Enhancing implicit change detection through action

Philip Tseng¶, Jan Tuennermann§, Nancy Roker-Knight, Dorina Winter#, Ingrid Scharlau§, Bruce Bridgeman
Psychology Department, University of California, Santa Cruz, CA 95064, USA; e-mail: tsengphilip@gmail.com (¶ also National Central University, Jhongli, Taiwan); § Universität Paderborn, Paderborn, Germany; # Georg-August-Universität Göttingen, Göttingen, Germany
Received 17 April 2010, in revised form 9 September 2010

Abstract. Implicit change detection demonstrates how the visual system can benefit from stored information that is not immediately available to conscious awareness. We investigated the role of motor action in this context. In the first two experiments, using a one-shot implicit change-detection paradigm, participants responded to unperceived changes either with an action (jabbing the screen at the guessed location of a change) or with words (verbal report), and sat either 60 cm or 300 cm (with a laser pointer) away from the display. Our observers guessed the locations of changes at a reachable distance better with an action than with a verbal judgment. At 300 cm, beyond reach, the motor advantage disappeared. In experiment 3, this advantage was also unavailable when participants sat at a reachable distance but responded with hand-held laser pointers near their bodies. We conclude that a motor system specialized for real-time visually guided behavior has access to additional visual information. Importantly, this system is not activated by merely executing an action (experiment 2) or presenting stimuli in one’s near space (experiment 3). It is activated only when both conditions are fulfilled, which implies that it is the actual contact that matters to the visual system.

1 Introduction
Change blindness demonstrates the impoverished nature of visual representations. Large changes, even at the center of view, often go unnoticed for an extended period of time if they are masked by disruptions to visual continuity (Becker and Pashler 2002; Levin and Simons 1997; Pashler 1988; Rensink et al 1997; Simons and Levin 1998; Simon and Rensik 2005). The robustness of this effect provides a strong argument for a poor visual representation, as most changes that do not affect the gist of a scene are hard to detect (Sampanes et al 2008). However, what people cannot consciously perceive sometimes can alter subsequent decisions without alerting awareness (Greenwald et al 2003). For example, people can perform visual searches faster among a previously-seen pattern of distractors although the distractor pattern cannot be recognized consciously (Chun and Jiang 1998). Other findings in the investigation of subliminal perception and inattentional blindness also support the idea that some usable information may not be picked up by explicit reports (Merikle et al 2001; Moore and Egeth 1997).

Following this logic, several researchers have looked at whether change detection can be assessed with implicit measures that do not rely on conscious report. Fernandez-Duque and Thornton (2000), for example, observed a guessing rate that was better than chance even when subjects were not aware of the changes. Henderson and Hollingworth (1999) also reported that when changes are not detected, people’s eye gazes linger around the areas of change longer as if they knew where to look for the changes. Recent neuroscientific studies utilizing event-related brain potentials were also able to identify electrophysiological markers that are exclusive to change-present trials and not change-absent trials when conscious change detection fails (Fernandez-Duque et al 2003; Kimura et al 2008). Together, these studies demonstrate that implicit change
detection can be a fruitful approach to studying visual representation (Fernandez-Duque and Thornton 2003; Laloyaux et al 2003; see Thornton and Fernandez-Duque 2002, for a review; but see Mitroff and Simons 2002 and Mitroff et al 2002 for an alternative view on how implicit change detection can also be explained by explicit mechanisms).

1.1 The role of action in change detection
In the present study, we propose a new way to assess visual representations by incorporating motor responses beyond the conventional verbal response. In other words, we ask whether actions, or motor responses, may have an effect on people's ability to detect changes implicitly. This research question is based on the qualitative differences between the two visual systems (Goodale and Milner 1992; Milner and Goodale 1995; Ungerleider and Mishkin 1982), which we refer to as the cognitive (ventral) and the sensorimotor (dorsal) system (Bridgeman et al 1997).

