Hills appear much steeper than they are (Kammann, 1967; Li & Durgin, 2009; Proffitt, Bhalla, Gossweiler, & Midgett, 1995; Ross, 1974). Proffitt has proposed that motor representations of hills are accurate and stable, whereas conscious representations are distorted and malleable (e.g., Creem & Proffitt, 2001; Witt & Proffitt, 2007). The entire body of evidence for accurate and stable motor representations of hills depends on data from a single measurement technique: the palm board. Proffitt has consistently asserted that adjusting a palm board is a "visually guided action." We sought to test this claim by asking participants to adjust a palm board to match a real surface with their unseen hand, their settings being nearly perfect when referenced to the central plane of the hand, as shown in Fig. 1. We have replicated this finding several times (Durgin, Li & Hajnal, 2010; Li & Durgin, 2010a,b). We believe it shows that hand gestures are well-calibrated for surfaces in reach. The fact that palm board estimates differ from free-hand gestures in our experiments requires explanation. We have observed that, in most experimental set-ups, palm board adjustments must be accomplished primarily by flexing the wrist. A freely-gestured hand does not have this constraint. We measured proprioception of wrist flexion and found that it was biased and noisy (Durgin, Hajnal, Li, Tonge, & Stigliani, 2010). Proffitt and Zadra (2010) claim that our demonstration that palm boards are highly inaccurate is irrelevant to interpreting past and present findings concerning dissociations between verbal reports and palm board estimates. In their paper they offer a theoretical representation of the findings of Bhalla and Proffitt (1999) and argue that our analysis is incompatible with their account. We offer here an alternative account of the findings of Bhalla and Proffitt, based on their actual data (which are fully compatible with our original analysis). We further show that our account generalizes to more recent studies that continue (1) to mistakenly describe null statistical effects on (insensitive) palm boards as evidence of a "dissociation" from (more sensitive) verbal measures that show a similar relative magnitude of change and (2) to introduce uncontrolled demand characteristics.

We recently showed that palm board measures are systematically inaccurate for full-cue surfaces within reach of one's hand, whereas free-hand gestures and reaching actions are quite accurate for such surfaces (Durgin, Hajnal, Li, Tonge, & Stigliani, 2010). Proffitt and Zadra (2010) claim that our demonstration that palm boards are highly inaccurate is irrelevant to interpreting past and present findings concerning dissociations between verbal reports and palm board estimates. In their paper they offer a theoretical representation of the findings of Bhalla and Proffitt (1999) and argue that our analysis is incompatible with their account. We offer here an alternative account of the findings of Bhalla and Proffitt, based on their actual data (which are fully compatible with our original analysis). We further show that our account generalizes to more recent studies that continue (1) to mistakenly describe null statistical effects on (insensitive) palm boards as evidence of a "dissociation" from (more sensitive) verbal measures that show a similar relative magnitude of change and (2) to introduce uncontrolled demand characteristics.

© 2010 Elsevier B.V. All rights reserved.
In their study of elderly participants, we were surprised to discover upon plotting their data that, contrary to their hypothesis, the elderly evidently gave significantly lower estimates than the younger controls for the majority of hills tested (see Durgin, Hajnal et al., 2010, Fig. 14); Bhalla and Proffitt (1999) summarized the data as showing that the elderly saw hills as steeper by separately analyzing only the steepest hills.

In sum, the evidence that Bhalla and Proffitt (1999) used to argue that hill perception is affected by “behavioral potential” is actually quite weak and equally well supports the conclusion that perception is unaffected by all of these factors, but that judgments of hills are sometimes swayed by the social context or methodological artifacts of the experiments.

As for the imputed dissociation between palm board measures and verbal measures, we have pointed out that this is typically reported as a null result on one measure (the palm board estimates) paired with a positive result on the other (e.g., verbal reports). But a meta-analysis of the two studies of fatigue reported by Proffitt et al. (1995) and by Bhalla and Proffitt (1999) showed that although both experiments reported a null effect of palm board estimates for 5° hills, the combined data indicated that the 20% increase in palm board settings found in each of the two experiments (e.g., from 7.9° to 9.5° in the report of Bhalla & Proffitt, 1999) was not due to chance (see Durgin, Hajnal et al., 2010, Table 2). That is, the supposed dissociation between verbal measures and palm board measures in the fatigue studies is not supported by a more complete look at the data. According to our view, this shows that palm board measures are also susceptible to experimental demand characteristics, but, because they are insensitive measures, they require a larger N to obtain the same statistical effect.

