Computer versus live delivery of the Sussex Waterloo Scale of Hypnotizability (SWASH)

Article (Accepted Version)

Computer vs live delivery of the Sussex Waterloo Scale of Hypnotisability (SWASH).

Lush, P.* [1,2], Scott, R. B., [1,3], Moga, G.[3] & Dienes, Z. [1, 3].

[1] Sackler Centre for Consciousness Science, University of Sussex, Falmer, BN1 9RH, UK

[2] Department of Informatics, Chichester Building, University of Sussex, Falmer, BN1 9RH, UK

[3] Department of Psychology, Pevensey Building, University of Sussex, Falmer, BN1 9RH, UK

Data and materials available at https://osf.io/3wg46/

*Corresponding author

Abstract

There exist reliable and stable trait differences in the ability to control phenomenology in response to imaginative suggestion. Hypnotisability scales measure response to imaginative suggestion within a hypnotic context. Because hypnotisability has recently been shown to predict measures of experiential change in psychological experiments (e.g., the rubber hand illusion), there is a need for easy to use screening tools which are accessible to researchers with little or no background in hypnosis or imaginative suggestion research. The SWASH is a time efficient group hypnotisability scale which can be administered to up to 50 participants simultaneously. Here we present norms from an undergraduate sample for a recorded version delivered by a computer program alongside norms for a live presentation. Reliability, validity and mean scores are similar across the two presentations. Computer delivery of a pre-recorded script provides a simple tool to rapidly screen for hypnotisability in large groups for researchers with no prior experience of hypnosis research.

Introduction

Hypnotisability scales measure trait differences in the ability to control phenomenology to meet expectancies within a hypnotic context. Hypnotisability is measured by standardised screening procedures in which a 'hypnotic induction' script is delivered followed by a series of imaginative suggestions (Woody & Barnier, 2008). Hypnotisability scores are then calculated for each participant by summing response to each suggestion. Scores on hypnotisability scales have been shown to be stable over a 25 year period (Piccione, Hilgard & Zimbardo, 1989) and are typically used to identify participants of high, low or medium hypnotisability for subsequent experiments (for a recent review see Acunzo & Terhune, 2019). Because imaginative suggestion in the hypnotic context reliably produces changes in experience, participants pre-screened on hypnotisability scales are frequently recruited for experiments in consciousness science. For example, recent studies have employed hypnotic suggestion to investigate changes in the sense of agency (Lush et al, 2017; Polito, Barnier & Woody, 2013), and as models of delusions (Cox & Barnier, 2009). Notably, it has long been known that trait differences in response to imaginative suggestion in a hypnotic context strongly predict response outside the hypnotic context (Hull, 1933, Weitzenhoffer and Sjoberg 1961, Braffman & Kirsch, 1999, Fassler, Lynn & Knox, 2008). Recently, this has been shown to extend to measures of subjective experience previously thought to be unrelated to imaginative suggestion: hypnotisability scores have been shown to predict measures of experiential change in embodiment paradigms (e.g., the rubber hand illusion) to a degree comparable to individual hypnotic suggestions, and trait differences in phenomenological control may confound many measures of experience (Lush et al, 2020; Dienes et al, 2019). There is therefore a need for efficient and accessible screening tools

which will enable researchers without prior experience of hypnosis or imaginative suggestion research to identify and control for trait differences in phenomenological control abilities.

