Discussion forum

Do visual illusions affect grasping? Considerable progress in a scientific debate. A reply to Whitwell & Goodale, 2016

Karl K. Kopiske a,*, Nicola Bruno b, Constanze Hesse c, Thomas Schenk d and Volker H. Franz e

a Center for Neuroscience and Cognitive Systems@UniTn, Istituto Italiano di Tecnologia (IIT), Rovereto, TN, Italy
b Dipartimento di Neuroscienze, Università di Parma, Unità di Psicologia, Parma, Italy
c University of Aberdeen, School of Psychology, William Guild Building, King’s College, Aberdeen, United Kingdom
d Ludwig-Maximilians Universität München, Department of Psychology, Munich, Germany
e University of Tübingen, Institute of Computer Sciences, Department of Experimental Cognitive Sciences, Tübingen, Germany

1. Our registered report and the illusion debate

When we set out to perform our preregistered study (Kopiske, Bruno, Hesse, Schenk, & Franz, 2016), our goal was to clarify whether or not grasping is affected by the Ebbinghaus illusion. This seemingly simple question has far-reaching theoretical consequences for our understanding of the functional architecture of the visual brain, and in particular for the two-visual systems hypothesis (TVSH; Milner & Goodale, 1995, 2006).

We preregistered our design before collecting any data, painstakingly trying to avoid any methodological pitfalls that might compromise the interpretation. Two expert reviewers (at least one of them being a strong advocate of the TVSH) provided detailed input for improving our design and we adapted our study accordingly. Only after the design had been approved did we perform our large study with \( N = 144 \) participants and collected data in parallel in four different labs, intending to provide the best test to-date of whether or not grasping is affected by visual illusions, as proposed by the TVSH.

However, Whitwell and Goodale argue that our study was methodologically weak and misguided from the outset because we presented only one Ebbinghaus display at a time, while the predictions of the TVSH could only be tested when simultaneously presenting a pair of two Ebbinghaus displays. In consequence, they think we missed our target and failed to contribute anything new. Here, we argue that this is far too grim a view. The methodological critique offered by Whitwell and Goodale is not justified, and the claim that nothing new has been contributed ignores that a de-facto consensus has been reached on a number of facts, as indirectly also acknowledged by Whitwell and Goodale. These facts will in the future facilitate the scientific debate by narrowing down the contentious issues in need of clarification. We will first describe this de-facto consensus before we turn our attention to Whitwell and Goodale’s main critique.
2. De-facto consensus: single Ebbinghaus displays affect grasping as well as perception

In contrast to previous papers (e.g., Goodale, 2008, 2011), Whitwell and Goodale no longer question that there is a clear effect of a single Ebbinghaus display on grasping and that this effect is of the same size as the effect on perception. This is substantial progress, such that scientists should be able to close the files on this question.

Whitwell and Goodale also concede that our study rules out obstacle avoidance mechanisms as the reason for the effects of single Ebbinghaus displays on grasping (“We actually have no issue with this aspect of their study”). This too is progress in the scientific debate, and notably so, given that obstacle avoidance has been the most frequent explanation of TVSH-advocates for why the effects of the Ebbinghaus illusion on grasping should not be attributed to the same processes as the effects of the illusion on perception (Goodale, 2008, 2011; Haffenden, Schiff, & Goodale, 2001; Milner & Goodale, 2008). This also has consequences for the interpretation of studies on other illusions. For example, Whitwell, Buckingham, Enns, Chouinard, and Goodale (2016) used this obstacle-avoidance hypothesis as an argument of why unwanted effects of the Ponzo illusion on grasping should be attributed to different processes than the illusion in perception.

3. Theoretical consequences of this consensus

Despite this de-facto consensus, there is disagreement with respect to its theoretical implications. While we have argued that this finding is not consistent with key notions of the TVSH (see: Aglioti, DeSouza, & Goodale, 1995; Milner & Goodale, 2006, p. 242; Goodale & Ganel, 2016), Whitwell and Goodale argue that single Ebbinghaus displays cannot be used at all to test the validity of the TVSH, that therefore our findings are irrelevant for the TVSH, and that the TVSH can only be tested using dual Ebbinghaus displays.