The effect of the sensorimotor system on visual perception has been reported in many different paradigms. For example, the Ebbinghaus illusion, where a disc appears to be smaller when surrounded by bigger peripheral discs (and vice versa), is stronger in verbal reports and diminishes by approximately 40% when actions (eg finger apertures) are used in place of verbal report (Aglioti et al 1995; Franz et al 2000, 2001; Vishton et al 2007). In a similar vein, the induced Roelofs effect—where the subjective midline is biased by an off-centered frame—can be observed only in cognitive reports (ie verbal report or multiple choice) and not when participants jabbed the stimulus display with fingers (Dassonville et al 2004). Therefore, accurate spatial action was possible despite illusions in the cognitive system. Together, these findings indicate the importance of actions in visual perception and representation, and it is reasonable to expect motor responses to affect people's ability to implicitly detect changes. In the context of implicit change detection, the closest measure to incorporating action was using a mouse to respond on a computer screen (Fernandez-Duque et al 2003; Mitroff et al 2002). However, since using a mouse relies on a completely different kind of representation (eg moving the cursor up requires pushing the mouse forward, and the location of action is not the location of its result), it is arguably a cognitive response that is expressed via motor actions, which is quite different from direct action initiated by the sensorimotor system.

Here, we investigated the role of action in implicit change detection by combining a one-shot change-detection paradigm with an action response—pointing and actually touching the screen that contains the stimulus image. To anticipate our results, in experiment 1 we observed a sensorimotor advantage over verbal report when participants jabbed the stimulus display with their fingers. In experiments 2 and 3 we investigated whether this sensorimotor advantage originated from the pointing action or hand–stimulus contact by varying the distances between the observers and the display so that the pointing motion did not always reach the stimulus display.

2 Experiment 1
Experiment 1 was designed to test whether there is a difference between the cognitive and sensorimotor systems in implicit detection performance. As mentioned before, the two systems react differently in terms of perception (Milner and Goodale 1995) and remembering spatial locations (Bridgeman et al 1997; Dassonville et al 2004). Since the sensorimotor system does not rely on conscious experience as much, if there exists a qualitative difference between the two systems, we expect the sensorimotor system to outperform the cognitive one in the context of implicit change detection.
2.1 Methods

2.1.1 Participants. Eighty-four students participated in the study in fulfillment of course requirement (twenty-one male, sixty-three female). Four (one male, three female) were excluded from the study owing to their tendencies to change responses. All had normal or corrected-to-normal vision. Participants were randomly divided into either the cognitive or the sensorimotor group. To prevent any learning effect, whether conscious or unconscious, that may alter participants’ implicit change detection performance, in this study we used a between-subjects design. Therefore, each participant only participated in the verbal report or in the pointing action group, but never in both. This design also minimized the possibility of participants being able to develop any search strategies based on their previous exposure to the pictures.

2.1.2 Materials. The stimuli consisted of 30 sets of natural-scene pictures. Each set contained an original photo (picture A) and its altered version (picture $A'$). For recording purposes, each picture was divided into six grids of equal size that were organized in three columns and two rows. On the response screen, participants were to respond either with a spoken number or a pointing motion, depending on their group assignment. For the verbal report group, they were shown a six-grid frame with a number (e.g., 1–6) in each grid immediately following picture A. In the action condition, only a photo frame was shown (see figure 1). All responses were recorded by blind raters. All changes in the photos were evenly distributed in the six areas of the photos so that there were five changes in total for each grid area. All changes were carefully made so that they were all well enclosed within each grid and at least 1 deg away from the grid border. Participants sat 60 cm away from a 17-inch CRT monitor. The entire display was 24 deg wide.

Since the goal of the study was to achieve implicit detection, we conducted several pilot studies in order to exclude picture pairs that contained changes that were easy to detect and replaced them with harder ones. The final 30 picture pairs used in this experiment were therefore the ‘elite’ pictures containing changes that were never detected by any of our fifteen pilot participants. This was done to ensure minimal conscious detection of changes. In addition, this approach also ensured that the degrees of difficulty of detecting changes between (pictures) and within (grids) each picture exceeded a difficulty baseline.

2.1.3 Procedure. Participants pressed a key to start each trial. There was no fixation target. Instead, participants were instructed to look at the entire image rather than just fixating at a few corners. Once the trial started, picture A was presented for 500 ms, followed by a 100 ms blank screen, and picture $A'$ presented for 500 ms. A response screen was then displayed for 1000 ms, where participants were to respond either by word or by hand, depending on their assigned group. In the action condition, participants were to jab the location of the change (they were given a pen if their index fingers could not reach the display). They were also told not to speak during the test, unless they consciously detected any change. This was done to prevent any ‘bleeding’ of cognitive influences via verbal report.