The Sussex-Waterloo Scale of Hypnotisability (SWASH; Lush et al, 2018) is an adaptation of the Waterloo-Stanford Group Scale of Hypnotic Susceptibility, form c (Bowers, 1993) with modifications to improve time efficiency (reducing delivery time by around 50%) and increased sample size, which for our group have enabled the screening of large participant samples in around 10% of the time required for WSGC screening. Note that if available samples are smaller than 50, the decrease in required time will be smaller than in this best case scenario. The SWASH produces similar scores to the WSGC and has good reliability. Here we present norms for a fully automated computer delivery of SWASH. Recorded imaginative suggestion and hypnotisability procedures have been shown to be similarly effective as live induction (Barber & Calverley, 1964a; Fassler, Lynn & Knox, 2008). Recorded delivery is common in norming studies (see e.g., Oakley et al, 2020) and allows greater consistency across testing sessions because it eliminates possible effects of differences in delivery style across experimenters (e.g. tone of voice; Barber & Calverley, 1964b). The use of a recorded delivery also allows screening to be fully automated and therefore performed without the presence of an experimenter designated as a ‘hypnotist’. This increases the accessibility of hypnotisability screening to researchers with little or no background in hypnosis research. Finally, the collection of data by computer increases efficiency by removing the need for the time-consuming data entry which follows traditional pen and paper data collection.

There were two adaptations necessary to convert the SWASH to an automated computer delivery. First, a picture of three coloured balls for the negative hallucination suggestion is displayed on the computer screen at a short distance from the participant rather than projected onto a large screen in a lecture theatre. This adaptation was necessary in order

to perform screenings using just a computer. Second, the post-hypnotic suggestion (in which participants are told they will respond to a suggestion after the session, but will not remember the suggestion) was changed from drawing a tree in the corner of the SWASH response booklet to pressing the computer keyboard space bar six times. t was necessary to replace this SWASH suggestion because participants would not have access to pen and paper and the report needed to be easily verifiable by computer. We did not have evidence that these suggestions were equivalent to the suggestions in the SWASH before conducting this study.

We tested participants on the Sussex-Waterloo Scale of Hypnotisability following either live or pre-recorded delivery of induction and imaginative suggestions. If response is comparable across the two delivery methods, then an automated, computer-based delivery of the SWASH would offer advantages in consistency, ease of use and time efficiency.

Method

Participants

143 participants (121 female, 22 male, Mean age = 19.1, SD = 2.4) completed the SWASH screening by recorded delivery of the script and 146 (124 female, 22 male, Mean age = 18.9, SD = 2.7) by live delivery. All participants performed the task as part of a practical on an undergraduate psychology course at the University of Sussex. Participants were run in groups of up to 50 people; each group was randomly assigned to condition, but participants within groups were determined by their practical assignment. Ethical approval was received from the University of Sussex Sciences & Technology Cross-Schools Research Ethics Committee (ER/RBS20/6) and informed consent was obtained.

Materials

The screening program was created in Matlab (Mathworks, 2017). Participants reported subjective and objective response (see Lush et al, 2018 for details). Subjective response was recorded on a scale from 0 to 5 by participants pressing number keys on the computer keyboard. See the materials at https://osf.io/3wg46/ for individual scale labels and response prompts provided to participants. Because the post-hypnotic suggestion was changed, the response prompt for this item in the previous response booklet (Lush et al, 2018) was changed to the following:

”You were told that you would press the space bar six times in a row, but that you would forget that you were told to do so.”

Two subjective responses were collected for this item, the first regarding urge (from no urge to clear urge) and the other regarding memory of the instruction given (from no memory of instruction to normal memory of instruction).

Procedure

The procedure was delivered in lab practical sessions run as part of the 1st year undergraduate course "Cognition in Clinical Contexts" at the University of Sussex. Participants were screened in groups of up to 50. In both the recorded and live conditions participants were given paper instructions explaining the steps necessary to start the computer program that would record their self-ratings and, in the case of the recorded condition, would also present the recorded induction and suggestions. The program first presented participant information and consent instructions together with the option to choose to provide contact details for inclusion in a participant recruitment database (materials available at https://osf.io/3wg46/). The experimenter talked through these screens answering any questions participants had about participation and emphasising their right to withdraw at any time. They were then prompted to enter standard demographic information. Participants were instructed to pause at this point so that all of them started the main experiment at the same time. The instructions differed at this point depending on the condition being run.

