Psychological Assessment
2005, Vol. 17, No. 2, 200–211

1040-3590/05/$12.00 DOI: 10.1037/1040-3590.17.2.200

Multiple Hypnotizabilities: Differentiating the Building Blocks of Hypnotic Response

Erik Z. Woody
University of Waterloo

Amanda J. Barnier and Kevin M. McConkey
University of New South Wales

Although hypnotizability can be conceptualized as involving component subskills, standard measures do not differentiate them from a more general unitary trait, partly because the measures include limited sets of dichotomous items. To overcome this, the authors applied full-information factor analysis, a sophisticated analytic approach for dichotomous items, to a large data set from 2 hypnotizability scales. This analysis yielded 4 subscales (Direct Motor, Motor Challenge, Perceptual–Cognitive, Posthypnotic Amnesia) that point to the building blocks of hypnotic response. The authors then used the subscales as simultaneous predictors of hypnotic responses in 4 experiments to distinguish the contribution of each component from general hypnotizability. This analysis raises interesting questions about how best to conceptualize and advance measurement of the ability to experience hypnosis.

Keywords: hypnosis, hypnotizability, hypnotizability scales, factor analysis, multidimensionality

Hypnotizability is “the capacity to produce those effects generally considered to be ‘hypnotic’” (Weitzenhoffer, 1997, p. 128). In the 1950s and 1960s, Hilgard and his colleagues developed standardized scales for measuring hypnotic ability; in their view, hypnotizability was best measured by taking “samples of hypnotic performance under standard conditions of induction and testing” (Hilgard, 1965, p. 69). Their hypnotizability scales involve administering a standard induction procedure, suggesting various types of hypnotic experiences (hypnotic items), and scoring observable responses according to predetermined pass/fail criteria; hypnotizability is calculated by adding the items passed.

The three main types of hypnotic items (ideomotor, challenge, and perceptual–cognitive) contain three phases (suggestion, test, and cancellation). During the suggestion phase, the hypnotist suggests the effect to be experienced; during the test phase, the hypnotist measures response to the suggestion according to some predefined criterion; and during the cancellation phase, regardless of the response, the hypnotist tells the person to stop experiencing the suggested effect. Ideomotor items involve “thoughts becoming action” and usually lead to motor responses, like feeling one’s arm getting heavier and lowering or imagining one’s hands moving farther and farther apart. Challenge items involve the inhibition of motor responses, like trying to pull one’s fingers apart after being told they are tightly interlocked or trying to open one’s eyes despite being told they are tightly shut. Perceptual–cognitive items involve alterations in perceptual and cognitive functioning. For example, with appropriate suggestion, the person might see or hear things that are not present (positive hallucinations), fail to see or hear things that are present (negative hallucinations), or experience distortions in memory (age regression, posthypnotic amnesia). Within these types of hypnotic items, an almost infinite range of responses can be suggested (for a review of hypnotizability scales, see Barnier & McConkey, 2004).

Hypnotizability measures are generally thought of as measuring a unitary underlying trait, whereby individuals who score at the low end of the scale (*low hypnotizable*; typically about 10%–15% of the population) possess little of this trait; those who score in the midrange (*medium hypnotizable*; typically 70%–80%) possess a moderate degree; and those who score at the high end (*high hypnotizable*; typically about 10%–15%) possess a high degree (Hilgard, 1965). A problem left unanswered by existing scales, however, is how finely to resolve the measurement of hypnotic response. Is hypnotic ability better conceptualized and measured as a unitary phenomenon or as a set of distinguishable component subskills? Hilgard and colleagues tried to have it both ways: The Stanford Hypnotic Susceptibility Scale, Form C (SHSS:C; Weitzenhoffer & Hilgard, 1962), measures a unitary general trait, whereas the Stanford Profile Scales of Hypnotic Susceptibility, Forms I and II (Weitzenhoffer & Hilgard, 1963; Revised Stanford Profile Scales of Hypnotic Susceptibility, Weitzenhoffer & Hilgard, 1967), measure a set of distinguishable component subskills (direct motor action, challenges to motor action, perceptual–cognitive distortions, and posthypnotic amnesia). Hilgard (1965) interpreted factor-analytic results of hypnotizability measures as broadly supportive of this hybrid position. However, in subsequent

Erik Z. Woody, Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada; Amanda J. Barnier and Kevin M. McConkey, School of Psychology, University of New South Wales, Sydney, New South Wales, Australia.

