
ORIGINAL RESEARCH REPORT

Demand Characteristics Confound the Rubber Hand Illusion

Peter Lush^{*†}

Reports of experiences of ownership over a fake hand following simple multisensory stimulation (the 'rubber hand illusion') have generated an expansive literature. Because such reports might reflect suggestion effects, demand characteristics are routinely controlled for by contrasting agreement ratings for 'illusion' and 'control' conditions. However, these methods have never been validated, and recent evidence that response to imaginative suggestion ('phenomenological control') predicts illusion report prompts reconsideration of their efficacy. A crucial assumption of the standard approach is that demand characteristics are matched across conditions. Here, a quasi-experiment design was employed to test demand characteristics in rubber hand illusion reports. Participants were provided with information about the rubber hand illusion procedure (text description and video demonstration) and recorded expectancies for standard 'illusion' and 'control' statements. Expectancies for 'control' and 'illusion' statements in synchronous and asynchronous conditions were found to differ similarly to published illusion reports. Therefore, rubber hand illusion control methods which have been in use for 22 years are not fit for purpose. Because demand characteristics have not been controlled in illusion report in existing studies, the illusion may be, partially or entirely, a suggestion effect. Methods to develop robust controls are proposed. That confounding demand characteristics have been overlooked for decades may be attributable to a lack of awareness that demand characteristics can drive experience in psychological science.

Keywords: Rubber hand illusion; demand characteristics; imaginative suggestion; suggestion effects; expectancies; embodiment; phenomenological control

In the rubber hand illusion (RHI; Botvinick & Cohen, 1998), synchronous brush strokes on a participant's concealed hand and a visible fake hand prompt reports of illusory sensations of touch and of ownership of the fake hand. The RHI is thought to reflect the role of multimodal integration in embodiment and to demonstrate that a fundamental aspect of conscious selfhood can be disrupted by a surprisingly simple intervention (for reviews see Braun et al, 2018; Riemer, Trojan, Beauchamp & Fuchs 2019). The validity of such claims ultimately rests upon the efficacy of methods to control for demand characteristics ("the totality of cues which convey an experimental hypothesis to the subject", Orne, 1962). Demand characteristics in the RHI (and elsewhere) may act as implicit imaginative suggestions, generating expectancies which are met by the voluntary top-down control of phenomenology (which is experienced as involuntary), just as in direct imaginative suggestion within the context of 'hypnosis' (for an extended treatment of phenomenological control and its relationship to imaginative suggestion

and effects including the RHI, see our preprint Lush et al, 2019; see also Dienes et al, in press). It is therefore necessary to test the validity of existing control methods. Experimental demand characteristics can be tested by 'quasi-experiments' in which participants provided with information about an experimental procedure predict their response (Orne, 1969). Here, a quasi-experimental design is employed to test the demand characteristics of illusion and control measures and therefore the validity of claims that the RHI is not a suggestion effect.

Demand characteristics are generally considered to drive behavioural compliance, but Orne recognised that they can also determine a subject's "actual experience" during an experimental procedure (Orne & Scheiber, 1964). Expectancies arising from direct imaginative suggestion can drive unusual experiences (e.g., hallucinations and apparently involuntary action) in a substantial proportion of the population (within and outside the context of 'hypnosis'; Braffman & Kirsch, 1999). Imaginative suggestion can be implicit rather than direct (Orne, 1959) and expectancies within scientific experiments can drive experiential change (e.g., gustatory hallucinations, Juhasz & Sarbin, 1966; psychedelic experiences, Heaton 1975; see Kirsch & Council, 1989). RHI response is predicted both by participant expectancies and by a stable trait measure of response to direct imaginative suggestion in a hypnotic context (hypnotisability).

* Sackler Centre for Consciousness Science, University of Sussex, Falmer, UK

† School of Informatics, Chichester Building, University of Sussex, Falmer, UK

petelush@gmail.com

Furthermore, expectancies for synchronous and asynchronous induction predict illusion scores (Lush et al, 2019). Illusory experience in the RHI is therefore likely to reflect the top-down control of phenomenology to meet expectancies ('phenomenological control') rather than, or in addition to multi-sensory integration or top-down processes which are not driven by demand characteristics (Lush et al, 2019; Dienes et al, 2019). Multi-sensory integration may play a role in the illusion, but it is also possible that multi-sensory stimuli in the RHI merely provide cues which participants interpret as implicit suggestions for particular experiences. Similarly, theories of the RHI which appeal to top-down processes (e.g., Longo et al, 2008) may merely be referring to effects arising from demand characteristics. Here I use the term 'hypnotisability' to refer to stable trait differences in response to imaginative suggestion within a hypnotic context and 'phenomenological control' to refer to a general ability to meet expectancies (e.g. arising from direct suggestion or from demand characteristics) with top down control of experience in a range of contexts (including hypnosis and scientific experiments).

The RHI has generated an extensive literature, with 3,289 citations of Botvinick and Cohen's 1998 paper at the time of writing (Google Scholar 9/12/2019). The experimental procedure has been extended to a wide range of effects, with illusion experience reported for a variety of body parts including faces (Sforza, Bufalari, Haggard & Aglioti, 2010) tongues (Michel, Velasco, Salgado-Montejo & Spence, 2014), and the whole body (Lenggenhager, Tadi, Metzinger & Blanke, 2007). In addition to subjective report, indirect measures are often claimed to reflect changes in embodiment mechanisms, e.g., perceived hand location (Botvinick & Cohen, 1998), skin conductance response (SCR; Armel & Ramachandran, 2003) and brain imaging

(Ehrsson, Holmes & Passingham, 2005). Because it is subjective report which links indirect measures to changes in experience, these claims depend on the validity of controls for demand characteristics in subjective reports. Note also that suggestion effects may account for indirect measures (see discussion and also Lush et al, 2019).

