizing condition order to provide counterbalanced groups can also serve to reduce the learning effect's impact on results and can be useful in observing group differences (Holmes, Calvert & Spence, 2007). However, this counterbalancing approach will effectively average out the learning effect, which would lead to reduced magnitudes of any observed differences between conditions.

It is not obvious what is being 'learned' through repeated task exposure. We suspect that subjects strengthen an internal model of the foot pedal location tied to the vibration stimulus, resulting in faster response times. Anecdotal support for this idea comes from one subject's comment: "When the feedback is in the opposite location to the light, I sometimes find myself thinking about which pedal to press only to realize I've already pressed the correct pedal". Somatosensory reorganization of an individual's body schema through training may also be implicated (Cardinali et al., 2009).

The initial session CCE scores (i.e. before learning effect stabilization) appear to be higher than scores reported by other researchers (Spence, Pavani & Driver, 2004; Zopf, Savage & Williams, 2010). One potential explanation for this observation is that other studies may have different baseline CCE scores due to the experimental inquiry, such as studies using spatial misalignment (Spence, Pavani & Driver, 2004) or artificial hands (Zopf, Savage & Williams, 2010; Marini et al., 2014). We would expect a lower degree of feedback incorporation, and thus a lower CCE score (Maravita et al., 2002), under these conditions compared to the ideal conditions we used with both vibratory feedback and visual distractors aligned in place on one of the subject's hands. Additionally, different experimental set-ups could account for differences in CCE scores. Others have recorded user inputs differently, for example with a rocking heel-toe foot pedal setup (Spence, Pavani & Driver, 2004) or with finger presses (Zopf, Savage & Williams, 2010). Other variable techniques include the exclusion of trials with extensive eye movement [over 9% omitted in one study (Spence, Pavani & Driver, 2004)], or the use of no-go trials as a control (Zopf, Savage & Williams, 2010). Furthermore, differences in the duration and parameters of practice or familiarization trials before testing could lead to differences in the resulting CCE scores. In a hypothetical experimental setup with all other factors held constant, this study would suggest that a longer practice phase could result in a more pronounced learning effect before testing began, resulting in a lower CCE score. It might also be helpful to explore more granular trial-to-trial variation in the CCT results using mixed effects modeling.

Deleted: e.g

Another possible explanation for the observed variability in CCE scores across studies is the different methods used to apply feedback. Different feedback modalities result in different CCE scores (Mayer et al., 2009; Frings & Spence, 2010). Even when comparing studies using vibration feedback, the method of application can vary and may result in CCE score differences. We used small vibratory motors (0.8cm diameter vibrating surface) but others have used larger bone conduction vibrators [1.6cm x 2.4cm vibrating surface, (Spence, Pavani & Driver, 2004)] or small speakers [0.9cm diameter vibrating surface, (Zopf, Savage & Williams, 2010)], all could result in different degrees of feedback incorporation. To focus on the learning effect, we kept all testing parameters in this study constant. Although we expect these results to be similar across other CCT implementations, the magnitude of the learning effect under different CCT parameters is unknown.

Differences in subject characteristics could also explain differences in CCE scores. All of the subjects in this study were initially naïve to the CCT, but it is unclear if that was an inclusion criteria in other studies. Even comparing the first session results from this study across the two cohorts of subjects tested with 8 CCT test blocks we see differences in mean CCE score. In Experiment 1, a mean CCE score of 140ms was observed (Fig. 2) compared to 115ms in Experiment 3 (Fig. 5). More careful monitoring of additional potentially confounding variables may be helpful. The ruler-drop task we used to measure reaction time is inherently variable and a more precise reaction time measurement method may be helpful to better track day-to-day within-subject reactivity baseline differences. To summarize, there are lots of factors that could affect CCE score and it is impossible to determine if results reported elsewhere are capturing pre-, mid- or post-learning effect scores.

To our knowledge, this is the first study that has demonstrated the modulation of CCE score due to task overexposure. We recognize some limitations with this study that could warrant additional studies. Although we collected a generalized reaction time metric with a ruler grasp task, more sophisticated tracking of mental and physical fatigue would be helpful to better inform our understanding of the observed learning effect. Although CCT testing for individual participants was conducted at the same time of the day for each exposure, no strict guidelines were used to ensure that participants were in similar mental or physical states before each test session. Nevertheless, the learning effect we have described has important implications for future use of the CCT. Going forward, the learning effect must be considered when measuring CCE scores by only using scores from after learning has stabilized, or by using a modified protocol that reduces exposure to the task.

Conclusion

This study identified a learning effect with an established psychophysics metric: repeat exposure to the crossmodal congruency effect task resulted in reduced scores. The effect persisted during follow-up testing after a six-month hiatus. This learning effect must be considered when using the CCT. For limited testing, we have presented a modified protocol with reduced trial blocks that can be used to limit task exposure. For ongoing testing where multiple tests are necessary, we suggest researchers allow this learning effect stabilize before relying on CCE scores for comparisons. Consideration of the learning effect is important to properly contextualize CCT results.