Before discussing Whitwell and Goodale’s dual-Ebbinghaus-only-conjecture, let us point out that their argument is inconsistent with earlier papers from the Goodale-group such that it does not strike us as very convincing. TVSH-proponents have themselves used single Ebbinghaus displays (Haffenden et al., 2001) and concluded that their single Ebbinghaus experiments provide “compelling evidence that the size-contrast illusion elicited by the Ebbinghaus display does not affect grasp scaling” (p. 180), a statement echoed by Ganel, Tanzer, and Goodale (2008). Why, if it was a-priori so clear that single Ebbinghaus displays are not appropriate to test the TVSH, were those displays used in those earlier studies with exactly that purpose? This concern has only now been raised by Whitwell and Goodale. That is, after our results have clearly shown that there is no dissociation between perception and action with single Ebbinghaus displays.

However, such post-hoc reasoning is scientifically problematic (see, e.g., Kerr, 1998). In fact, precluding post-hoc reasoning was one of the main reasons to implement pre-registration in Cortex and other journals (Chambers, Dienes, McIntosh, Rotshtein, & Willmes, 2015). Nevertheless, we will consider below the suggested possibility that the TVSH can be meaningfully tested only with dual Ebbinghaus displays, but not with single Ebbinghaus displays.

4. Are dual-Ebbinghaus displays the only valid tests of the TVSH predictions?

Whitwell and Goodale argue that the illusion effects of single Ebbinghaus displays are too small to test the proposed dissociation. However, the size of the illusion effects cannot be the problem because many studies did find effects of single Ebbinghaus displays on grasping as well as on perception.

Now, one could argue that the purported differences between illusion effects on grasping and on perception are too small in single Ebbinghaus displays and that those differences only show up reliably in dual Ebbinghaus displays. However, the large sample size and corresponding a-priori power analysis in our registered report (eight times as many participants as in the largest dual display study; Haffenden & Goodale, 1998), as well as using Bayes factors, and a condition with perceptually matched discs designed specifically to be sensitive to small differences, all rule out the size of the effect of or the differences as potential problems.

In consequence, to make the case that our single-Ebbinghaus-display data should be dismissed, Whitwell and Goodale would have to assume that the dissociation between perception and grasping only exists if we use dual Ebbinghaus displays. By this they assume a qualitatively different, new illusion process, which is active only in dual Ebbinghaus displays, and only for this illusion process the purported dissociation between perception and grasping is existent. This would be a completely new assumption, and we are unaware of any evidence that supports it. The assumption would also be inconsistent with the logic of the TVSH: The TVSH assumes that grasping is unaffected by the Ebbinghaus illusion because it is a contextual effect (Milner & Dyde, 2003). Why then should the single Ebbinghaus illusion (which also is a contextual effect) be allowed by the TVSH to affect grasping? Finally, we want to stress that single Ebbinghaus displays have been typically used in classic studies of the perceptual illusion (e.g., in Coren & Enns, 1993; Coren & Gignus, 1972; Gignus, Coren, & Agdern, 1972), so why should they be inappropriate to test for a possible dissociation between perception and grasping?

However, again, it is a logical possibility that for some hitherto unknown reason the dissociation between perception and grasping can only be detected with dual Ebbinghaus displays but not with single Ebbinghaus displays. Therefore,

1 Note that this process cannot be the superadditivity of the Ebbinghaus illusion (cf. Foster & Franz, 2014; Franz et al., 2000), because superadditivity can be switched on and off in perceptual measures depending on the task demands (cf. Experiment 3 of Franz et al. (2000). If task demands are matched for perceptual measures and grasping there is no difference between illusion effects on perception and grasping; see also our discussion of superadditivity in the next paragraphs.
let us briefly review whether there is empirical evidence for this notion.