In the recorded condition they were instructed to put on the headphones provided and to adjust the volume to a comfortable level using a reference tone. These participants then listened to a pre-recorded introductory statement, hypnotic induction and a series of 10 imaginative suggestions. In the live delivery group, participants listened to the same experimenter (G. Moga M.D.) reading the same introductory statement, induction and imaginative suggestions in person. Following delivery of the script participants in both conditions reported their experience by entering ratings on the computer keyboard. Therefore,

the conditions differed only in whether the induction and suggestions were presented pre-recorded or in person.

Analyses

Data and analysis scripts are available at https://osf.io/3wg46/. Mean objective and subjective scores were calculated. Due to a programming error, objective scale scores for the negative visual hallucination item were incorrectly recorded. Objective scale analyses are therefore presented for nine items only.

Scoring was calculated as in the SWASH (Lush et al, 2018). Objective scale items were scored as pass or fail according to the criteria for each item. For two suggestions there were two subjective responses requested. Subjective scale scores between 0 and 5 were calculated from the average of subjective scale responses. The final subjective response score for taste is the mean of the sweet and sour responses. For the post-hypnotic suggestion, the geometric mean of the urge and amnesia responses for the item was calculated, so that a subjective response for this item would be zero if either of the components of the suggestion did not generate a subjective response.

Scale scores for recorded and live delivery are reported for comparison. Scale validity was assessed using point-biserial correlations between subjective and objective responses. Correlations were interpreted according to Cohen (1988). Scale reliability was assessed using Cronbach's alpha, alpha with item-dropped and corrected same scale item-total correlations. Alphas were interpreted according to recommendations given in Field (2013). Note that omega values (Dunn, Baguley & Brunsden, 2014) were almost identical to the alpha values given here. 95% Confidence Intervals are reported throughout, which can be interpreted as 95% Credibility Intervals with uniform priors. The credibility interval can be interpreted as the plausible range of population values.

Results

Scores

For the 143 participants who completed the screening following a recorded induction delivery, mean total score on the subjective scale was 1.6 (SD = .84) out of a maximum of five and the total objective score was 3.7 (SD = 1.8) out of a maximum of nine. For the 146 participants for whom the induction was delivered live, the mean total subjective scale score was 1.5 (SD = .78) and the total objective scale score was 4.0 (SD = 1.8). Table 1a shows mean subjective scores for each suggestion for live and recorded delivery and table 1b shows percentage pass rate for each suggestion. Scores and pass rates were for each item numerically similar or greater for recorded delivery compared to live delivery.

	Live delivery		Recorded delivery		Difference
Suggestion	*M*	*SD*	*M*	*SD*	*M* [95% CI]
1. Hand lowering	3.4	1.4	3.3	1.5	0.1 [-0.2, 0.4]
2. Moving hands together	2.7	1.5	2.8	1.5	-0.1 [-0.4, 0.3]
3. Mosquito hallucination	0.8	1.1	0.9	1.2	-0.1 [-0.3, 0.2]
4. Taste hallucination	0.9	1.1	1.1	1.2	-0.2 [-0.4, 0 .1]
5. Arm rigidity	2.4	1.6	2.4	1.5	0.0 [-0.3, 0.4]
6. Arm immobilisation	2.2	1.6	2.2	1.5	-0.0 [-0.4, 0.3]
7. Music hallucination	0.2	0.6	0.2	0.6	0.0 [-0.1, 0.2]
8. Negative visual hallucination	0.2	0.7	0.2	0.9	-0.0 [-0.2, 0.2]
9. Amnesia	1.7	1.4	1.6	1.4	0.2 [-0.2, 0.5]
10. Posthypnotic suggestion	0.9	1.4	1.0	1.5	-0.1 [-0.5, 0.2]

Table 1a. Mean subjective scores for each suggestion for live and recorded SWASH delivery.