Proffitt and Zadra (2010) argue that because adjusting a palm board involves motor manipulation of a surface, it is a “motoric” dependent measure that is unlike gesturing with a free hand. But free-hand measures are much more accurate for near surfaces (Durgin, Hajnal et al., 2010; Li & Durgin, 2010a). If the palm board is a privileged measure and the free hand is not, why is the free-hand measure more accurate for surfaces in reach? Our alternative hypothesis is that palm board adjustments reflect errors in wrist-proproception (because flexing the elbow when adjusting a palm board tends to remove the hand from the palm board), whereas free-hand gestures allow participants to use the hand in a more typical fashion. Consistent with this hypothesis, we have recently found that free-hand gestures are more accurate when participants spontaneously use more elbow flexion than wrist flexion when they depend primarily on wrist flexion (Li & Durgin, 2010b). Typically, successful gesturing of orientation is accomplished by a combination of elbow and wrist flexion with about 80% of the rotation being produced by the elbow joint and only 20% by the wrist (Li & Durgin, 2010b). We are not proposing that a free-hand gesture is a visually guided action. Rather, we think this evidence shows that free-hand gestures achieved by typical postures of elbow and wrist are calibrated to experience.

Does our account generalize? Proffitt (2009) and Proffitt and Zadra (2010) argue that our experiments, many of which were conducted indoors, do not generalize to outdoor hills. But it is our theoretical conclusions that are meant to generalize. That is, it is known that farther portions of hills appear steeper than nearer portions (Bridgeman & Hoover, 2008; Feresin & Agostini, 2007; Ross, 1974). Based on our findings indoors, this should mean that free-hand estimates of hills will increase and no longer be accurate, and that palm board estimates will also increase and seem to become more accurate; this inference was supported by our observations and those of Proffitt et al. (1995) and of Bridgeman and Hoover. What our experiments suggest is that palm board estimates of (misperceived) hill surfaces may sometimes seem accurate because palm boards are also misperceived. After all, even a broken clock is right twice a day. In addition, imputed dissociations between palm board measures and verbal measures.

may sometimes result from the fact that palm boards are insensitive measures and therefore less statistically powerful, as in the fatigue study.

This claim generalizes quite nicely to more recent reports. For example, Schnall, Zadra, and Proffitt (2010) recently reported that the effect of wearing a backpack on hill estimation differed between participants with different levels of blood sugar (all participants arrived at the lab with depleted blood sugar; half were given a sugar drink and half an artificially-sweetened drink matched for taste). In their study they reported a reliable effect of their blood sugar manipulation on verbal and visual measures and a null result on palm board measures. This is the kind of finding that they interpret as evidence of a dissociation. However, the relative magnitudes of the (non-significant) effect on the palm board and of the (significant) effect on the verbal measure were identical (18% of the respective means). Specifically the palm board effect was 2.2° (18% of the mean palm board estimates of 12.4°—note that the hill was only 6°) while the effect on the verbal measure was 5.1° (18% of the mean verbal estimate of 27.3°). Of course, the proportional variability of palm board measures was much higher than that of other measures, which is sufficient to explain why the same proportional palm board change is not statistically reliable. This provides a good (new) example of one of the problems identified by our study: When two measures are differentially sensitive, finding a statistical null result with the less sensitive one is not good scientific grounds for arguing for a dissociation. (To demonstrate a meaningful dissociation between measures like these, with different scales, it would be better to test for the presence of a statistical interaction in an analysis of log-transformed data.)