All participants were instructed to respond as quickly as possible while the response screen was still up, and choose a location even if they did not see any change. In addition, participants were told to guess freely and quickly, and trust their ‘gut feelings’ if they could. They were also told to report any conscious detections of change (3%). There were 4 practice trials before the experiment began. No feedback regarding accuracy was given during practice or the experiment.
2.2 Results

Data from the eighty participants were submitted to an independent-sample $t$-test. Overall, participants from the sensorimotor condition performed better than their counterparts from the cognitive condition. That is, they averaged more correct guesses or hits (26% versus 22%). This difference between the sensorimotor and cognitive condition was significant in an independent sample $t$-test ($t_{78} = 2.829, \ p = 0.006$). In addition, separate one-sample $t$-tests also revealed that hit rates from both groups

---

**Figure 1.** [In colour online, see http://dx.doi.org/10.1068/p6711] Picture A was presented for 500 ms, followed by a 100 ms grey screen, and a 500 ms picture A'. A response screen then prompted participants to respond within 1 s. Participants in the motor response group viewed a grey response frame and were instructed to point to the location of change. Participants in the verbal report group saw six grids on the response screen and were instructed to verbally report the number that corresponds to the location of change.
were significantly above chance level of 16.7% (sensorimotor: $t_{39} = 6.271, p < 0.001$; cognitive: $t_{39} = 3.111, p = 0.003$). This implies that, although the sensorimotor visual system outperformed the cognitive in implicit detection, both systems were able to access the implicit visual representation to some extent and yield a hit rate significantly above chance. These results support the growing literature suggesting the existence of an implicit visual representation that is often overlooked by the standard change-blindness paradigm (Fernandez-Duque and Thornton 2000).

In order to make sure our observed pattern was not driven by hit rates from a particular region (out of the six grids) in our stimuli, we also submitted the hit rates of each image (30 in total) into a one-way ANOVA with an independent variable of ‘region of change’ (six levels: region 1 to 6). The result was not significant in the one-way ANOVA ($F_{5,24} = 0.622, p = 0.684$). Therefore, we did not observe a systematic difference between hit rates from each region.

2.3 Discussion

Our data show that action responses provided something unique over and above cognitive verbal reports. By changing the way people responded, the implicit detection paradigm picked up the qualitative differences between the systems. This has two implications for the current change-detection literature. First, participants in the cognitive group, although outperformed by their counterparts in the sensorimotor group, still guessed above chance level. This extends the evidence of previous findings on implicit change detection (Fernandez-Duque and Thornton 2000) by confirming the possibility of implicit detection occurring in natural scenes. Previous studies on implicit detection have mostly relied on simple visual-search displays with singletons of various orientations. Since the means from both cognitive and sensorimotor conditions were significantly above chance level of 6 hits, the present experiment suggests that implicit detection in natural scenes is possible, regardless of the mode of response. Since most of the change-blindness studies up to date have exclusively relied on verbal response, our conclusion regarding an impoverished visual representation may need to be reexamined. Second, the main finding from the current experiment revealed that people could guess better with actions than words even when conscious change detection failed. This is a novel manipulation that has not been incorporated into an implicit change-detection paradigm before. In experiments 2 and 3 we further investigate the origin of such sensorimotor advantage.

One word of caution regarding our data is that, although hit rates in the sensorimotor condition were significantly higher than in the cognitive condition, the difference was approximately 4%, which corresponds to 1.2 more hits out of 30 trials. Therefore, size of the effect was rather small. Note that in a flicker paradigm participants need to successfully attend to the correct locations both during picture A and A′ for change detection to succeed (O’Regan et al 2000). Thus, although we encouraged our participants to try to cover the entire image rather than focusing on a specific location alone, the actual number of pictures that were properly encoded and compared was small. Even so, we proceeded with the one-shot paradigm to avoid a type-I error. Therefore, it is possible that the effect of action affordance may be greater than our data suggest.