This research was supported by a visiting fellowship from the Australian Psychological Society to Erik Z. Woody and Amanda J. Barnier, an operating grant from the Natural Sciences and Engineering Research Council of Canada to Erik Z. Woody, and an Australian Research Council Queen Elizabeth II Fellowship and Large Grant to Amanda J. Barnier. We are grateful for that support. We are grateful also for research assistance by Lyndel Mayoh and Rochelle Cox. Finally, we thank Auke Tellegen and John Kihlstrom for their penetrating comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Erik Z. Woody, Department of Psychology, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada. E-mail: ewoody@uwaterloo.ca

years, the unitary measurement approach of the SHSS:C has had the widest influence, whereas the component subskills approach of the Stanford Profile Scales of Hypnotic Susceptibility, Forms I and II, and the Revised Stanford Profile Scales of Hypnotic Susceptibility, Forms I and II, has had almost none (McConkey & Barnier, 2004).

Following Hilgard's (1965) lead and innovative work by Tellegen and Atkinson (1976), various researchers have considered this subskills issue, mainly by submitting data from standardized scales to a variety of factor analytic approaches. These researchers, however, have reached various conclusions about the underlying nature of hypnotizability, and a number of methodological problems have made some factor analytic work difficult to interpret (Balthazard, 1993; Balthazard & Woody, 1985). Chief among these problems is the issue of so-called difficulty factors, which are artifacts that can occur in factor analyses of product–moment correlations (phi coefficients) among dichotomously scored (pass or fail) items. For example, if a set of continuous measures of the same underlying skill were dichotomized at two very different difficulty levels, a subset of "easy" items would tend to define one factor, and a subset of "difficult" items would tend to define a second factor. An interpretive problem therefore arises when two or more factors derived from a hypnosis scale consist of items that differ not only in content but also in mean difficulty level: Are such factors substantively distinct latent variables, or are they merely spurious difficulty factors?

Tellegen and Atkinson (1976) undertook to resolve this interpretive dilemma by unconfounding content and difficulty level. To the original behavioral scoring for seven ideomotor and challenge items of the Harvard Group Scale of Hypnotic Susceptibility, Form A (HGSHS:A; Shor & Orne, 1962), they added polychotomous subjective response scales. They also asked participants to rate the difficulty of each hypnotic experience from four alternatives. Tellegen and Atkinson (1976) dichotomized the subjective response items in such a manner as to equate the difficulty levels of the direct and challenge items, thus unconfounding item content and item difficulty level. Factor analysis of the seven resulting variables yielded two orthogonal factors, consistent with Hilgard's (1965) results. Unfortunately, this analysis did not include cognitive items, and the changed scoring method limits comparability with the standardized scales. More recently, Kihlstrom (2001) also considered this issue by developing a new scale of hypnotic susceptibility that scored behavioral response along a 9-point continuous, rather than dichotomous, scale. Using his 16-item scale, the Arizona Motor Scale of Hypnotizability, Kihlstrom matched direct and challenge items for difficulty. Principal components analysis of six direct and six challenge suggestions again yielded two factors, which suggested that the distinction between these items is not an artifact of difficulty. Again, however, this analysis did not include a truly representative set of items and changed the standard scoring procedure.

Sadler and Woody (2004) noted that the recent development of multidimensional item-response theory (IRT) offers sophisticated statistical methods for addressing difficulty artifacts without compromising the standardized scales. These techniques are based on a model of dichotomous responses that avoids difficulty artifacts (Bock, Gibbons, & Muraki, 1988; Waller, Tellegen, McDonald, & Lykken, 1996). Applying IRT-based factoring methods to a large 39-year sample of responses to the HGSHS:A (Shor & Orne, 1962), Sadler and Woody (2004) found that this scale had a consistent two-factor structure involving Direct Motor (e.g., hand lowering) and Motor Challenge (e.g., arm rigidity) subscales.

We applied and extended an IRT approach to the SHSS:C, which is the gold standard of hypnotizability scales (Perry, Nadon, & Button, 1992). In doing this, there were two logistical problems with which we had to deal. First, the usual procedure in many hypnosis research laboratories, including those of the present authors, is to test potential participants on the HGSHS:A, preselect the highest and lowest scorers, and then retest only these preselected individuals on the SHSS:C. Although this is an efficient way of finding high and low hypnotizable individuals, for our present purposes it creates an artificial bimodal distribution of scores, in which potentially separable factors probably cannot be distinguished. To overcome this problem, we assembled a data set from those periods of hypnosis research at the University of New South Wales and Macquarie University (both located in Sydney, New South Wales, Australia) in which there was no preselection on the HGSHS:A before testing on the SHSS:C, thus allowing the sample to contain the full range of individual ability levels. Second, the SHSS:C includes a particular sample of hypnotic items with an overrepresentation of the perceptual–cognitive type (e.g., hallucinations). Whereas 7 of its 12 items are perceptual–cognitive, there are only 2 or fewer items for each of the other putative subtraits (direct motor, motor challenge, and posthypnotic amnesia). This limited representation of non-perceptual–cognitive components makes factor analysis of this scale potentially problematic. To overcome this problem, we used the total item pool, combining both the SHSS:C and the HGSHS:A, to obtain a more adequate sampling of the full range of hypnotic-item content.