Second, although Schnall et al. (2010) argued that their hidden manipulation of blood sugar could not have produced experimental demands, they introduced an uncontrolled demand characteristic by requiring all their participants to wear a heavy backpack while making their slant judgments. We have pointed out that the effects of experimental demand depend on factors that affect the rate of compliance, rather than on the mere awareness of the demand (Durgin et al., 2009). These factors are not all known, but Schnall, Harber, Stefanucci, and Proffitt (2008) have shown, for example, that having a friend along reduces slant judgments when wearing a heavy backpack. Because we have shown that backpack effects can be driven by the compliance of a subset of participants, we interpret the friend effect as evidence that social support reduces the rate of compliance with experimental demand. A similar interpretation applies to the sugar study: Participants who were feeling less depleted (i.e., who had been given a sugary drink) might be more willing (or able) to resist the obvious experimental demand characteristic introduced by the heavy backpack. Given that the study concerned blood sugar, and was intended to control for experimental demand, it is noteworthy that a backpack was introduced at all.

It is also noteworthy that Proffitt and Zadra (2010) defend Proffitt’s theory by presenting “idealized” data rather than real data. In a typical instance Proffitt and Zadra state the following:

“For example, when participants who were low in physical fitness viewed a 5° hill, they reported that it was 25° but made an accurate 5° palmboard adjustment. When asked to make a palmboard adjustment of 25°, they provided a 5° response. These individuals exhibited an internal consistency between their conscious representations and motoric actions (Bhalla & Proffitt, 1999).”

(Proffitt & Zadra, 2010, p. 2)

This statement is not a factual description of any actual data. In response to our objections during the review process, Proffitt and Zadra have added a footnote indicating that the palm board estimates are “idealized” to “exaggerate its accuracy for purposes of expository clarity” (p. 2). But the actual data collected by Bhalla and Proffitt (1999) contradicts several of the theoretical claims made earlier. We have already pointed out that fitness was confounded with sex as a result of recruiting male athletes for the study. Here are four additional theoretically-relevant facts concerning the original data of this study:

1. Dissociation? A null effect was found for palm boards in the fitness study. No numeric data is available concerning this, but their graph suggests that for the 5° hill, this null palm board effect was (again) accompanied by a 20% numeric difference in the same direction as the reliable verbal difference (perhaps due to sex differences as well).

2. Palm board accuracy? On average, participants in this study made highly inaccurate 10.3° palm board adjustments to the 5° hill (Bhalla & Proffitt, Table 2, p. 1082).

3. Internal consistency? In this study, the 5° hill was estimated (verbally) to be 20.4°. To be internally consistent, this implies that when asked to produce “20” by haptic adjustment, these participants should have made an adjustment equivalent to their palm board estimates of the hill (i.e., about 10.3°). In fact, their mean haptic adjustment to an instruction of “20°” was 14.9° (SE=0.6) (Bhalla & Proffitt, 1999, Table 4), which is reliably higher than the 10.3° (SE=1.3) palm board adjustments made outdoors by the same participants. The actual data therefore contradict the claim of “internal consistency.” The discrepancy may signal that waist level was not consistently used for haptic production indoors where keeping the palm board out of view was not an issue.

4. Recalibration? Although Bhalla and Proffitt predicted that those of low fitness would make lower productions of palm board measures in response to verbal prompts, they actually found no statistical evidence of this except at the two highest requested orientations (60° and 75°) where wrist-joint flexibility becomes an issue. (This was already discussed by Durgin, Hajnal et al., 2010.)

In short, in the actual data of Bhalla and Proffitt (1999), the palm board results were either ambiguous (null effects) or directly contradicted the theoretical claims (e.g., of internal consistency) that Proffitt and Zadra (2010) made in the quote earlier.

Palm boards are noisy, inaccurate measures (Durgin, Hajnal et al., 2010, Experiments 1 and 2). They are affected by minor changes in posture (Durgin, Hajnal et al., 2010, Experiment 5). In the aforementioned quote Proffitt and Zadra (2010) have actually paraphrased a section of Bhalla and Proffitt’s General Discussion in which Bhalla and Proffitt were explicitly providing a dramatization of an imaginary observer.1 Proffitt and Zadra conclude that it is “impossible to explain” these (dramatized) findings based on our theory. They are correct. However, our theory can easily explain the actual data of Bhalla and Proffitt (1999). Finally, although this point is of little theoretical import, Proffitt and Zadra (2010) say that we “misrepresent” (p. 4) the experimental conditions of Creem and Proffitt (1998). They write that they “informed” (p. 3) us that participants in that study were standing. What they actually told us during a review process was that they did not recall having participants sitting during that study.2 We did not regard this as clarifying information because human memory is

---

1 The original passage from Bhalla and Proffitt (1999) was explicitly a dramatization: “Now consider another observer at a low level of physical fitness. When viewing the 5° hill, he reports that it is 25°, but makes an accurate 5° motor adjustment. When asked to make a motor adjustment of 25°, he provides a 5° response.” (Bhalla & Proffitt, 1999, p. 1092).