	Live delivery		Recorded delivery		Difference
Suggestion	*M*	*SD*	*M*	*SD*	*M* [95% CI]
1. Hand lowering	.72	.45	.69	.46	.01 [-.08, .13]
2. Moving hands together	.74	.44	.76	.43	-.02 [-.12, .09]
3. Mosquito hallucination	.19	.40	.31	.47	-.12 [-.22, -.023]
4. Taste hallucination	.48	.50	.51	.50	-.03 [-.15, .09]
5. Arm rigidity	.50	.50	.47	.50	.03 [-.09, .15]
6. Arm immobilisation	.41	.49	.38	.49	.03 [-.08, .15]
7. Music hallucination	.03	.18	.05	.22	-.02 [-.06, .03]
8. Negative visual hallucination	N/A		N/A		N/A
9. Amnesia	.18	.38	.20	.40	-.03 [-.12, .07]
10. Posthypnotic suggestion	.41	.49	.66	.47	-.25 [-.37, -1.14]

Table 1b. Mean response on the objective criterion in live and recorded delivery conditions.

Validity

The correlation between total subjective scale score (ten items) and total objective scale score (nine items) for recorded delivery *was* $r(143) = .67$, 95% CI [.57, .75] and for live delivery, $r(146) = .72$, 95% CI [.63, .79]. Table 2 shows mean subjective score and point biserial correlations between objective and subjective responses for each item. Correlation coefficients between objective and subjective responses in the live condition (mean coefficient of .50) and recorded condition (mean coefficient of .53) were large, except for the post-hypnotic suggestion item. The subjective response for this item is the geometric mean of the subjective ratings of an urge to press the space bar 6 times as instructed and of subjective ratings of amnesia for the instructions to press the space bar. The correlation between ratings

of the urge to complete the task and objective response in the PHS was small to medium for recorded delivery, $r(143) = .34$, 95% CI [.18, .48], and live delivery, $r(146) = .26$, 95% CI [.10, .40]. However, the plausible range of the correlation between amnesia for the suggestion and objective response to the PHS was negative for the recorded delivery, $r(143) = -.12$, 95% CI [-.28, .041] and also for the live delivery, $r(146) = -.22$, 95% CI [-.06, .-.37]. The plausible range of the correlation between the two subjective measures of the PHS was small or negative for both recorded delivery, $r(143)= -.075$, 95% CI [-.24, .091], and live delivery, $r(146) = -.19$, 95% CI [-.34, -.029].

Suggestion	Live delivery *r*pb	Recorded delivery *r*pb
1. Hand lowering	.57 [.45, .67]	.59 [.47, .69]
2. Moving hands together	.51 [.38, .62]	.38 [.23, .51]
3. Mosquito hallucination	.59, [.47, .69]	.60 [.48, .69]
4. Taste hallucination	.37 [.22, .50]	.49 [.36, .61]
5. Arm rigidity	.58 [.46, .68]	.55 [.43, .66]
6. Arm immobilisation	.48 [.34, .60]	.55 [.42, .65]
7. Music hallucination	.61 [.49,.70]	.70 [.61, .78]
8. Negative visual hallucination	N/A	N/A
9. Amnesia	.26 [.11, .41]	.25 [.092, .40]
10. Posthypnotic suggestion	.030 [-.13, .19]	.12 [-.044, .28]

Table 2. Point biserial correlations between behavioural and experiential scoring of suggestions for live and recorded delivery

Reliability

For the recorded induction, subjective scale alpha was .84, 95% CI [.80, .87], indicating good internal consistency. Live delivery subjective scale alpha was .81, 95% CI [.76, .85], also indicating good consistency.

Table 3a shows Cronbach's alpha if the item is dropped for each SWASH suggestion on the subjective scale. Point estimates of the coefficient were equal or similar to overall alpha in each condition, indicating that both scales are reliable.