Since experiment 1 demonstrated a sensorimotor advantage in implicit detection, the next logical question becomes when, or under what condition, this advantage occurs. Milner and Goodale (1995) proposed that our sensorimotor system deals with action, and therefore should mostly be tuned to information within our arms’ reach. This idea is similar to the concept of affordance, proposed by J J Gibson several decades ago (1979), which encouraged psychologists to look at objects by the possible actions which they support. If this is true, the observed sensorimotor advantage should diminish when the visual scene no longer affords possibility for actions. Experiments 2 and 3 were designed to answer this question.
3 Experiment 2
In this experiment we stretched the distance between our participants and the display from 60 cm to 300 cm while keeping the visual angle constant. If the possibility to exert action was responsible for the observed sensorimotor advantage in experiment 1, this advantage should disappear with stretched distance. Alternatively, if the sensorimotor advantage came from the actual execution of the pointing action, then the same sensorimotor advantage should persist.

3.1 Methods
3.1.1 Participants. Eighty students participated in the study in fulfillment of course requirement (twenty-eight male, fifty-two female). All had normal or corrected-to-normal vision. Participants were randomly divided into either the cognitive or the sensorimotor condition.

3.1.2 Materials. The pictures shown were the same as those used in experiment 1. To accommodate the longer distance, we used a projector to display the images. The visual angle was kept the same as in experiment 1.

3.1.3 Procedure. Procedures were identical to those used in experiment 1 except for two differences: (a) participants sat 300 cm away from the screen, and (b) participants in the sensorimotor condition used a laser pointer to indicate their guesses for the location of the change. They were told to stretch their arms while using the laser pointer as if they were to jab directly at the screen.

3.2 Results
Overall, participants had a hit rate of 23% and 24% (see figure 2) in the sensorimotor and cognitive conditions, respectively. An independent-sample t-test did not reveal any significant difference between the two modes of response ($t_{78} = 0.77, p = 0.44$). Therefore, when the distance between the observer and the stimuli is stretched, the advantage of a sensorimotor response over a cognitive one is now eliminated. Note that we did not collapse the current data with those from experiment 1 for comparison, because participants were not randomly assigned to the four conditions across the two experiments.\(^1\)

![Figure 2. Hit rates from all three experiments. Only the difference in experiment 1 was statistically significant. The dashed line represents chance level performance. The error bars represent ±1 SEM.](image)

\(^1\)Although inappropriate for the reasons given above, the data from the first two experiments were directly compared in a 2 × 2 ANOVA with two between-subjects variables of mode of response (verbal report versus jabbing motion) and distance (60 cm versus 300 cm). The analysis revealed a significant interaction between the two factors ($F_{1,78} = 7.684, p < 0.01$). Separate t-tests revealed that performance from the motor action and verbal report groups were significantly different in the short-distance condition (experiment 1, $p = 0.002$), but not in the long-distance condition (experiment 2, $p = 0.444$), which is consistent with what we report above.
3.3 Discussion
The results showed that the sensorimotor advantage disappeared when the distance was stretched. This suggests that our findings from experiment 1 were due to the presence of action possibilities—the potential possibility of reaching and manipulating with one’s actions, rather than the execution of such action. As a result, the same pointing action was no longer advantageous when it could not bring changes to the external environment.

One interesting observation is that, even when action possibility was taken away, participants’ performance in the sensorimotor condition did not drop to chance level of 17%, but to 23%. If the possibility of asserting action was entirely responsible for the sensorimotor advantage in experiment 1, one would expect that forcing oneself to operate with the sensorimotor system in the absence of any possibility for action would be disadvantageous. One speculation would be that perhaps the laser pointer provided our participants in the long-distance condition a possibility for action to some degree. Findings have been reported, however, documenting how laser pointers do not afford as much action as other tools (Longo and Lourenco 2006). Another plausible explanation would be that as action possibility was taken away, the cognitive visual system took over and therefore no difference in performance was observed.

Before we conclude that action possibility was mainly responsible for the results in experiments 1 and 2, there is a confounding factor that needs to be resolved. That is, as we stretched the distance between the observers and the display, we not only took away the possibility for asserting actions; the extended distance itself also became a potential variable between the two experiments. Thus there remains the possibility that it was the distance, and not the possibility for asserting actions, that was responsible for our results in experiment 1. To account for this confounding variable, we designed experiment 3 to take away any direct-action possibility while keeping the stimulus distance the same as in experiment 1 (60 cm).