Among the options for factor-analyzing dichotomous items while avoiding difficulty artifacts, we chose the full-information factor analysis (FIFA) method of Bock and colleagues (Bock & Aitkin, 1981; Bock et al., 1988). Bock et al. (1988) detailed how this method overcomes shortcomings of the older technique of principal factor analysis of tetrachoric correlation coefficients, including problems with indeterminate tetrachoric values for very easy or very difficult items, failure to converge on an admissible solution because of missing items and other data configurations, and the lack of a test of statistical significance for the addition of successive factors. The FIFA method uses tetrachoric correlations only to determine starting values; thus, problems with imputing and smoothing the tetrachoric matrix do not bias the resulting solution. Bock and colleagues have shown that the FIFA method yields parameter estimates and statistical tests for successive factors that compare closely with those of another important method, the limited-information generalized-least-squares technique of Muthén (1978); also, they have shown that the FIFA method is highly effective in guarding against artifactual difficulty factors.

Our first goal, addressed in Part 1, was to provide a clearer view of the component abilities, or building blocks, of hypnotic response. Our second goal, addressed in Part 2, was to show that the subscales that measure these building blocks differentially predict a range of hypnotic behaviors. In particular, our aim was to show that specific abilities predict hypnotic performance over and above prediction due to general hypnotizability and thus help to illuminate the diverse nature of hypnotic performance.

Part 1: Psychometric Analysis of the Combined SHSS:C and HGSHS:A Items

We used factor analysis and other psychometric analyses to construct subscales that distinguish component traits underlying hypnotic response. Our first step was to submit the combined set of dichotomous SHSS:C and HGSHS:A items to an IRT-based factoring method specially developed for use with dichotomous items. We expected support for the conceptual analysis of hypnotic items that divides them into three or four groups. Our second step was to investigate whether these item groups yield subscales with good psychometric properties, such as high internal consistency and reasonable unique variance. We also submitted the subscales to separate IRT analyses to examine whether they discriminate among people at various levels of the underlying traits.

Method

The participants were 616 (176 male and 440 female; age, $M = 20.72$ years, $SD = 5.51$) undergraduate psychology students at the University of New South Wales or Macquarie University, Sydney, New South Wales, Australia, who voluntarily participated in hypnosis testing in return for research credit. The data came from 5 years in which, because of the nature of the research being conducted at that time, there was no preselection of high and low hypnotizables on the HGSHS:A prior to testing on the SHSS:C. These years were 1986 (79 participants), 1994 (127 participants), 1995 (141 participants), 2000 (132 participants), and 2001 (137 participants).

The participants first were administered the HGSHS:A in groups of 3 to 30 and then returned for the individually administered SHSS:C. Because slightly modified or tailored versions of the scales (Hilgard, Crawford, Bowers, & Kihlstrom, 1979) were used for some research projects, there are some constraints on the data set; it is most important to note that there are no data available for the SHSS:C hallucinated voice item (No. 10 in the original scale). Although there are between 610 and 616 observations for almost all the other items of the two scales, there is a smaller number of observations for the following four items: SHSS:C anosmia to ammonia (No. 9; 220 observations), SHSS:C negative visual hallucination (No. 11; 220 observations), HGSHS:A arm immobilization (No. 4; 345 observations), and HGSHS:A arm rigidity (No. 6; 347 observations).

Results

Factor analysis of the combined HGSHS:A and SHSS:C items. We submitted the 23-item data set to the FIFA method as implemented in the program TESTFACT (Wilson, Wood, & Gibbons, 1991). This technique is based on the marginal maximum-likelihood method of estimating parameters for multidimensional item-response models (Bock & Aitkin, 1981; Bock et al., 1988) and also provides a statistical test of the contribution of each additional factor. Wilson et al. (1991) indicated that these tests should be interpreted cautiously, particularly in data combining multiple samples, and recommended reevaluating the probability by dividing the chi-square by a design factor of two or three. For our 23-item data set, the addition of a second factor was clearly significant, $\Delta\chi^2(22, N = 616) = 163.11$, yielding $p < .001$, whether corrected or not. The addition of a third factor was more equivocal, $\Delta\chi^2(21, N = 616) = 52.57$, yielding an uncorrected $p < .001$, but an insignificant probability ($p > .10$) when corrected; in addition, with three factors the program succeeded in meeting only a relatively lax convergence criterion for the parameter estimates. Finally, the addition of a fourth factor was clearly significant, $\Delta\chi^2(20, N = 616) = 106.71$, yielding $p < .001$, whether corrected or not.