	Live delivery	Recorded delivery
1. Hand lowering	.79 [.72, .83]	.81 [.76, .85]
2. Moving hands together	.80 [.74, .84]	.82 [.77, .85]
3. Mosquito hallucination	.80 [.74, .84]	.83 [.79, .86]
4. Taste hallucination	.79, [.74, .84]	.82, [.77,.85]
5. Arm rigidity	.78 [.72, .83]	.81 [.76, .84]
6. Arm immobilisation	.77 [.70, .82]	.81 [.75, .84]
7. Music hallucination	.82 [.78, .86]	.84 [.80, .87]
8. Negative visual hallucination	.81 [.76, .85]	.84 [.80, .87]
9. Amnesia	.80 [.75, .84]	.82 [.77, .85]
10. Posthypnotic suggestion	.81 [.76, .85]	.83 [.79, .86]

Table 3a. Subjective scale alpha (if item dropped) (95% CI)

Alpha for the objective scale was .54, 95% CI [.41, .63], suggesting that internal consistency was not good for this scale. Table 3b shows alpha coefficient if the item is dropped for objective scale response to each SWASH suggestion.

	Live delivery	Recorded delivery
1. Hand lowering	.50 [.35, .61]	.48 [.33,.58]
2. Moving hands together	.50 [.33, .61]	.52 [.39, .62]
3. Mosquito hallucination	.51 [.36, .62]	.49 [.33,.59]
4. Taste hallucination	.49 [.34, .60]	.51 [.37,.61]
5. Arm rigidity	.48 [.31, .59]	.46 [.27, .58]
6. Arm immobilisation	.45 [.26, .57]	.46 [.29,.58]
7. Music hallucination	.53 [.38, .63]	.52 [.38,.62]
8. Negative visual hallucination	N/A	N/A
9. Amnesia	.52 [.37, .632	.52 [.38, .62]
10. Posthypnotic suggestion	.60 [.49, .69]	.57 [.45,.66]

Table 3b. Alpha (if item dropped) (95% CI in brackets)

Table 4a shows each subjective scale item correlated with the corrected total scale score (corrected by removal of that item). Mean item-total correlations were medium to large and mean item-total correlations were comparable between live delivery (r = .49) and recorded delivery (r = .53) delivery. Table 4b shows each subjective scale item correlated with the corrected total scale score. Item-total correlations were not strong for the objective scale.

	Live delivery	Recorded delivery
1. Hand lowering	.60 [.48, .69]	.66 [.55, .74]
2. Moving hands together	.52 [.38, .63]	.59 [.47, .69]
3. Mosquito hallucination	.50 [.37, .61]	.45 [.31, .57]
4. Taste hallucination	.53 [.40. 63]	.58 [.46, .68]
5. Arm rigidity	.63 [.52. 72]	.68 [.58, .76]
6. Arm immobilisation	.72 [.63. 79]	.68 [.58, .76]
7. Music hallucination	.20 [.04, .35]	.32 [.16, .46]
8. Negative visual hallucination	.35 [.20, .48)	.27 [.11, .42]
9. Amnesia	.48 [.35, .60)	.57 [.45, .67]
10. Posthypnotic suggestion	.38 [.24, .51)	.45 [.31, .57]

Table 4a. Subjective scale corrected same-scale item-total correlations (95% CI)

	Live delivery	Recorded delivery
1. Hand lowering	.28 [.12, .42]	.32 [.17, .46]
2. Moving hands together	.30 [.14, .44]	.18 [.01, .33]
3. Mosquito hallucination	.25 [.09, .39]	.30 [.14, .44]
4. Taste hallucination	.31 [.15, .45]	.23 [.06, .38]
5. Arm rigidity	.34 [.18, .47]	.36 [.21, .49]
6. Arm immobilisation	.41 [.26, .53]	.37 [.22, .51]
7. Music hallucination	.23 [.07, .38]	.20 [.04, .35]
8. Negative visual hallucination	N/A	N/A
9. Amnesia	.22 [.05, .36]	.20 [.04, .35]

10. Posthypnotic suggestion	-.03 [-.20, .13]	.032 [-.13, .20]

Table 4b. Objective scale corrected same-scale item-total correlations (95% CI)

.