4 Experiment 3
This experiment represents an attempt to dissociate the confounding factors of observer–stimulus distance and action possibility. To do this, we kept the display within observers’ direct action space as in experiment 1 (60 cm), but instructed our participants to respond with a laser pointer instead of a jabbing motion. Thus we were asking whether the sensorimotor advantage from experiment 1, and the absence of such effect in experiment 2, was due to distance or action possibility.

4.1 Methods
4.1.1 Participants. Eighty students participated in the experiment, forty in the verbal and forty in the pointing condition.

4.1.2 Material and procedure. The experimental setup and procedure were identical to those in experiment 1 with one exception: participants in the pointing group held a laser pointer close to their body and responded with the pointer instead of directly jabbing the display with their finger.

4.2 Results
Hit rates of 22.8% and 23.5% were recorded for the verbal report and laser pointer condition, respectively. A t-test revealed no significance (t_{38} = 0.47, p = 0.64), indicating an absence of the sensorimotor advantage over verbal report in this setup. This implies that the null result we have observed in experiment 2 was not due to the extended distance between the observer and the stimulus.

4.3 Discussion
This experiment was designed to rule out the confounding variable of observer–display distance from the effect of action possibility. To do this, we kept the distance
at 60 cm, which was within observers’ direct action space, but took away the possibility of performing an action by instructing the observers to respond with a laser pointer. If the sensorimotor advantage was still present, then the results from experiments 1 and 2 could be interpreted as effect of near space, which did not necessarily require actual actions or the possibility for actions. Alternatively, if the sensorimotor advantage disappeared as in experiment 2, then we could conclude that it was the absence of action possibility that was responsible for the null results in experiments 2 and 3. This was indeed what we found: when action possibility was taken away, it did not matter whether the stimulus was within the action space (experiment 3) or out of reach (experiment 2). Therefore, it seems that the distance between the observers and the stimuli was not the key here. These results are consistent with the current literature suggesting that the distinction between near and far space cannot be clearly defined because one can enlarge the extent of near space with tools (Berti and Frassinetti 2000; Witt et al 2005) and shrink it by weighting the arms (Longo and Lourenco 2009), which suggests a gradual transition between near and far space instead of an abrupt boundary separating the two (Longo and Lourenco 2006, 2007).

5 General discussion
Our data here show that action responses provided something unique over and above cognitive verbal reports in implicit change detection. This advantage disappeared when the possibility of manipulating the stimulus via one’s action was taken away, either by extending the distance beyond arm’s reach (experiment 2) or by restricting one’s movement (experiment 3). Specifically, participants in experiment 2 could freely point to the screen and participants in experiment 3 could see the stimulus in their near space, yet no sensorimotor advantage was observed. Therefore, the effect from experiment 1 was most likely due to action possibility rather than the actual pointing action or close distance.

Interestingly, when action possibility was taken away, performance in the pointing condition of experiments 2 and 3 did not drop to chance level of 17%. As mentioned before, if the opportunity to assert action was entirely responsible for the motor advantage in experiment 1, one would expect that forcing oneself to operate with the sensorimotor system in the absence of any possibility for action would be disadvantageous. This was not the case according to our results, which suggest that an advantage, whether cognitive or motor, is to some extent preserved. One plausible explanation, which has been suggested before (Bridgeman et al 1997), would be that, as action possibility was taken away, the cognitive visual system took over and therefore no difference in performance was observed. Hence the nearly identical hit rates in experiments 2 and 3 could have been driven by the same mechanisms underlying verbal report. In this light, the effects from experiment 1 can be seen as an additive effect over and above the advantage that the cognitive visual system has over chance. This was especially true for experiment 3, since participants held the pointer close to their body; holding a pointer in that way results in having the arm bent and using slight movements of the hand to direct the laser beam. This form of movement could differ too much from the natural pointing action that a direct extraction from the motor-action-related visual maps could not be possible, and requires slower access to the cognitive maps because complex planning of the movement (and possible correction due to feedback from the projected laser point) is required. Therefore, it is plausible that the sensorimotor visual system was only actively involved in implicit change detection in experiment 1 but not in experiments 2 and 3.