Botvinick and Cohen (1998) measured subjective experience using Likert scale agreement scores for three 'illusion' statements, two describing experiences of referred touch and one describing ownership of the fake hand. A further six statements described other experiences, for example hallucinations of the hand drifting or turning rubbery (see **Table 1**). These three 'illusion' statements and six 'control' statements (or modifications and subsets of them) have since appeared in the majority of RHI research (see Riemer et al, 2019 for a thorough review of RHI methodology), though it is worth noting that a small minority of researchers consider the 'control' statements not as controls for suggestion but as part of the illusion (e.g. Haans et al 2012).

Demand characteristics in RHI subjective report measures are typically controlled by comparing agreement scores for 'illusion' statements (describing referred touch and ownership experience which the experimenter expects to occur) and 'control' statements (experiences that the experimenter does not expect to occur). If agreement is greater for 'illusion' statements than for 'control' statements (e.g., Kalckert & Ehrsson, 2012), 'illusion' scores are interpreted as evidence of experience which cannot be attributed to demand characteristics or suggestion. Alternatively, the difference between response to 'illusion' and 'control' statements is calculated to provide an 'illusion index' in which demand characteristics are apparently accounted for (e.g., Abdulkarim & Ehrsson, 2016). A crucial assumption

Table 1: Statements, questions and response labels used to generate subjective report scores. All statements taken from Botvinick & Cohen (1998).

'Illusion' and 'control' statements	Scale labels
'Illusion'	
S1. It seemed as if I were feeling the touch of the paintbrush in the location where I saw the rubber hand touched	+3. I am certain I will feel some effect
S2. It seemed as though the touch I felt was caused by the paintbrush touching the rubber hand	+2. I am fairly certain I will feel some effect
S3. I felt as if the rubber hand were my hand	+1. I think I will feel some effect
'Control'	
C1. It felt as if my (real) hand were drifting toward the rubber hand	0. I have no idea either way
C2. It seemed as if I might have more than one left hand or arm'	-1. I think I won't feel any effect
C3. It seemed as if the touch I was feeling came from somewhere between my own hand and the rubber hand'	-2. I am fairly certain I won't feel any effect
C4. It felt as if my (real) hand were turning "rubbery"	-3. I am certain I won't feel any effect.
C5. It appeared (visually) as if the rubber hand were drifting towards the left (towards my hand)	
C6. The rubber hand began to resemble my own (real) hand, in terms of shape, skin tone, freckles or some other visual feature.	

of these control methods is that demand characteristics (and therefore expectancies) for 'control' and 'illusion' statements are closely matched. If this assumption is not met, differences in agreement scores may merely reflect differing demand characteristics. There is no evidence supporting the validity of RHI control methods. As Riemer et al (2019; p. 277) note:

“It is remarkable that the use of control items has become common practice although empirical support justifying this practice is lacking. There is neither a psychometric examination of whether the “control items” are adequate to assess suggestibility nor of whether they are indeed unspecific.”

To control for demand characteristics, what participants expect to happen (and what they believe the experimenter expects to happen) must be matched across 'control' and 'illusion' conditions. If participants can discern which statements are intended as controls, or if implicit task demands are more consistent with one statement than another, expectancies will differ. To illustrate, consider the following statement intentionally designed to generate very low expectancies (both because it is ridiculous and because it is hard to imagine how the described experience could arise from the procedure): “Your hand turned into an octopus tentacle and crawled out of the experimental apparatus toward your face”. Participants might well realise that the experimenter does not really expect them to have this experience, and the brush stroking induction seems unlikely to be interpreted as requiring this experience. Now consider a statement intended to measure referred touch in the RHI: “It seemed as though the touch I felt was caused by the paintbrush touching the rubber hand”. Presented with both statements, participants may well be able to guess which is intended as the experimental 'illusion' measure and which the 'control', and the brush stroke induction seems more likely to be interpreted as requiring referred touch than a hallucinated tentacle. An expectancy driven difference in report between referred touch and tentacle statements should therefore occur even if response to both statements is entirely driven by demand characteristics. This example is purposely extreme in order to clarify the argument. However, there is no reason to believe that expectancies in the more plausible examples (e.g., the participant's hand “turning rubbery” or the rubber hand changing appearance) used by RHI researchers are not confounded in this way.

To what extent is the RHI confounded by phenomenological control? In our previous preprint (Lush et al, 2019), agreement with 'illusion' experience was predicted by hypnotisability with a regression slope of .57 in a sample of 353 participants. For each 1 point increase in SWASH hypnotisability score (0 to 5 scale), RHI 'illusion' statement agreement (7 point scale from -3 to +3) increased by more than half a point. Expectancy for synchronous induction illusion predicted synchronous induction 'illusion' report with a slope of .33 units (both measures on a 7 point scale). For each one point increase in expectancy, illusion report increased by 1/3 of a point. For

proprioceptive drift (a measure of how far toward or away from the fake hand the participant feels their own hand has moved), hypnotisability predicted .53 cm of synchronous condition drift toward the hand per SWASH scale point. These are substantial relationships. Considering the RHI ownership statement in isolation (the key measure which links the illusion to experiences of ownership; Wu, in press), the slope is .76 and the intercept -0.50 . This model predicts that low hypnotisable participants scoring zero on the SWASH will disagree with the ownership statement, but that participants with maximum hypnotisability score will score maximum (+3) agreement for ownership experience. In sum, correlations between hypnotisability and RHI report are comparable to correlations between individual SWASH scale suggestions and the overall SWASH score (with that item dropped; Dienes et al, in press). It is therefore plausible that the RHI is entirely a suggestion effect. Any theory of the RHI has to account for this relationship. The theory that both the RHI and imaginative suggestion in a hypnotic context are driven by phenomenological control accounts for this relationship. It therefore offers considerable advantages in terms of parsimony compared to existing RHI theories. The theory is employed to form the hypothesis tested here: if the RHI is driven by phenomenological control, then existing methods to control suggestion effects in the RHI must be flawed.