Discussion

We collected self-reports of response to suggestion in live presentation and recorded presentation of the Sussex Waterloo Scale of Hypnotisability (Lush et al, 2018). Mean subjective report scores were similar and both scales showed comparable internal validity across both methods. While the SWASH consists of two scales, an objective scale (adapted from the WSGC, Bowers, 1993) and a subjective scale, we do not have complete data for the objective scale for this study. However, for the nine items for which we have complete data, objective scale pass rates were in all cases similar or numerically greater for recorded delivery compared to live delivery. Because we do not have objective scale data for the negative visual hallucination item, more research will be required to provide information regarding objective scores for this item and any changes this would make to objective scale results.

Post hypnotic suggestion to press the space bar six times produced similar subjective ratings and similarly weak correlations between objective and subjective scale post-hypnotic suggestion scores to the post-hypnotic suggestion to draw a tree in previous live delivery (Lush et al, 2018). The weak correlations may reflect the inadequacy of behavioural scoring for a suggestion which requires both amnesia (participants must not remember the suggestion) and some felt compulsion to perform the action. For example, participants may have full memory of the suggestion, but perform the suggested action. Such responses would generate a positive objective score, but a score of zero on the subjective scale. We have previously noted this issue with objective scoring of post-hypnotic suggestion items (Lush et al, 2018) and agree with Sadler & Woody (2004) that objective scoring has over-estimated post-hypnotic suggestion pass rates in the hypnosis literature. This points to a more general issue with behavioural scoring in hypnotisability scales. Because the validity and reliability of the objective scale is numerically lower than that of the subjective scale in published

norms (Lush et al, 2018; note that this is also the case for WSGC validity and reliability), and because it is subjective experience rather than behavioural response which is the target of hypnosis research (Kihlstrom, 2008), we recommend the use of the subjective scale over the objective scale. Recently we have moved to reporting subjective hypnotisability scores alone (Lush et al, 2019. Lush, Roseboom, Cleeremans, Scott, Seth & Dienes, 2019).

While subjective ratings for a negative hallucination were numerically lower here when presented on a computer screen than in previous group screening when the image was projected onto a large lecture hall screen, reliability measures were similar for this item and previously published norms.

Computer delivery of a hypnosis scale offers advantages over a traditional delivery, providing uniformity of presentation to a degree which would not be possible in a screening programme spread over multiple sessions with live presentation of a script. The automated procedure provides an accessible way to screen participants for trait differences in phenomenological control abilities for researchers with no prior experience of hypnosis research. It therefore provides an effective way of controlling for implicit imaginative suggestion effects in experimental paradigms (e.g., we have used this scale to identify correlations between hypnotisability and the rubber hand illusion or mirror experience, Lush et al, 2020). Finally, the practical advantages of the SWASH have been retained. The procedure has been designed to be administered easily within the space of a typical lecture, and the recorded delivery can be run with as many participants as there are computers available. We note that, while this scale provides a relatively simple procedure for rapidly screening large numbers of participants, two stage screening procedures may provide more detailed information for particular applications.

This study was conducted in a sample of UK undergraduates, drawn from a relatively young population. Further research will be required to rule out the possibility of different

responses to this scale in other populations. Note that these results do not provide information regarding remote delivery of the SWASH (Palfi et al, 2019, report norms for remote delivery).

While there is no evidence that the experience of hypnosis in an academic setting is associated with more negative experiences than other non-hypnotic experimental procedures, and in fact hypnosis is associated with less negative experiences than listening to a lecture (for review, Heap et al., 2001), the experimenter following any protocol needs to be sensitive to possible negative experiences. This may be addressed by having an experimenter physically present when screening occurs, by Zoom calls before and after an online screening for debriefing, or at least by having contact details to report any negative reaction.

In conclusion, an automated, computer-based delivery of the Sussex Waterloo Scale of Hypnotisability provides an accessible and time efficient procedure for testing trait hypnotisability in an undergraduate sample.