Despite the statistically equivocal nature of the third factor, we first briefly describe the promax-rotated three-factor solution (including items and their factor loadings), because this solution corresponds most directly with the prevailing conception of hypnosis scales. The first factor was Motor Challenge, defined by SHSS:C arm immobilization (No. 8, .45) and HGSHS:A arm immobilization (No. 4, .93), finger lock (No. 5, .86), arm rigidity (No. 6, .94), communication inhibition (No. 8, .63), and eye catalepsy (No. 10, .93). The modest loading of SHSS:C arm rigidity (No. 5, .27) on this factor, even though it is very similar to HGSHS:A arm rigidity (No. 6), was surprising, as was the fact that HGSHS:A posthypnotic amnesia (No. 12, .30) had its highest (albeit modest) loading on this factor. The second factor was Direct Motor, defined by SHSS:C hand lowering (No. 1, .93) and hands apart (No. 2, .88), and HGSHS:A head falling (No. 1, .36), eye closure (No. 2, .36), hand lowering (No. 3, .64), and hands moving (No. 7, .49). The loadings on this factor of SHSS:C mosquito hallucination (No. 3, .42) and HGSHS:A fly hallucination (No. 9, .57), as well as HGSHS:A posthypnotic suggestion (No. 11, .24) to a more modest extent, were surprising; however, passing these items involves a direct motor response (e.g., shooing away the mosquito or fly). The third factor was Perceptual–Cognitive, defined by SHSS:C mosquito hallucination (No. 3, .46), taste hallucination (No. 4, .52), dream (No. 6, .41), age regression (No. 7, .29), anosmia to ammonia (No. 9, .56), negative visual hallucination (No. 11, .98), and posthypnotic amnesia (No. 12, .61). No HGSHS:A items loaded on this factor, which is consistent with the findings of Sadler and Woody (2004) that the HGSHS:A lacks a clear third, perceptual–cognitive factor. An important anomaly was the pattern of loadings of the two posthypnotic amnesia items. Although SHSS:C posthypnotic amnesia (No. 12) loaded on this Perceptual–Cognitive factor, the very similar HGSHS:A posthypnotic amnesia (No. 12) did not.

The first four columns in Table 1 present the four-factor solution, which was statistically superior to the three-factor version and is the one on which we focus. The first three factors are essentially similar to those in the three-factor solution but are trivially reordered, with the posthypnotic amnesia items removed from the Perceptual–Cognitive factor. The first factor is now Perceptual–Cognitive, the second is Motor Challenge, and the third is Direct Motor. The fourth factor contains the two posthypnotic amnesia items (SHSS:C No. 12 and HGSHS:A No. 12) to form a Posthypnotic Amnesia dimension.

The four-factor solution is conceptually appealing, but there are some surprising aspects to it. First, the Perceptual–Cognitive dimension is more cleanly defined by negative (SHSS:C Nos. 9 and 11) than by positive (SHSS:C Nos. 3 and 4) hallucinations. In addition, the dream and age-regression items (SHSS:C Nos. 6 and

Table 1
Promax-Rotated Four-Factor Solution: Factor Loadings

Item	Primary factor 1	Primary factor 2	Primary factor 3	Primary factor 4	Second-order general factor
SHSS-C					
1. Hand lowering	−.02	−.16	.92	.03	.63
2. Hands apart	−.08	−.04	.86	.05	.62
3. Mosquito hallucination	.42	−.12	.43	.14	.61
4. Taste hallucination	.49	−.04	.35	.17	.67
5. Arm rigidity	.73	.22	.00	−.01	.78
6. Dream	.29	.22	.06	.23	.47
7. Age regression	.31	.13	.21	.08	.54
8. Arm immobilization	.68	.32	.02	−.15	.84
9. Anosmia to ammonia	.88	−.04	−.15	.00	.57
11. Negative visual hallucination	.80	−.14	−.07	.32	.49
12. Posthypnotic amnesia	.23	−.13	.13	.76	.19
HGSHS:A					
1. Head falling	−.18	.33	.40	.10	.45
2. Eye closure	.11	.29	.35	−.08	.62
3. Hand lowering	.00	−.18	.63	−.04	.38
4. Arm mobilization	.01	.90	−.11	.02	.65
5. Finger lock	.19	.79	.01	−.21	.81
6. Arm rigidity	−.02	.99	−.16	.06	.66
7. Hands moving	−.11	.16	.56	−.07	.51
8. Communication inhibition	.30	.52	.06	.00	.73
9. Fly hallucination	−.11	.06	.68	.05	.53
10. Eye catalepsy	−.08	.86	.11	−.01	.73
11. Posthypnotic suggestion	.22	−.01	.28	−.19	.40
12. Posthypnotic amnesia	−.10	.38	−.04	.45	.19

Note. SHSS:C hallucinated voice item (No. 10) was omitted. Primary factor 1 = Perceptual-Cognitive; Primary factor 2 = Motor Challenge; Primary factor 3 = Direct Motor; Primary factor 4 = Posthypnotic Amnesia; SHSS:C = Stanford Hypnotic Susceptibility Scale, Form C; HGSHS:A = Harvard Group Scale of Hypnotic Susceptibility, Form A.