It has previously been shown that expectancies for synchronous and asynchronous RHI induction differ (Lush et al, 2019). The validity of claims that the RHI is not a suggestion effect therefore rest on the untested assumption that demand characteristics are closely matched across established RHI 'control' and 'illusion' statements. If they are not, then RHI researchers have for more than twenty years been using methods to control demand characteristics which are themselves confounded by demand characteristics. In this study, this assumption is tested using a quasi-experimental design. Participants were provided with video and text information about the RHI procedure and asked to record their expectations for each statement and condition. It is predicted that expectancy ratings will, like in illusion reports, be higher for 'illusion' than for 'control' conditions and also that, as reported in Lush et al (2019), 'illusion' expectancy ratings will be higher for synchronous than asynchronous conditions.

Method

Participants

Data from 32 participants was recorded. In accordance with pre-registered exclusion criteria (preregistration document available at <https://osf.io/9c8mq>), 12 participants were excluded: 6 for spending less than ten seconds reading the information page and 6 for reporting previous participation in a procedure similar to that shown in the video. Bayes factors were calculated once data for 20 participants (after exclusion) had been collected. Because the Bayes factors for each pre-registered analysis were greater than the preregistered stopping rule threshold (greater than 6 or less than 1/6), data collection ceased. Data from 20 participants (17 females and 3 males) were therefore included (mean age = 20.8, SD = 6.0). Participants were

compensated with course credit. All participants provided informed consent and ethical approval was granted by the University of Sussex ethics committee.

Procedure

Participants completed the study on their own computers. Study materials are available at <https://osf.io/9c8mq>. After providing consent and demographic information, participants were asked to read the following short passage describing the RHI:

“In this procedure, a participant’s own hand is hidden from their view and a fake hand is placed in front of them. An experimenter then uses brushes to stroke the participants hidden real hand and the visible fake hand. The location of the brush strokes on the real and fake hands is matched, so that a downward brush stroke on the participants index finger will be accompanied by a downward brush stroke on the fake hand. Participants can therefore see a paintbrush brushing down the finger on a fake hand while they feel a paintbrush brushing down the finger on their real hand (which they cannot see). There are two conditions in the experimental procedure: **Synchronous condition:** The brush strokes on the real hand and on the fake hand occur at the same time (they are synchronous).

Asynchronous condition: The brush strokes on the participants real hand and on the fake hand occur at different times (they are asynchronous)”.

Participants then watched a 62 second video in which a rubber hand illusion procedure was shown. Onscreen subtitles described the displayed procedure (for example, the synchronous and asynchronous induction conditions). The video showed the induction procedure from the participant’s and experimenter’s perspectives. Following the video, participants were asked to read another short passage:

“In the video, an experimenter performed brush strokes on the participant’s hand and a fake hand in matching locations (the index fingers) on each hand. The participant could see the fake hand but could not see their real hand. There were two conditions. In the synchronous condition, the brush strokes on the real and fake hands occurred at the same time. In the asynchronous condition there was a delay between the brush strokes on the fake hand and the real hand.”

To encourage consideration of the procedure described in the text and shown in the video, participants were then asked to freely respond to the following question: “What do you think this procedure is supposed to cause (what is the participant expected to experience)?”. Participants were then asked to report whether or not they had heard of the procedure before and whether or not they had previously participated in an experiment in which this procedure was used. Participants then reported their

expectancies for each of the 9 statements, in fixed order. Synchronous and asynchronous condition expectancies were recorded for each statement in turn (the order was statement 1 synchronous, statement 1 asynchronous, statement 2 synchronous, statement 2 asynchronous and so on). An anonymous reviewer pointed out that this presentation may lead to order effects. This is plausible, and the approach was chosen to reflect common practice in RHI research. A fixed-order procedure matches that of roughly 80% of RHI studies (see supplemental material at <https://osf.io/9c8mq/>); following majority practice ensures greater generality of conclusions.

Measures

Table 1 shows the ‘illusion’ and ‘control’ statements taken from Botvinick and Cohen (1998) and the scale labels used to measure expectancies for each statement. The seven-item scale is taken from Lush et al, 2019 and is based on the seven point scale which measures agreement and disagreement with RHI statements devised by Botvinick and Cohen (1998). Response to the three ‘illusion’ statement expectancies (S1–S3) was used to calculate a mean ‘illusion’ expectancy score and response to the six ‘control’ statements (C1–C6) was used to calculate a mean ‘control’ expectancy score.

Analyses

Preregistered analyses

Pre-registered analyses were designed to mimic common approaches to testing agreement scores in the RHI and are registered at <https://osf.io/89m7j>. The following text is adapted from the preregistration document.

A t-test of the difference between synchronous condition ‘illusion’ (mean of S1–3) and synchronous condition ‘control’ statement (mean of S4–9) expectancy scores was conducted. A Bayes factor was calculated using a half normal based on the 1 scale point difference in expectancy between synchronous and asynchronous induction reported in Lush et al (2019). If synchronous condition ‘illusion’ statement expectancies (S1–3) are greater than synchronous condition ‘control’ statement expectancies (S4–9), then RHI ‘control’ statements are not suitable controls for suggestion effects because scores would be expected to differ because of differing expectancies even if response to both ‘control’ and ‘illusion’ agreement statements entirely reflect suggestion effects.

A t-test of the difference between synchronous condition ‘illusion’ (mean of S1–3) and asynchronous condition ‘illusion’ (mean of S1–3) expectancy scores was conducted to replicate the result reported in Lush et al (2019). A Bayes factor was calculated using a half normal based on the 1 scale point difference in expectancy between synchronous and asynchronous induction reported in Lush et al (2019). If synchronous condition ‘illusion’ statement expectancies (S1–3) are greater than asynchronous condition ‘control’ statement expectancies (S1–3), then asynchronous induction is not a suitable control for suggestion effects because scores would be expected to differ because of differing expectancies

even if response to both synchronous and asynchronous induction agreement statements entirely reflect suggestion effects.