As Whitwell and Goodale point out, there are two prominent grasping studies that used dual Ebbinghaus displays: Aglioti et al. (1995) and Haffenden and Goodale (1998). Both have been taken as evidence for a dissociation between grasping and perception. However, in the first study (Aglioti et al., 1995), task demands were not well matched between grasping and perception (Franz, Gegenfurtner, Bülthoff, & Fahle, 2000; Pavani, Boscaglì, Benvenuti, Rabuffetti, & Farnè, 1999): In grasping, participants operated on only one Ebbinghaus display at a time, while in perception they performed a direct comparison between the target discs of the two Ebbinghaus displays, thereby simultaneously operating on both Ebbinghaus displays. This mismatch is known—as also acknowledged by Whitwell and Goodale—to create an increase of the illusion effect of about 50% (Franz et al., 2000; see also Foster & Franz, 2014), which corresponds well to the difference Aglioti et al. (1995) found between perception and grasping. Therefore, Aglioti et al. (1995) cannot be considered strong evidence for the TVSH. This leaves the Haffenden and Goodale (1998) study, which will be discussed in the next section.

5. Is Haffenden and Goodale (1998) the most decisive study?

Whitwell and Goodale suggest that the study by Haffenden and Goodale (1998) is currently the best test of the TVSH. They argue that the problem of mismatched task demands was avoided in that study (despite using a dual illusion display) by using manual size estimation (ME), where participants indicate the size of an object with index finger and thumb. ME is interpreted as a perceptual measure in the framework of the TVSH. Because participants estimated only one of the central discs of the dual Ebbinghaus display at a time (operating only on one disc, just as in grasping), Whitwell and Goodale argue that there was no mismatch of task demands. Furthermore, Whitwell and Goodale present a reanalysis of the data of Haffenden and Goodale (1998), and calculated for the first time the slope-corrected illusion effects for grasping and ME. They demonstrate that even after slope correction, the illusion effects in ME are much bigger than in grasping.

It is commendable that the appropriate quantitative estimates for the illusion effects are now available for Haffenden and Goodale’s (1998) study. However, there are problems that make us reluctant to accept these recalculations as a strong argument for the proposed dissociation between perception and grasping in visual illusions:

Firstly, the study is only one of multiple studies that investigated the predictions of the TVSH for the Ebbinghaus illusion. If the other studies were now essentially to be ignored, this would constitute a strategy that vastly increases the chances of finding support for just about any given hypothesis (see e.g., Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). If, therefore, Haffenden and Goodale’s (1998) study should from now on be the most central argument for the proposed dissociation between grasping and perception in the Ebbinghaus illusion, it would need to be replicated and tested. In Kopiske et al. (2016) we did such a replication and test of Haffenden et al. (2001), another study that was considered to be decisive evidence. Haffenden et al.’s (2001) conclusions did not stand the empirical test—as also acknowledged by Whitwell and Goodale (most notably the idea that the effects of a single Ebbinghaus display on grasping are caused by obstacle-avoidance mechanisms independent of perception).

Secondly, a serious problem of the Haffenden and Goodale (1998) study are the discrepant findings obtained for the two perceptual measures Haffenden and Goodale (1998) measured not only ME, but also a standard perceptual size-matching task. In this task, participants directly compared and matched two central discs in the dual Ebbinghaus displays until they perceived these discs to be equal in size. This yielded a perceptual illusion effect of approximately 2.4 mm. In comparison the newly calculated illusion effect in ME is almost twice as big: About 4.7 mm (our Fig. 1 and Figure 2 of Whitwell & Goodale).³

This strong inconsistency between the two perceptual measures is even more surprising if we take into account that in ME there is no superadditivity to be expected (as also argued by Whitwell & Goodale). This is so, because participants operated on only one of the two Ebbinghaus displays at a time (just as in grasping). In the standard perceptual size-matching task, on the other hand, the illusion effect should be increased by approximately 50% due to the superadditivity induced by the direct comparison of the two illusory displays (as also acknowledged by Whitwell & Goodale). If we take into account this mismatch in task demands, we obtain an illusion effect of approximately 1.6 mm for standard perception (2.4*100/150 = 1.6) as the most appropriate value to be compared to the illusion effect in ME (cf. Fig. 1). This demonstrates that the two measures of perception in Haffenden and Goodale’s (1998) study are dramatically different. In contrast, studies that systematically compared ME to standard perceptual measures (Franz, 2003; Kopiske et al., 2016) obtained similar illusion effects for both larger.