7) seem relatively weak.[1] Second, the Motor Challenge dimension is better defined by the motor challenge items of the HGSHS:A (Nos. 4, 5, and 6) than of the SHSS:C (Nos. 5 and 8), which tend to load more on the Perceptual–Cognitive factor. For example, HGSHS:A arm rigidity (No. 6) loads exclusively on the Motor Challenge factor, whereas the very similar SHSS:C arm rigidity (No. 5) loads mainly on the Perceptual–Cognitive factor. Given that this SHSS:C item is preceded by perceptual–cognitive items but that the corresponding HGSHS:A item is not, this difference suggests that the context of the preceding suggestions may affect what such an item taps. Third, although the Direct Motor dimension comprises the direct motor items of both scales (SHSS:C Nos. 1 and 2; HGSHS:A Nos. 3 and 7), HGSHS:A fly hallucination (No. 9) loads on it, as do, to a lesser extent, the similar SHSS:C mosquito hallucination (No. 3) and HGSHS:A posthypnotic amnesia (No. 11). Although these three items would usually be regarded as involving perceptual–cognitive alterations, passing them requires an explicit motor response.

As can be seen by the interfactor correlations in Table 2, the factors are substantially intercorrelated, although somewhat less so for the Posthypnotic Amnesia dimension. Our hypothesis is that these intercorrelations reflect the presence of a single higher order factor of general hypnotizability (Hilgard, 1965). To evaluate this, we submitted the matrix of interfactor correlations to a confirmatory factor analysis (CFA) by using the program Amos 5.0 (Arbuckle & Wothke, 1999). A model positing one higher order factor fit very well, $\chi^2(2, N = 616) = 4.47$, $p = .11$, comparative fit index = .998, root-mean-square error of approximation = .045,[2] and yielded the estimated loadings shown in the last column in Table 2. The loading of each item on the higher order general factor can be estimated via the sum of cross-products of these second-order loadings with the loadings for the primary factors (Gorsuch, 1983). These values, which appear in the last column in Table 1, show that Motor Challenge items tend to be strongly related to general hypnotizability (e.g., HGSHS:A Nos. 5, 8, and 10; SHSS:C Nos. 5 and 8), whereas Posthypnotic Amnesia items are only modestly related to it (HGSHS:A No. 12; SHSS:C No. 12).

[1] The emphasis on the two negative hallucination items over other perceptual–cognitive items may be an artifact of the smaller number of observations for these two items. In the Perceptual–Cognitive subscale formed later, all types of items (e.g., negative hallucinations, positive hallucinations, and cognitive suggestions) showed strong item-total correlations.

[2] Widely acknowledged criteria for excellent model fit are a nonsignificant chi-square, a comparative fix index greater than .95, and a root-mean-square error of approximation less than .05 (Arbuckle & Wothke, 1999).

Table 2
Promax-Rotated Four-Factor Solution: Interfactor Correlations

Primary factor	Primary factor 1	Primary factor 2	Primary factor 3	Primary factor 4	Loadings on general factor
1	—				.83
2	.67	—			.82
3	.69	.69	—		.83
4	.48	.43	.43	—	.54

Note. Primary factor 1 = Perceptual-Cognitive; Primary factor 2 = Motor Challenge; Primary factor 3 = Direct Motor; Primary factor 4 = Posthypnotic Amnesia.

In summary, the four-factor solution yields reasonably well-defined Perceptual–Cognitive, Motor Challenge, Direct Motor, and Posthypnotic Amnesia dimensions. Some details of the solution that depart from an expected ideal solution are likely attributable to the mixed quality of some items or to problems of local dependence. Some items may call on two or more relevant individual differences. For example, the SHSS:C mosquito hallucination (No. 3) suggests both a positive hallucination and a direct motor response. Also, responses elicited by a suggestion may depend somewhat on the preceding suggestions (Balthazard & Woody, 1985); for instance, whereas the HGSHS:A seems to evoke a motor challenge set, the SHSS:C seems to evoke a perceptual–cognitive set. Finally, the correlations among the four factors are consistent with the presence of a single higher order factor of general hypnotizability, which contributes substantial and important nonspecific variance to each of the primary factors.