References

Acunzo, D., & Terhune, D. (2019, September 16). A critical review of standardized measures of hypnotic suggestibility. https://doi.org/10.31234/osf.io/m93uy

Bowers, K. S. (1993). The Waterloo-Stanford Group C (WSGC) scale of hypnotic susceptibility: Normative and comparative data. *International Journal of Clinical and Experimental Hypnosis*, *41*(1), 35-46.

Barber, T. X., & Calverley, D. S. (1964a). Comparative effects on" hypnotic-like" suggestibility of recorded and spoken suggestions. *Journal of Consulting Psychology*, *28*(4), 384.

Barber, T. X., & Calverley, D. S. (1964b). Effect of E s Tone of Voice on “Hypnotic-Like” Suggestibility. *Psychological Reports*, *15*(1), 139-144.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum

Cox, R. E., & Barnier, A. J. (2010). Hypnotic illusions and clinical delusions: Hypnosis as a research method. *Cognitive Neuropsychiatry*, *15*(1-3), 202-232.

Dienes, Z., Lush, L., Palfi, B., Rooseboom, W., Scott, R., Parris, B., Seth, A., & Lovell, M. (in press). Phenomenological control as cold control: Hypnosis and beyond. Psychology of Consciousness: Theory, Research, and Practice.

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British journal of psychology, 105(3), 399-412.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: Sage .

Fassler, O., Lynn, S. J., & Knox, J. (2008). Is hypnotic suggestibility a stable trait? *Consciousness and cognition*, *17*(1), 240-253.

Heap, M., Alden, P., Brown, R. J., Naish, P., Oakley, D. A., Wagstaff, G., & Walker, L. G. (2001). *The Nature of Hypnosis: A report prepared by a Working Party at the request of The Professional Affairs Board of The British Psychological Society*. British Psychological Society.

Hull CL. Hypnosis and Suggestibility. New York: Appleton Century, 1933.

Kihlstrom, J. F. (2008). The domain of hypnosis, revisited. *The Oxford handbook of hypnosis: Theory, research and practice*, 21-52.

Lush, P., Botan, V., Scott, R. B., Seth, A. K., Ward, J., & Dienes, Z. (2020). Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. Nature communications, 11(1), 1-10.

Lush, P., Caspar, E. A., Cleeremans, A., Haggard, P., Magalhães De Saldanha da Gama, P. A., & Dienes, Z. (2017). The power of suggestion: posthypnotically induced changes in the temporal binding of intentional action outcomes. *Psychological science*, *28*(5), 661-669.

Lush, P., Roseboom, W., Cleeremans, A., Scott, R. B., Seth, A. K., & Dienes, Z. (2019). Intentional binding as Bayesian cue combination: Testing predictions with trait individual differences. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(9), 1206.

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.

MATLAB (2017). V*ersion 9.2 (R2017a)*. Natick, Massachusetts: The MathWorks Inc.

Oakley, D. A., Walsh, E., Lillelokken, A. M., Halligan, P. W., Mehta, M. A., & Deeley, Q. (2020). United Kingdom Norms for the Harvard Group Scale of Hypnotic Susceptibility, Form A. *International Journal of Clinical and Experimental Hypnosis, 68*(1), 80-104.

Palfi, B., Moga, G., Lush, P., Scott, R. B., & Dienes, Z. (2019). Can hypnotic suggestibility be measured online? *Psychological research*, 1-12.

Piccione, C., Hilgard, E. R., & Zimbardo, P. G. (1989). On the degree of stability of measured hypnotizability over a 25-year period. *Journal of Personality and Social Psychology*, *56*(2), 289.

Polito, V., Barnier, A. J., Woody, E. Z., & Connors, M. H. (2014). Measuring agency change across the domain of hypnosis. *Psychology of Consciousness: Theory, Research, and Practice, 1*(1), 3.

Weitzenhoffer, A. M., & Sjoberg, B. M. (1961). Suggestibility with and without "induction of hypnosis". *Journal of Nervous and Mental Disease*.