³ We will not discuss the question of what exactly ME measures in further detail here. It seems clear, however, that if ME is a perceptual measure, it should yield results consistent with traditional perceptual measures as, e.g., the methods of adjustment or constant stimuli. To our knowledge, the only systematic investigations into this question have been performed by Franz (2003) and Kopiske et al. (2016), who show that ME can respond with quite a different gain (slope) to a variation of physical size than traditional perceptual measures. In these cases, we need to accurately measure and correct for the response-slope, as now seems to be acknowledged by Whitwell and Goodale (but was questioned in earlier publications of this group).

³ The slopes of standard perception were not measured in Haffenden and Goodale (1998), therefore it is not possible to slope-correct those standard-perception illusion effects. However, we know that the slope of standard perception is typically close to 1; therefore we can use the uncorrected data as a fairly good approximation and can compare this approximation to the slope-corrected illusion effects of ME.

---

212 C O R T E X 8 8 ( 2 0 1 7 ) 2 1 0 – 2 1 5
Fig. 1 – Illusion effects in studies comparing grasping (MGA) to manual estimation (ME) as well as a standard perceptual measure. Illusion effects are in percent relative to the physical size of the stimuli; all illusion effects are slope-corrected (Bruno & Franz, 2009; Franz, 2007; Franz, Scharnowski, & Gegenfurtner, 2005; Hesse, Franz, & Schenk, 2016; Kopiske et al., 2016, p. 139); except for standard perception of HG98 (see footnote 3). All studies used roughly similar Ebbinghaus displays (HG98: SN/LF, F03: SN/LN, Kopiske et al., 2016: SN/LF, see Kopiske et al., 2016 for nomenclature). Aggregated data for HG98 were kindly provided by M. Goodale and R. Whitwell (personal communication, July, 29th and Aug, 19th, 2016). Error bars indicate the SEM of the corrected illusion effect, estimated using a Taylor-approximation (cf. Kopiske et al., 2016, p. 139 and Hesse et al., 2016, p. 94 for an equivalent but simplified formula). Note, that Whitwell and Goodale used in their Figure 2 the problematic ‘zero-variance method’, that in general underestimates the size of the SEMs. Although for Whitwell and Goodale this effect is not dramatic, we show here the more appropriate Taylor-approximated SEMs (cf. Franz, 2007; Franz et al., 2005 for a discussion of the zero-variance method).

measures, as long as the slope correction was performed and task demands were matched (cf. Fig. 1).

So what do Haffenden and Goodale’s (1998) unusual perceptual illusion effects mean for the comparison to grasping? It is clear that no matter whether we take into account superadditivity or not, the difference between the illusion effects in grasping and standard perception is much smaller than that between standard perception and ME (Fig. 1). Therefore, even this data provides no evidence for a “perceptual cluster” (guided by the ventral stream) versus a “motoric cluster” (guided by the dorsal stream). If anything, Haffenden and Goodale’s (1998) data (but not those of the other studies) suggests that grasping and standard perception are similar but different from ME. Thus, before drawing far-reaching conclusions from this data it will be necessary to clarify why the ME data of this study is so unusual and unexpected – even from the viewpoint of the TVSH.

6. Did we ignore the perceptually-matched condition?

Before closing, we want to discuss a more specific issue: Whitwell and Goodale argue that, historically, we simply ignored the perceptually-matched condition of Aglioti et al. (1995) and Haffenden and Goodale (1998), thereby ignoring a substantial part of the data of those studies. In consequence, it would be no surprise if we came to wrong and biased conclusions. This argument has been brought up repeatedly before and has been responded to (e.g., Franz & Gegenfurtner, 2008). It also seems ironic that it is now raised against Kopiske et al. (2016), a study in which we took great care to laboriously implement such a perceptually-matched condition.