2 We were informed, in part, as follows: “I doubt that participants in our studies were seated when making palm board adjustments; I have no recollection that they were so, but I cannot say for sure.” (Proffitt, 2009, personal communication).
fallible. Here is a direct quote from the relevant methods section of Creem and Proffitt: “The subjects were seated...Among more general questions, the subjects were asked to judge the incline of the hill relative to the horizontal, giving a verbal and haptic measure.” (p. 29, emphasis added). There is no contradiction in adjusting a palm board while seated, and we stand by our statement that Creem and Proffitt reported that their participants were seated.

We emphasize that Experiment 5 of Durgin, Hajnal et al. (2010) is theoretically significant even if Creem and Proffitt’s (1998) participants were not seated. We discovered that a highly reliable change in palm board outputs (by about 5°) could be produced as an artifact of a slight change in palm board height (by about 4 in. or 10 cm) for standing observers. This had not been previously demonstrated, and it is an important observation both because it addresses Proffitt’s (2009) claim that palm boards are demonstrably sensitive measures (we would say sensitive to the wrong things), and because it may provide an explanation of why isolated published examples can be found of very accurate palm board estimates of hills (e.g., Witt & Proffitt, 2007) that differ from other published reports (e.g., Schnall et al., 2010) where palm board adjustments routinely overestimate the slants of low hills by a factor of 2.

In summary, we understand that Proffitt and Zadra (2010) remain committed to their theoretical perspective, but their paper does not address the theoretical or methodological issues that were raised by our article. Contrary to their claims, it is reasonable to make theoretical generalizations of the kind we have made: (1) Free-hand gestures are accurate for surfaces within reach, but palm boards are not (Durgin et al., Experiments 1 and 3; Li & Durgin, 2010a); this implies that palm boards are not special conduits to the dorsal stream. (2) Proprioception of wrist flexion is exaggerated; this can account for why palm boards feel steeper than they are (Durgin, Hajnal et al., 2010, Experiment 4; Durgin, Li et al., 2010; Durgin et al., in press; Li & Durgin, 2010b). (3) Participants asked to wear heavy backpacks usually suspect that the experimenter wants them to provide higher slant estimates (Durgin et al., 2009; Durgin, et al., in press; Russell & Durgin, 2008); this is not avoided by taking them outdoors or giving them sugary drinks. (4) Palm boards are measurably noisier measures than verbal estimates (Durgin, Hajnal et al., 2010, Experiment 2) and noisier than free-hand measures (Durgin, Hajnal et al., 2010, Experiments 1 and 2); this can account for many “dissociations,” which are often simply null effects paired with positive effects of the same proportional magnitude. (5) Small changes in posture can produce large (artifactual) changes in palm board measures (Durgin, Hajnal et al., 2010, Experiment 5); this means any particular palm board estimate is hard to interpret (i.e., as “accurate”). None of these observations is specific to the stimuli employed.

The empirical observations of Proffitt et al. (1995) were indeed a valuable contribution to the study of geographical slant perception, but their theory that palm boards are “visually guided action measures” has been falsified, and Proffitt and Zadra (2010) did not seek to defend that claim. Their claims that palm boards are usually accurate for outdoor hills and that they dissociate from verbal measures have been overstated. Proffitt and Zadra go so far as to “idealize” the numbers they discuss in order to avoid “confusing” the reader with the fact that typical palm board estimates of low hills, though lower than verbal measures, are not particularly accurate. We believe that palm board measures may now be of more sociological interest than scientific. We advocate adopting new approaches to studying perceived slant (Durgin & Li, 2010; Li & Durgin, 2010a,b,c).

References


http://www.journalofvision.org/7/9/282/


3 Feresin and Agostini (2007) varied a number of other factors besides height and did not do a direct comparison of palm board heights. He, Hong, and Ooi (2007) reported effects of varying palm board height, but they varied it much more dramatically.