95% CIs (interpreted as Bayesian credible intervals with uniform priors) were used to estimate 'illusion' (mean of S1–3) and 'control' scores (mean of S4–9). It is predicted that 'control' expectancy CIs will be negative and illusion score expectancies will be positive.

Robustness regions are reported, to indicate the range of scales that qualitatively support a given conclusion (i.e. evidence as insensitive, or as supporting H_0 , or as supporting H_1). Robustness regions are notated as: RR [$\times 1$, $\times 2$], where $\times 1$ is the smallest SD that supports the conclusion and $\times 2$ is the largest.

Exploratory analysis

The preregistered analysis included expectancy ratings from participants who reported previously having heard of the procedure (but not having taken part in a similar experiment). It was assumed that, as participants in RHI studies are typically psychology undergraduates that this approach would provide a sample representative of participants in contemporary RHI research. Therefore 9 participants who reported not having heard of the procedure before were also analysed separately. To test whether participants who had not heard of the procedure before reported lower expectancies for 'control' than 'illusion' experience statements, t -tests of expectancy differences between 'illusion' and 'control' conditions (details as pre-registered) were also conducted in this sub-sample.

Results

Synchronous condition 'illusion' expectancy ratings ($M = 1.7$, $SD = 0.9$) were greater than synchronous condition 'control' expectancy ratings ($M = 0.0$, $SD = 1.3$,

$t(1,19) = 6.24$, $p < .001$, Cohen's $d = 1.40$, $B_{H(0,1)} = 3.93 \times 10^7$, RR [0.05, 622.40].

Synchronous condition 'illusion' expectancy ratings were also greater than asynchronous 'illusion' expectancy ratings ($M = 0.4$, $SD = 1.3$), $t(19) = 4.68$, $p < .001$, Cohen's $d = 1.05$, $B_{H(0,1)} = 1.14 \times 10^4$, RR [0.07, 404.50].

Consistent with predictions, participants expected to experience the effects described in 'illusion' statements more than those described in 'control' statements.

Figure 1 shows mean scores for 'illusion' (S1–S3) and 'control' (C1–C9) expectancies. The pattern shown for expectancies is characteristic of agreement scores following illusion induction in RHI studies. That is, mean synchronous condition 'illusion' scores indicate agreement (scores of 1 or greater) and mean 'control' statements in either condition indicate either disagreement (scores below -1) or neither agreement nor disagreement (scores of 0) (see **Figure 1** in Botvinick & Cohen, 1998). 95% CIs contain only positive values for the 'illusion' score in the synchronous condition and CIs for the 'control' synchronous condition score and asynchronous condition score contained values consistent with negative and positive scores. Note that, while contrasting sensitive evidence with insensitive evidence (for example, contrasting a significant p value with a non-significant p value) is not informative (Dienes, 2014), this approach is used as a control method in RHI research (e.g., Rohde, di Luca & Ernst, 2011). **Figure 2** shows mean expectancy ratings for each of the three 'illusion' statements and six 'control' statements for both synchronous and asynchronous induction.

Exploratory results

In the nine participants who had not encountered the procedure previously, synchronous condition 'illusion' expectancy ratings ($M = 1.9$, $SD = 0.9$) were greater

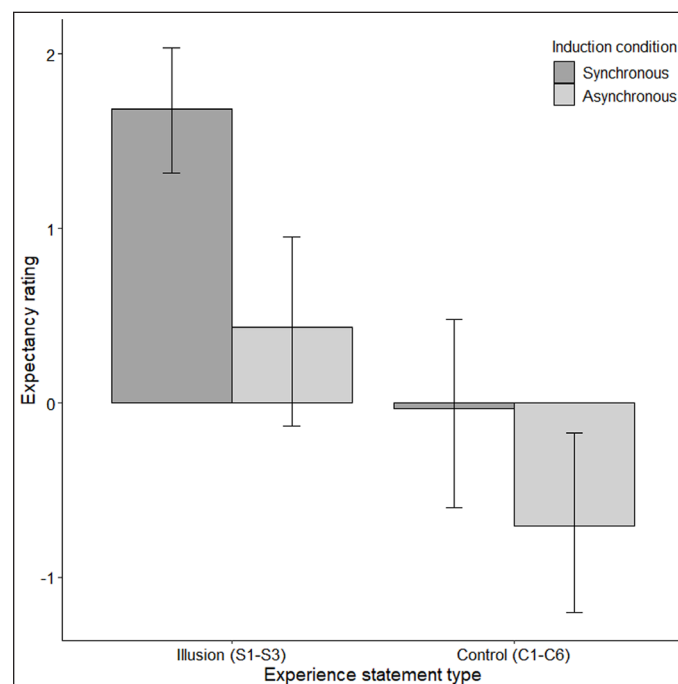


Figure 1: Mean expectancy ratings (-3 to $+3$) for synchronous and asynchronous condition 'illusion' statements (S1–S3) and 'control' statements (C1–C6). Error bars show 95% CIs.