References

Beglinger, L.J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D.A., Crawford, Fastenau, P.S., & Siemers, E.R. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing, *Archives of Clinical Neuropsychology*, *20*(4), 517–529. doi:10.1016/j.acn.2004.12.003

Blustein, D., Wilson, A., & Sensinger, J. (2018). Assessing the quality of supplementary sensory feedback using the crossmodal congruency task. *Scientific Reports*, *8*, 6203. doi:10.1038/s41598-018-24560-3

Cardinali, L., Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A.C., & Farnè, A. (2009). Tool-use induces morphological updating of the body schema. *Current Biology, 19*(12), R478–9. doi:10.1016/j.cub.2009.05.009

Collie, A., Maruff, P., Darby, D., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. *Journal of the International Neuropsychological Society, 9*(3), 419-428. doi:10.1017/S1355617703930074

Davidson, D.J., Zacks, R.T., & Williams, C.C. (2003). Stroop interference, practice, and aging. *Neuropsychology, development, and cognition. Section B: Aging, neuropsychology and cognition*, 10(2), 85-98. doi:10.1076/anec.10.2.85.14463

Frings, C., & Spence, C. (2010). Crossmodal congruency effects based on stimulus identity. *Brain Research, 1354*, 113–22. doi:10.1016/j.brainres.2010.07.058

Holmes, N.P., Calvert, G.A., & Spence, C. (2007). Tool use changes multisensory interactions in seconds: evidence from the crossmodal congruency task. *Experimental Brain Research*, *183*(4), 465–76. doi:10.1007/s00221-007-1060-7

Holmes, N.P., & Spence, C. (2004). The body schema and the multisensory representation(s) of peripersonal space. *Cognitive Processing, 5*(2), 94–105. doi:10.1007/s10339-004-0013-3

Maravita, A., Spence, C., & Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Current Biology, 13*(13), R531–9. doi:10.1016/S0960-9822(03)00449-4

Maravita, A., Spence, C., Kennett, S., & Driver, J. (2002). Tool-use changes multimodal spatial interactions between vision and touch in normal humans. *Cognition, 83*(2), B25–34. doi:10.1016/S0010-0277(02)00003-3

Marini, F., Tagliabue, C.F., Sposito, A.V., Hernandez-Arieta, A., Brugger, P., Estévez, N., Maravita, A. (2014). Crossmodal representation of a functional robotic hand arises after extensive training in healthy participants. *Neuropsychologia, 53*, 178–86. doi:10.1016/j.neuropsychologia.2013.11.017

Mayer, A.R., Franco, A.R., Canive, J., & Harrington, D.L.. (2009). The effects of stimulus modality and frequency of stimulus presentation on cross-modal distraction. *Cerebral Cortex, 19*(5), 993–1007. doi:10.1093/cercor/bhn148

Poliakoff, E., Ashworth, S., Lowe, C., & Spence, C. (2006). Vision and touch in ageing: crossmodal selective attention and visuotactile spatial interactions. *Neuropsychologia, 44*(4), 507–17. doi:10.1016/j.neuropsychologia.2005.07.004

Sengül, A., Rognini, G., van Elk, M., Aspell, J.E., Bleuler, H., Blanke, O. (2013). Force feedback facilitates multisensory integration during robotic tool use. *Experimental Brain Research, 227*(4), 497–507. doi:10.1007/s00221-013-3526-0

Sengül, A., van Elk, M., Rognini, G., Aspell, J.E., Bleuler, H., & Blanke, O. (2012). Extending the body to virtual tools using a robotic surgical interface: evidence from the crossmodal congruency task. *PLoS ONE, 7*(12), e49473. doi:10.1371/journal.pone.0049473

Spence, C. (2015). The cognitive neuroscience of incorporation: body image adjustment and neuroprosthetics. In: K. Kansaku, L. Cohen, & N. Birbaumer (Eds.), *Clinical Systems Neuroscience* (pp. 151–68). Tokyo: Springer. doi:10.1007/978-4-431-55037-2_9

Spence, C., Kingstone, A., Shore, D.I., & Gazzaniga, M.S. (2001a). Representation of visuotactile space in the split brain. *Psychological Science, 12*(1), 90–3. doi:10.1111/1467-9280.00316

Spence, C., Shore, D.I., Gazzaniga, M.S., Soto-Faraco, S. & Kingstone, A. (2001b). Failure to remap visuotactile space across the midline in the split-brain. *Canadian Journal of Experimental Psychology,* 55(2), 133-40. doi:10.1037/h0087360

Spence, C., Pavani, F., & Driver J. (1998). What crossing the hand can reveal about crossmodal links in spatial attention. *Abstracts of the Psychonomic Society*, 3, 13.

Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience, 4*(2), 148–69. doi:10.3758/CABN.4.2.148

Spence, C., Pavani, F., Maravita, A., & Holmes, N. (2004). Multisensory contributions to the 3-

D representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *Journal of Physiology - Paris, 98*(1-3), 171–89. doi:10.1016/j.jphysparis.2004.03.008

Weir, J.P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength & Conditioning Research, 19*(1), 231–40. doi:10.1519/15184.1

Zopf, R., Savage, G., & Williams, M.A. (2010). Crossmodal congruency measures of lateral distance effects on the rubber hand illusion. *Neuropsychologia, 48*(3), 713–25. doi:10.1016/j.neuropsychologia.2009.10.028

Zopf, R., Savage, G., & Williams, M.A. (2013). The crossmodal congruency task as a means to obtain an objective behavioral measure in the rubber hand illusion paradigm. *Journal of Visualized Experiments*, 77:e50530. doi:10.3791/50530