Subscale construction and psychometric properties. We next divided the item set to form subscales measuring each of the four dimensions. Where details of the factor analysis departed from the prevailing conception of hypnotic item content, we retained the conceptually guided distinctions. For example, we put SHSS:C arm rigidity (No. 5) into the Motor Challenge subscale rather than the Perceptual–Cognitive subscale, as a literal reading of the four-factor solution might indicate. This tempering of some of the factor analytic results makes good sense given the complications of mixed-item quality and local dependence.[3]

As many methodologists have recommended, we unit-weighted and summed the items for each subscale to compute subscale scores (Gorsuch, 1983; Nunnally & Bernstein, 1994; Wainer, 1976).[4] Thus, scores for the Direct Motor, Motor Challenge, and Perceptual–Cognitive subscales are the sums of the relevant set of dichotomously scored items from the two instruments. However, rather than basing scores for the Posthypnotic Amnesia subscale on the original, dichotomously scored posthypnotic amnesia items (i.e., individuals fail this item if they recall three or more hypnotic items during the test of amnesia, and pass if they recall two or fewer items), we used the actual number of suggestions the participant remembered for each scale. This graded scoring retained more information than did the dichotomous scoring and improved the reliability of the resulting subscale scores.[5] So that higher scores on the Posthypnotic Amnesia subscale would denote greater amnesia, we computed the subscale scores by subtracting the total number of suggestions remembered on each scale from 23 (the maximum total number that could be recalled).

The resulting subscales are listed in Table 3, together with their internal-consistency reliabilities and intercorrelations. The Cronbach's alphas for the subscales were reasonable, although those for the Direct Motor and Posthypnotic Amnesia subscales were lower. The intercorrelations among the subscales were substantial, but they were not so high as to be inconsistent with indexing discriminable individual differences. For example, the alpha for the Direct Motor subscale indicated that 65% of its variance is true-score variance. A multiple regression with the Direct Motor subscale as the criterion and the other three subscales as the predictors accounted for 37% of the variance. Taking the difference, we can infer that about 28% of the variance of this subscale is true-score variance unique to it. Likewise, 38% of the variance of the Motor Challenge subscale, 48% of the variance of the Perceptual–Cognitive subscale, and 27% of the variance in the Posthypnotic Amnesia subscale are true-score variance unique to that subscale. Thus, there is a reasonable amount of potentially useful unique variance in each subscale. Finally, a CFA of the subscale intercorrelations again confirmed the presence of a single higher order factor of general hypnotizability, $\chi^2(2, N = 616) = 3.20$, $p = .20$, comparative fit index = .998, root-mean-square error of approximation = .031.

IRT assessment of the subscales. Because hypnotic items were not designed to measure the distinctions of the subscales, the subscales may have inadvertent shortcomings. Accordingly, we tested the ability of the Direct Motor, Motor Challenge, and Perceptual–Cognitive subscales to discriminate among people at various levels of the underlying traits via IRT analyses. Using the two-parameter model as implemented by the program MULTILOG (Thissen, 1988), we generated separate information curves for each of these three subscales; the Posthypnotic Amnesia subscale could not be readily analyzed in this way because its two items have a wide range of possible scores, rather than being dichotomous.

See Sadler and Woody (2004) for illustrations of information curves for the HGSHS:A and their interpretation; here we simply

[3] An alternative would be to omit items from subscales if they did not load on the conceptually expected factor. Hence, SHSS:C arm rigidity (No. 5) and arm immobilization (No. 8) might be omitted from the Motor Challenge subscale, and HGSHS:A fly hallucination (No. 9) and posthypnotic suggestion (No. 11) might be omitted from the Perceptual–Cognitive subscale. Subscales that excluded these items showed the same patterns of significant correlations with experimental behaviors as the subscales that included them.

[4] More sophisticated alternatives would include calculating factor scores with TESTFACT or expected a posteriori estimates for each subscale with MULTILOG. Although intriguing, these possibilities would make the present work more difficult for other researchers to use and replicate.

[5] We settled on this scoring of amnesia after trying various ways of incorporating reversal (i.e., recovery of memories after the cancellation of the amnesia suggestion). For both hypnosis scales, we looked at the difference between the number of suggestions remembered before versus after cancellation; reminiscent of Kihlstrom and Register's (1984) findings, neither continuous nor dichotomous versions of this index improved the reliability of the resulting subscale. Also, incorporating reversal neither changed the role of these items in the factor analyses nor increased the subscale's correlations with other variables, such as the experimental posthypnotic amnesia effects obtained in Barnier et al. (2004).