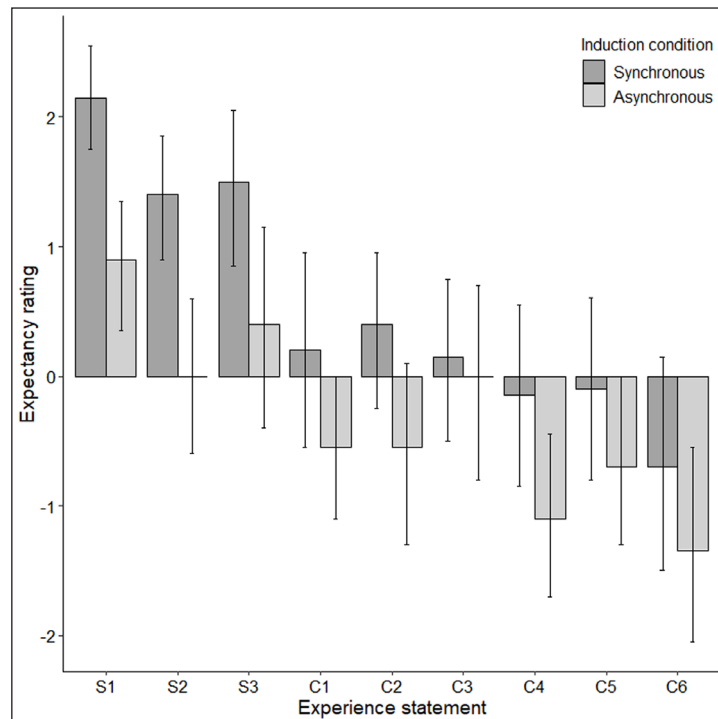


Figure 2: Mean expectancy ratings (−3 to +3) for synchronous and asynchronous condition for all individual statements. Error bars show 95% CIs.

than synchronous condition ‘control’ expectancy ratings ($M = -0.1$, $SD = 1.2$, $t(8) = 6.90$, $p < .001$, Cohen’s $d = 2.30$, $B_{H(0,1)} = 1.48 \times 10^5$, $RR [0.09, 715.40]$.

Synchronous condition ‘illusion’ expectancy ratings were also greater than asynchronous ‘illusion’ expectancy ratings ($M = 0.7$, $SD = 1.0$), $SE = 0.2$, $t(8) = 5.05$, $p = .001$, Cohen’s $d = 1.68$, $B_{H(0,1)} = 444.67$, $RR [0.10, 316.70]$. Therefore, participants who reported having no previous knowledge of the RHI expected the classic RHI pattern of results.

Discussion

As predicted, participants provided with information about the RHI induction procedure reported higher expectancy ratings for ‘illusion’ statements than for ‘control’ statements. Any reported difference between control and illusion measures in published RHI studies may reflect differences in expectancies. RHI control methods are not fit for purpose and cannot provide evidence that the RHI is not a suggestion effect. Because RHI control methods have been adapted to investigate a wide range of related effects (e.g., the full body illusion, enfacement etc), and there has been to date no attempt to control the demand characteristics of these control methods, much contemporary empirical and theoretical work on embodiment is undermined.

Ineffective control methods present considerable problems for the interpretation of illusion reports. Even in the absence of implicit imaginative suggestion effects, differences in ‘illusion’ and ‘control’ response may arise from differing expectancies (e.g., the ‘good participant’ effect; Orne 1962). However, expectancies can drive striking experiential change (e.g., in imaginative suggestion or placebo; Council, Kirsch & Grant, 1996)

and reports of referred touch and ownership of a fake hand in the RHI are likely to reflect, partially or entirely, the control of phenomenology to meet contextually derived expectancies (Lush et al, 2019). Future research attempting to establish whether or not there is a rubber hand illusion beyond demand characteristics and phenomenological control (e.g., driven by multimodal synchrony) will require control methods which account not only for demand characteristics, but also for a potentially confounding characteristic of imaginative suggestion effects – suggestion difficulty.

In hypnotisability research, demand characteristics regarding experimenter expectations are relatively well matched because suggestions are direct and explicit. Despite this, response varies reliably for different types of suggestion (perhaps reflecting differing cognitive requirements; Woody & Barnier, 2008). For example, 77% of participants respond to a suggestion that their hands will be drawn together as though they were magnets, but only around 26% respond to auditory and tactile hallucination of a suggested mosquito (see Lush et al, 2018). Mean subjective report is consequently greater for the moving hands suggestion than for the mosquito suggestion, but this should not be interpreted as evidence there is a ‘real’ magnet effect which is not attributable to suggestion. Once agreement statements are matched in expectancies it will also be necessary to match suggestion difficulty.

Controlling for demand characteristics in measures of experiential change therefore requires matching both expectancies and suggestion difficulty (using statistical tests which can support inferences of no difference, e.g., Bayes factors; see Dienes, 2014). Expectancies can be measured with a simple questionnaire procedure as

described here. Suggestion difficulty could be assessed by response to direct imaginative suggestion in participants matched on trait phenomenological control ability (e.g., hypnosis scales or a phenomenological control scale; Lush et al, in prep) so that the effect of trait differences is minimized. Suggestions should be tested in a task closely matched to the illusion induction procedure, but for which posited mechanisms (e.g., multimodal synchrony) can be ruled out. Higher scores for 'illusion' than 'control' statements matched on both expectancies and on suggestion difficulty would provide compelling evidence of an illusion not attributable to suggestion. If no effect exists when tested by valid control methods, reports of experience of illusory experience of ownership of a fake hand may be attributable to creative, interpretative acts of imagination which are experienced as unintentional (Lush et al, 2019; Dienes et al, in press).

Given an effect size well within the range of visibility to the naked eye (Cohen, 1992), it is remarkable that this issue has been overlooked by psychological scientists during two decades of studies. RHI participants often, unprompted, remark on their experiences with enthusiasm and RHI researchers are likely to have experience of the illusion themselves. This may provide a clue as to why the demand characteristics of methods to control demand characteristics were not critically examined. There appears to be little awareness amongst contemporary researchers that demand characteristics can drive striking changes in experience (perhaps because the bulk of contemporary imaginative suggestion research has focused on the hypnotic context). Researchers may therefore have disregarded the possibility that participants' compelling reports (or their own RHI experiences) could be suggestion effects, and consequently not considered demand characteristics to be a serious threat. Recognition that demand characteristics can drive experience may reveal similar issues in a wide range of paradigms across psychological science. Some readers may be tempted to dismiss the arguments presented here because they have personal experience of the rubber hand illusion and do not believe such a compelling experience could be a suggestion effect. It is worth again noting that imaginative suggestions drive striking changes in experience (including analgesia) and that the ability to control phenomenology to meet expectancies is not rare. As previously stated, around 80% of the population respond to the easiest 'hypnotic' suggestions. It is not safe to rule out expectancies as a cause of one's own experiences. A rigorous approach to controlling demand characteristics and suggestion effects in experimental manipulations of participant experience is necessary even when a researcher has personal experience of a particular effect.