Before describing this condition in Kopiske et al. (2016), let us first comment on the perceptually-matched condition in general: The perceptually-matched condition is a nulling-procedure: A pair of discs is selected that appears perceptually equal in size if one of the disc is surrounded by the enlarging context of the illusion and the other by the shrinking context. If the condition works as intended and perception is equalized, then we can attribute differences in grasping the discs to a different size of the illusion effect between grasping and perception.

However, the perceptually-matched condition has a big disadvantage: Because physical size and illusion are confounded, it is not easy to quantify the size of the illusion effect in grasping. This is a problem if we want to quantitatively compare illusion effects between perception and grasping. Such a quantitative comparison is necessary because studies typically did find at least some illusion effects on grasping (even Aglioti et al., 1995), thereby ruling out ‘strong’ versions of the TVSH that would state complete immunity of grasping to those illusions (as opposed to just a smaller illusion effect in grasping than in perception, as ‘weaker’ versions of the TVSH would state). Thus, all studies (including Aglioti et al., 1995 and the recalculations in Whitwell & Goodale) used the physically-matched conditions to quantify the illusion effect, such that quantitative estimates of the illusion effect are only available for this condition. Note, however, that this is not very critical because there is no reason to assume the illusion effect to be drastically different between perceptually-matched and physically-matched conditions. This is so
because (a) the conditions are very similar (the only difference is that one disc has a slightly different size in the perceptually-matched condition to achieve perceptual equivalence), and (b) we explicitly tested for such a difference between perceptually-matched and physically-matched conditions in Kopiske et al. (2016) and found no differences (Figure 8 of Kopiske et al., 2016).

Finally, let us comment on the perceptually-matched condition of our study: We included this condition for many methodological reasons (as detailed in Kopiske et al., 2016) and as suggested by one reviewer. This condition was performed in a much more controlled way than in Aglioti et al. (1995) and in Haffenden and Goodale (1998): (a) The earlier studies selected the pair of matched discs in a pilot phase by the experimenter using trial and error, while in Kopiske et al. (2016) we used a psychophysical constant stimuli method. (b) Previous studies did not quantitatively test whether the matching actually worked or whether there was a residual mismatch of the pair of discs. We tested this laboriously in a second condition. (c) In those earlier studies, participants could only choose between discs that varied in 1 mm steps. This is much too coarse for an illusion effect of, on average, only 2.4 mm (Haffenden & Goodale, 1998). We used step sizes of .25 mm (which is still not perfect, but much better). Given all these advantages, it is quite surprising that Whitwell and Goodale seem to dismiss the relevance of our perceptually-matched condition.

7. Summary and conclusions

Whitwell and Goodale concede that single Ebbinghaus displays seem to affect grasping to a similar degree as perception and that these effects cannot be attributed to non-perceptual, purely motor processes (obstacle avoidance, awkward grasping). However, they argue that a test of the TVSH can only and exclusively be performed using dual Ebbinghaus displays but not with single Ebbinghaus displays. They therefore suggest that Haffenden and Goodale (1998) is the decisive study to test for a dissociation between grasping and perception. However, as we discussed here, this study has serious problems, because the perceptual measures yielded highly inconsistent illusion effects. Future research should first focus on finding consistent perceptual illusion effects in the Haffenden and Goodale (1998) paradigm before these can be meaningfully compared to grasping data.

In contrast, the extensive tests in Kopiske et al. (2016) have demonstrated consistent illusion effects across a wide variety of perceptual measures and also between perception and grasping. The design of Kopiske et al. (2016) was the result of intensive efforts of four independent research groups and an in-detail critique by two anonymous expert reviewers. Here we have outlined why we think that Whitwell and Goodale’s methodological critique is post-hoc and not convincing, and why we believe that Kopiske et al. (2016) provides a strong and valid test of the claim that certain illusions affect perception more than grasping. The outcome of this test suggests that there is no difference in the effects of the Ebbinghaus illusion on grasping and perception.

References