Table 3
Internal-Consistency Reliabilities and Intercorrelations of Four Subscales

Subscale	1	2	3	4
1. Direct Motor	.65			
2. Motor Challenge	.58	.86		
3. Perceptual-Cognitive	.55	.64	.79	
4. Posthypnotic Amnesia	.33	.44	.43	.58

Note. Cronbach's alphas are on the main diagonal; for all correlations, $p < .001$. Direct Motor subscale = Stanford Hypnotic Susceptibility Scale, Form C (SCSS:C), hand lowering (No. 1) + hands apart (No. 2) + Harvard Group Scale of Hypnotic Susceptibility, Form A (HGSHS:A), head falling (No. 1) + eye closure (No. 2) + hand lowering (No. 3) + hands moving (No. 7); Motor Challenge subscale = SHSS:C arm rigidity (No. 5) + arm immobilization (No. 8) + HGSHS:A arm immobilization (No. 4) + finger lock (No. 5) + arm rigidity (No. 6) + communication inhibition (No. 8) + eye catalepsy (No. 10); Perceptual–Cognitive subscale = SHSS:C mosquito hallucination (No. 3) + taste hallucination (No. 4) + dream (No. 6) + age regression (No. 7) + anosmia to ammonia (No. 9) + negative visual hallucination (No. 11) + HGSHS:A fly hallucination (No. 9) + posthypnotic suggestion (No. 11); Posthypnotic Amnesia subscale = 23 - SHSS:C posthypnotic amnesia (No. 12, items recalled) - HGSHS:A posthypnotic amnesia (No. 12, items recalled).

provide the summary implications of the curves obtained for the present subscales. The information function for the Direct Motor subscale showed that this subscale is useful mainly for distinguishing among quite low levels of its underlying trait (centered around a standard deviation below the mean); it provides little information in the middle to high range. In contrast, the information provided about individual differences by the Motor Challenge and the Perceptual–Cognitive subscales peaks at the population mean for each underlying trait. Thus, these two subscales appear to strike a better balance of information than does the Direct Motor subscale. The Motor Challenge subscale is particularly tuned to discriminating in the midrange (i.e., trait levels near the average), whereas the Perceptual–Cognitive subscale is more tuned to discriminating among high levels of its underlying trait.

Discussion

Informed by our factor analysis as well as conceptual insights about hypnotic responses, we derived four subscales with acceptable psychometric properties from the combined pool of SHSS:C and HGSHS:A items. Although the moderate correlations among these subscales reflect the common influence of a general trait of hypnotizability, each subscale captures unique variance that can be disentangled from the common variance by using multivariate methods. Hence, our subscales offer the promise of measuring building blocks that underlie particular hypnotic responses (Woody & McConkey, 2003). Our analyses also suggest important ways to improve the subscales in future work. For instance, the capacity of the Direct Motor subscale to discriminate individuals could be enhanced by adding two or three more difficult items (e.g., automatic writing). Also, the reliability of the Posthypnotic Amnesia subscale could be improved by adding another amnesia item to the existing two items; suggested amnesia for an autobiographical memory is a promising candidate (Barnier, 2002; Barnier, Wright, & McConkey, 2004; Cox & Barnier, 2003).

An intriguing sidelight of the factor analyses is evidence that the suggestions preceding an item may to some extent affect responses to that item. In particular, the surprising difference between the two scales in the nature of responses to the arm immobilization and arm rigidity items may be due to the fact that in the SHSS:C they are preceded by perceptual–cognitive suggestions, whereas in the HGSHS:A they are preceded only by motor suggestions. Such a carryover or context effect is reminiscent of Hull's (1933) concept of *heteroactive hypersuggestibility,* except that previous suggestions may tend to inhibit, as well as facilitate, the role of particular subskills in an individual's response to a suggestion.[6]

It is useful to consider the four types of items delineated by the present subscales from the vantage point of facet theory (Shye, Elizur, & Hoffman, 1994), which is an experimental-design-like approach to measurement construction. The four types of items can be thought of as comprising the cells of a 2 × 2 matrix, for which the facets are motor versus perceptual–cognitive crossed with direct versus challenge. For instance, SHSS:C hand lowering (No. 1) falls into the Motor × Direct cell because participants are directed to experience a specific motor action; HGSHS:A finger lock (No. 5) falls into the Motor × Challenge cell because participants are told that their fingers are tightly interlocked and are then challenged to separate them (i.e., inhibition of a motor action); SHSS:C taste hallucination (No. 4) falls into the Perceptual–Cognitive × Direct cell because participants are directed to experience an hallucinated taste; and SHSS:C (No. 12) posthypnotic amnesia items falls into the Perceptual–Cognitive × Challenge cell because participants are told that they will not be able to recall the events of hypnosis and are then challenged to try (i.e., inhibition of a cognitive action). According to this reconceptualization of our 23 hypnosis items, 7 are direct motor (HGSHS:A hand lowering and hands apart; SHSS:C head falling, eye closure, hand lowering, hands moving, and posthypnotic suggestion), 7 are motor challenge (SHSS:C arm rigidity and arm immobilization; HGSHS:A arm immobilization, finger lock, arm rigidity, communication inhibition, and eye catalepsy), 5 are direct perceptual–cognitive (SHSS:C mosquito hallucination, taste hallucination, dream, and age regression; HGSHS:A fly hallucination), and 4 are perceptual–cognitive challenge (SHSS:C anosmia to ammonia, negative visual hallucination, and posthypnotic amnesia; HGSHS:A posthypnotic amnesia).