There are many parallels between imaginative suggestion effects and the RHI (Lush et al, 2019). Standard measurement statements support a wide range of possible interpretations (Wu, in press) and there is great variety in unconstrained reports (Valenzuela Moguillandsky, O'Regan & Petitmengin, 2013). RHI report may reflect creative and interpretative acts of imagination experienced

as unintentional 'happenings' rather than the disruption of multisensory embodiment mechanisms (Dienes et al, 2019; see also Alsmith, 2015).

Some may discount the possibility that the RHI could be entirely a suggestion effect because of evidence from 'implicit' measures. It is therefore worth considering these measures in some detail. There are two major issues. The first is that these measures may also be suggestion effects confounded by demand characteristics. For example, it is not clear why proprioceptive drift is considered resistant to demand characteristics. Indeed, the measure bears striking similarity to a standard measure of response to a hypnotic 'magnetic hands' suggestion which asks participants how far they felt their hands moved (see Bowers, 1998), and to my knowledge the demand characteristics have never been tested. It is also not clear why RHI researchers consider skin conductance to be resistant to suggestion effects. Claims in influential RHI papers that participants cannot voluntarily control SCRs are, on inspection, assertions not backed by reference to evidence (e.g. Armel and Ramachandran; 2003; Ma & Hommel, 2013). There is in fact considerable evidence for top-down effects on skin conductance measures. For example, in a highly cited paper (1566 citations, Google Scholar, 03/03/2020), Levenson, Ekman & Friesen (1990) report voluntary changes in skin conductance relating to changes in facial expression. Furthermore, imaginative suggestion has been repeatedly demonstrated over more than half a century to affect electrodermal activity (e.g., Barber & Coules, 1959; Kekecs; Szekely & Varga, 2016). Note also that imaginative suggestion can drive changes in fMRI measures and histamine reactivity (see Lush et al, 2019). Demand characteristics need to be carefully controlled in 'implicit' measures, and any claims that a particular measure is resistant to voluntary control or suggestion effects in the RHI must be backed by evidence, not merely asserted. The second issue is that 'implicit' measures are considered to reflect illusion experience because they correlate with subjective report of illusion experience. If reports of illusion experience are driven by demand characteristics, then any problems this causes for interpretation of subjective report extends to 'implicit' measures. Without valid subjective reports of ownership experience, proprioceptive drift is a measure of confusion about where one's hand is, skin conductance is a proxy measure of emotional arousal, fMRI is a proxy measure of blood-flow, and crossmodal congruency (Zopf et al, 2013) is a measure of reaction time and accuracy. None of these measures are directly informative about experiences of ownership. One might argue that implicit measures appear to converge on 'ownership experience'. Given that implicit measures have often been developed with ownership in mind, this should not be surprising, even if the illusion is entirely a suggestion effect. Note that the discovery of substantial relationships between response to imaginative suggestion and the RHI considerably extends the range of effects with which RHI illusion reports would be likely to correlate, if experimental demand characteristics suggested them (e.g., paralysis, amnesia, hallucinations, involuntary movement, analgesia, etc;

see Lush et al, 2018 for SWASH hypnotisability scale suggestions).

An anonymous reviewer raised the possibility that participants in an RHI study would not necessarily identify the conditions as involving either synchronous and asynchronous brush stroking, and so the task demand demands of the present study and a typical RHI procedure may differ. The reviewer suggested this issue could be resolved by asking participants to describe the difference between conditions after they had participated in an RHI procedure. We followed this suggested procedure and found that all seven participants were able to identify this difference (see supplemental material).

These results do not provide evidence that the RHI is a suggestion effect, but rather that existing claims to the contrary are invalid. It has long been known that implicit suggestion effects can drive apparently involuntary experience. Perhaps the best known such effect, mesmerism, was once believed to involve the perturbation of a 'magnetic fluid' as a magnetic rod or the mesmerist's hands were moved at a short distance from the subject's body (Pintar & Lynn, 2009). RHI procedures are strikingly similar to mesmeric induction, and in some cases virtually identical (e.g., the magnetic touch illusion, in which brushes are moved at a short distance from the subject's body; Guterstam, Zeberg, Özçiftci, & Ehrsson, 2016). Magnetic fluid was a scientifically plausible explanation for imaginative suggestion effects in the 18th century. Without effective controls for suggestion effects, claims that embodiment illusions are driven by multisensory integration mechanisms are on no firmer ground than the claims that mesmeric convulsions were induced by the manipulation of magnetic fluid.

This study provides evidence that established measures to control demand characteristics in the RHI used in hundreds of studies over the last 22 years are confounded by demand characteristics. This investigation of demand characteristics in the illusion was motivated by the discovery of substantial relationships between hypnotisability and RHI measures (Lush et al, 2019). A parsimonious explanation for these relationships, given that demand characteristics have not been controlled in the RHI is that the RHI is at least partially an implicit imaginative suggestion effect which is driven by expectancies arising from demand characteristics. Valid control methods must be developed and applied to establish whether or not there is a rubber hand illusion beyond suggestion effects. Establishing evidence to support existing theories of the RHI which attribute illusion reports to something other than demand characteristics (including those which appeal to top down processes, e.g., Longo et al, 2008) will require effective controls for demand characteristics.