5.1 Mechanisms of sensorimotor advantage
Recent neuroscientific studies comparing ERPs from no-change trials and change-blindness trials have found neurological markers that are unique to implicit change detection in frontal and central regions (Fernandez-Duque et al 2003; Kimura et al 2008).
However, no work has been done on the neural basis of perceiving action possibilities in the context of implicit change detection. Although it is tempting to attribute the advantage of affordance to the dorsal and ventral visual pathways (Milner and Goodale 1995), our data do not warrant an anatomical explanation at this point. At a functional level, our result is analogous to many studies that reported an absence of an illusion when responses are measured with motor actions (Bridgeman et al 1981, 1997; Milner and Goodale 1995; Wong and Mack 1981). Notably, an induced Roelofs effect can result in static mislocalizations when measured with a verbal probe but not with a pointing action (Bridgeman et al 1997). This difference in response can be explained with the `cognitive' and `sensorimotor' visual maps humans possess, where the cognitive map includes object identity information as well as object position in relation to the visual context (hence the Roelofs effect), and the sensorimotor map strictly processes action possibilities and object position in relation to egocentric bodily coordinates.

But why should the sensorimotor system be sensitive to changes in pictures in the present paradigm? To begin, we do not think this increased sensitivity is specific to natural-scene pictures, although these pictures do provide more cues for actions. Instead, we think this facilitation is available to all visual stimuli within the intended reaching distance. There is a growing number of studies demonstrating an increased visual sensitivity (Dufour and Touzalin 2008) and attentional engagement (Abrams et al 2008; Lloyd et al 2010; Reed et al 2006, 2010) for objects near the hands. This nearby-hand facilitation can also directly benefit change detection (Tseng and Bridgeman 2010). Interestingly, such facilitation does not actually require the hands to be physically near, as Davoli and Abrams (2009) also showed that mere imagination of a nearby hand position can bring forth the same effect. Studies by Witt, Proffitt, and colleagues (Bhalla and Proffitt 1999; Linkenauger et al 2009; Witt and Proffitt 2008; Witt et al 2004, 2005) also highlight the importance of one's intent to act by demonstrating how such intent can change the perceived distance of an object. Notably, Vishton et al (2007) have found that even preparing to reach can result in a reduced Ebbinghaus illusion, which fits well with the current findings. This converging evidence implies that it is possible that our participants' specific intent to 'jab' the screen has activated the same underlying mechanisms that processes the nearby-hand facilitation effect. This would also explain why we did not observe an improved performance in experiments 2 and 3, because the participants were planning to use the laser pointer rather than physically touching the visual stimuli. Neuroscientific studies have suggested that this 'intent to reach' relies heavily on the activity of the posterior parietal cortex (Andersen and Buneo 2002; Connolly et al 2003; Iriki et al 1996; Snyder et al 2000), a parietal region that marks the beginning of the dorsal sensorimotor pathway. Therefore, it is likely that a motor intent indeed shares many processes with the actual action. Interestingly, the posterior parietal cortex has also been found to play a critical role in encoding and maintaining information in VSTM in a change-detection task (Tseng et al 2010). Together, these findings suggest that the improved performance may have originated from the early sensorimotor processes in the parietal area.

6 Conclusion
The present study is the first study that attempts to explore the effect of action in the context of implicit detection. We found that when the stimulus can be acted upon, people guess the locations of change better with a motor response than a verbal report, although they do not consciously detect any changes. This advantage, however, disappears when the stimulus is moved further away (experiment 2), or when movements are restrained (experiment 3), because in both cases the stimulus can no longer afford any action. Together, these results suggest that action, or the possibility to exert action, plays an important role in the way we implicitly perceive and process visual scenes.
Acknowledgments. We would like to thank Dennis Proffitt, Peter Vishton, Ronald Rensink, and an anonymous reviewer for their insightful comments and suggestions that have greatly improved this manuscript. NRK and DW contributed equally to this work.

References
Aglioti S, DeSouza J F, Goodale M A, 1995 “Size–contrast illusions deceive the eye but not the hand” Current Biology 5 679 – 685
Davoli C C, Abrams R A, 2009 “Reaching out with the imagination” Psychological Science 20 293 – 295
Dufour A, Touzalin P, 2008 “Improved visual sensitivity in the perihand space” Experimental Brain Research 190 91 – 98
Gibson J J, 1979 The Ecological Approach to Visual Perception (Boston, MA: Houghton Mifflin)
Iriki A, Tanaka M, Iwamura Y, 1996 “Coding of modified body schema during tool use by macaque postcentral neurons” NeuroReport 7 2325 – 2330