Demand characteristics are not controlled in the RHI. This is particularly problematic because the RHI and other embodiment effects (mirror touch and pain experience) are likely to be at least partly driven by the control of phenomenology to meet expectancies (Lush et al, 2019). It is plausible that 'implicit' measures of the RHI (e.g., skin conductance and proprioceptive drift) are similarly confounded by demand characteristics. Further

research will be required to test this and measures used in related illusions. This issue is unlikely to be limited to embodiment research, however. The extent to which phenomenological control confounds behavioural science is currently unknown, but may be substantial. In recent years, psychological science has been shaken by problems arising from poor methodology, with much of the focus on the 'replication crisis' (see Chambers, 2019). However, attention is now turning to the problem of generalisability (see Yarkoni, 2019). Demand characteristics do not appear to have been taken seriously in recent years (Sharpe & Whelton, 2016), perhaps because of a lack of awareness that they can drive experience and not merely compliance. If demand characteristics have been driving experience in other measures (and given the wide range of imaginative suggestion effects, this is not unlikely), psychology will be faced with a crisis of generalisability. This prompts a reconsideration of demand characteristics in measures of experiential change across psychological science.

Data Accessibility Statement

This study was preregistered; the preregistration document can be accessed at <https://osf.io/89m7j>. An additional analysis which was not pre-registered is clearly identified as exploratory in the manuscript. De-identified data and code-book, data analysis (JASP) output, JASP file and study materials are available at <https://osf.io/9c8mq/>.

Acknowledgements

Thanks to Zoltan Dienes for discussion, to Julie McDermott, Maxine Sherman, Warrick Roseboom and Marie-Luise Schreiter for helpful feedback on the draft manuscript and to Marta Stojanovic for supplementary data collection.

Funding Information

This work was supported by the Dr Mortimer and Theresa Sackler Foundation.

Competing Interests

The author has no competing interests to declare.

Author Contributions

P. Lush is the sole author of this article and is responsible for its content.

References

- Abdulkarim, Z., & Ehrsson, H. H.** (2016). No causal link between changes in hand position sense and feeling of limb ownership in the rubber hand illusion. *Attention, Perception, & Psychophysics*, *78*(2), 707–720. DOI: <https://doi.org/10.3758/s13414-015-1016-0>
- Alsmith, A.** (2015). Mental activity & the sense of ownership. *Review of Philosophy and Psychology*, *6*(4), 881–896. DOI: <https://doi.org/10.1007/s13164-014-0208-1>
- Armell, K. C., & Ramachandran, V. S.** (2003). Projecting sensations to external objects: evidence from skin conductance response. *Proceedings of the Royal Society of London. Series B: Biological Sciences*,

- 270(1523), 1499–1506. DOI: <https://doi.org/10.1098/rspb.2003.2364>
- Barber, T. X., & Coules, J.** (1959). Electrical skin conductance and galvanic skin response during “hypnosis”. *International Journal of Clinical and Experimental Hypnosis*, 7(2), 79–92. DOI: <https://doi.org/10.1080/00207145908415811>
- Botvinick, M., & Cohen, J.** (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669), 756. DOI: <https://doi.org/10.1038/35784>
- Bowers, K. S.** (1998). Waterloo-stanford group scale of hypnotic susceptibility, form C: Manual and response booklet. *International Journal of Clinical and Experimental Hypnosis*, 46(3), 250–268. DOI: <https://doi.org/10.1080/00207149808410006>
- Braffman, W., & Kirsch, I.** (1999). Imaginative suggestibility and hypnotizability: an empirical analysis. *Journal of Personality and Social Psychology*, 77(3), 578. DOI: <https://doi.org/10.1037/0022-3514.77.3.578>
- Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H. H., & Philippen, A.** (2018). The senses of agency and ownership: a review. *Frontiers in psychology*, 9, 535. DOI: <https://doi.org/10.3389/fpsyg.2018.00535>
- Chambers, C.** (2019). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. New Jersey: Princeton University Press. DOI: <https://doi.org/10.1515/9780691192031>
- Cohen, J.** (1992). A power primer. *Psychological bulletin*, 112(1), 155. DOI: <https://doi.org/10.1037/0033-2909.112.1.155>
- Council, J. R., Kirsch, I., & Grant, D. L.** (1996). Imagination, expectancy, and hypnotic responding. *Hypnosis and imagination*, 41–65. DOI: <https://doi.org/10.4324/9781315224374-3>
- Dienes, Z.** (2014). Using Bayes to get the most out of non-significant results. *Frontiers in psychology*, 5, 781. DOI: <https://doi.org/10.3389/fpsyg.2014.00781>
- Dienes, Z., Lush, P., Palfi, B., Roseboom, W., Scott, R., Seth, A., & Lovell, M.** (in press). Phenomenological control as cold control. *Psychology of Consciousness*.
- Ehrsson, H. H., Holmes, N. P., & Passingham, R. E.** (2005). Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *Journal of Neuroscience*, 25(45), 10564–10573. DOI: <https://doi.org/10.1523/JNEUROSCI.0800-05.2005>
- Guterstam, A., Zeberg, H., Özçiftci, V. M., & Ehrsson, H. H.** (2016). The magnetic touch illusion: A perceptual correlate of visuo-tactile integration in peripersonal space. *Cognition*, 155, 44–56. DOI: <https://doi.org/10.1016/j.cognition.2016.06.004>
- Haans, A., Kaiser, F. G., Bouwhuis, D. G., & IJsselstein, W. A.** (2012). Individual differences in the rubber-hand illusion: Predicting self-reports of people’s personal experiences. *Acta psychologica*, 141(2), 169–177. DOI: <https://doi.org/10.1016/j.actpsy.2012.07.016>
- Heaton, R. K.** (1975). Subject expectancy and environmental factors as determinants of psychedelic flashback experiences. *The Journal of nervous and mental disease*, 161(3), 157–165. DOI: <https://doi.org/10.1097/00005053-197509000-00002>
- Juhasz, J. B., & Sarbin, T. R.** (1966). On the false alarm metaphor in psychophysics. *The Psychological Record*, 16(3), 323–327. DOI: <https://doi.org/10.1007/BF03393675>
- Kalckert, A., & Ehrsson, H. H.** (2012). Moving a rubber hand that feels like your own: a dissociation of ownership and agency. *Frontiers in human neuroscience*, 6, 40. DOI: <https://doi.org/10.3389/fnhum.2012.00040>
- Kekecs, Z., Szekely, A., & Varga, K.** (2016). Alterations in electrodermal activity and cardiac parasympathetic tone during hypnosis. *Psychophysiology*, 53(2), 268–277. DOI: <https://doi.org/10.1111/psyp.12570>
- Kirsch, I., & Council, J. R.** (1989). Response expectancy as a determinant of hypnotic behavior. In N. P. Spanos & J. F. Chaves (Eds.), *Hypnosis: The cognitive-behavioral perspective* (pp. 360–379). Buffalo, NY: Prometheus Press.
- Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O.** (2007). Video ergo sum: manipulating bodily self-consciousness. *Science*, 317(5841), 1096–1099. DOI: <https://doi.org/10.1126/science.1143439>
- Levenson, R. W., Ekman, P., & Friesen, W. V.** (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, 27(4), 363–384. DOI: <https://doi.org/10.1111/j.1469-8986.1990.tb02330.x>
- Longo, M. R., Schüür, F., Kammers, M. P., Tsakiris, M., & Haggard, P.** (2008). What is embodiment? A psychometric approach. *Cognition*, 107(3), 978–998. DOI: <https://doi.org/10.1016/j.cognition.2007.12.004>
- Lush, P., Botan, V., Scott, R. B., Seth, A., Ward, J., & Dienes, Z.** (2019, April 16). Phenomenological control: response to imaginative suggestion predicts measures of mirror touch synaesthesia, vicarious pain and the rubber hand illusion. *PsyArxiv preprint*. DOI: <https://doi.org/10.31234/osf.io/82jav>
- Lush, P., Moga, G., McLatchie, N., & Dienes, Z.** (2018). The Sussex-Waterloo Scale of Hypnotizability (SWASH): measuring capacity for altering conscious experience. *Neuroscience of consciousness*, 2018(1), niy006. DOI: <https://doi.org/10.1093/nc/niy006>
- Ma, K., & Hommel, B.** (2013). The virtual-hand illusion: effects of impact and threat on perceived ownership and affective resonance. *Frontiers in psychology*, 4, 604. DOI: <https://doi.org/10.3389/fpsyg.2013.00604>
- Michel, C., Velasco, C., Salgado-Montejo, A., & Spence, C.** (2014). The butcher’s tongue illusion. *Perception*, 43(8), 818–824. DOI: <https://doi.org/10.1068/p7733>
- Orne, M. T.** (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American psychologist*, 17(11), 776. DOI: <https://doi.org/10.1037/h0043424>
- Orne, M. T.** (1969). Demand characteristics and the concept of quasi-controls. In R. Rosenthal & R. Rosnow (Eds.), *Artifact in behavioral research* (pp. 143–179). New York: Academic Press.

- Pintar, J., & Lynn, S. J.** (2009). Hypnosis: A brief history. John Wiley & Sons. DOI: <https://doi.org/10.1002/9781444305296>
- Riemer, M., Trojan, J., Beauchamp, M., & Fuchs, X.** (2019). The rubber hand universe: On the impact of methodological differences in the rubber hand illusion. *Neuroscience and biobehavioral reviews*, 104. DOI: <https://doi.org/10.1016/j.neubiorev.2019.07.008>
- Rohde, M., Di Luca, M., & Ernst, M. O.** (2011). The rubber hand illusion: feeling of ownership and proprioceptive drift do not go hand in hand. *PLoS one*, 6(6), e21659. DOI: <https://doi.org/10.1371/journal.pone.0021659>
- Sforza, A., Bufalari, I., Haggard, P., & Aglioti, S. M.** (2010). My face in yours: Visuo-tactile facial stimulation influences sense of identity. *Social neuroscience*, 5(2), 148–162. DOI: <https://doi.org/10.1080/17470910903205503>
- Sharpe, D., & Whelton, W. J.** (2016). Frightened by an old scarecrow: The remarkable resilience of demand characteristics. *Review of General Psychology*, 20(4), 349–368. DOI: <https://doi.org/10.1037/gpr0000087>
- Valenzuela Moguillansky, C., O'Regan, J. K., & Petitmengin, C.** (2013). Exploring the subjective experience of the “rubber hand” illusion. *Frontiers in Human Neuroscience*, 7, 659. DOI: <https://doi.org/10.3389/fnhum.2013.00659>
- Woody, E. Z., & Barnier, A. J.** (2008). Hypnosis scales for the twenty-first century: What do we need and how should we use them. *The Oxford handbook of hypnosis: Theory, research, and practice*, 255–282. DOI: <https://doi.org/10.1093/oxfordhb/9780198570097.013.0010>
- Wu, W.** (forthcoming). Mineness and introspective data. In M. Guillot & M. Garcia Carpintero (Eds.), *The Sense of Mineness*. New York; Oxford: Oxford University Press.
- Yarkoni, T.** (2019, November 22). The Generalizability Crisis. DOI: <https://doi.org/10.31234/osf.io/jqw35>
- 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. *JoVE (Journal of Visualized Experiments)*, 77, e50530. DOI: <https://doi.org/10.3791/50530>

Peer review comments

The author(s) of this paper chose the Open Review option, and the peer review comments can be downloaded at: <http://doi.org/10.1525/collabra.325.pr>

How to cite this article: Lush, P. (2020). Demand Characteristics Confound the Rubber Hand Illusion. *Collabra: Psychology*, 6(1): 22. DOI: <https://doi.org/10.1525/collabra.325>

Senior Editor: Simine Vazire

Editor: Alex Holcombe

Submitted: 15 January 2020

Accepted: 10 March 2020

Published: 07 April 2020

Copyright: © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.