This reconceptualization also reveals possible confounds that could be addressed in future research. First, direct versus challenge is confounded with the manner of responding in the items. Whereas for most direct items participants must do or say something to pass, for most challenge items they must fail to do or say something (Sadler & Woody, 2004). This confounding seems inadvertent rather than necessary, because a challenge item that requires a response for a pass is readily conceivable: for instance, the suggestion that a weight in the hand will become too heavy to hold up, yet the subject should try to lift it up. Second, there are only a few examples of perceptual–cognitive challenge items, and

[6] It would be an interesting challenge to represent these possible context effects in a CFA: for example, by using a lag-one set of error covariances or by including a method factor for each hypnosis scale. However, because of the dichotomous nature of the items, such an analysis would require a much larger sample (e.g., Marsh, 1996).

the wording of their challenge is less explicit than for motor items. In motor challenge items, participants are clearly told to actively challenge the reality of the suggested experience, but in perceptual–cognitive challenge items, this challenge is only implied. Because current perceptual–cognitive challenge items may be closer in their impact to direct suggestions, it would be worth inventing or revising items for this cell. Third, there are only three posthypnotic items in the entire set, and they fall unevenly across the cells (HGSHS:A posthypnotic suggestion in direct motor, and HGSHSA and SHSS:C posthypnotic amnesia in perceptual–cognitive challenge). Because posthypnotic responding is both theoretically and clinically important, a broader and more systematic sampling of posthypnotic suggestions is necessary; for examples of possible posthypnotic suggestions and their pass percentages, see Barnier and McConkey (1996, 1998a, 1998b, 1999a).

Part 2: Multivariate Prediction of Hypnotic Responses

Given that the combined SHSS:C and HGSHS:A items yielded four subscales with good psychometric properties, we used the subscales as simultaneous predictors of hypnotic responses. Multivariate methods—logistic regression for dichotomous outcomes and multiple regression for continuous ones—allowed us to distinguish the unique contributions of the subscales from the common contribution of general hypnotizability. Assuming that these components represent specific building blocks of hypnotic response, we expected to show that different types of response call on different component skills.

Method

We reanalyzed data from four experiments conducted at the University of New South Wales and Macquarie University, Sydney, New South Wales, Australia. These were experiments on (a) suggested color blindness (Mallard & Bryant, 2001), (b) posthypnotic amnesia for autobiographical memory (Barnier et al., 2004), (c) verbal posthypnotic suggestion (Barnier & McConkey, 1996, Experiment 2), and (d) motor posthypnotic suggestions (Barnier & McConkey, 1999b). For each experiment, data were available for participants across the full range of hypnotizability; note that in some of the articles, only the data for the high hypnotizables were presented. In addition, for each experiment, scores on the items of the SHSS:C and HGSHS:A were available for the calculation of subscale scores. (In all the analyses reported below, the regression coefficients reported are for the subscales in standard-score form.)

Results

Suggested color blindness. Mallard and Bryant (2001) tested 146 participants, of whom 37 (25.3%) passed a hypnotic suggestion for color blindness and 109 (74.7%) did not. With logistic regression, we analyzed this dichotomous criterion variable as a function of our four subscales of Direct Motor, Motor Challenge, Perceptual–Cognitive, and Posthypnotic Amnesia. Because the subscales partly measure general hypnotizability and because general hypnotizability is relevant to passing most hypnotic suggestions, we anticipated that each subscale would show a significant univariate relationship with suggested color blindness. Confirming this reasoning, the univariate logistic-regression coefficients for each subscale as a single, separate predictor of color blindness were all highly significant ($p < .001$). However, the univariate relationships will necessarily also be affected by any unique contributions of the subscales; thus, a multivariate strategy is required to disentangle the various contributions. Our prediction was that when the four subscales served as simultaneous (multivariate) predictors of suggested color blindness, the Perceptual–Cognitive subscale would emerge as a significant unique predictor. This follows from the hypothesis that people who show hallucination-like changes in hypnosis possess a specific skill that other high hypnotizables do not, and that the Perceptual–Cognitive subscale measures this. For example, Szechtman, Woody, Bowers, and Nahmias (1998) distinguished between two groups of participants who were equally highly hypnotizable: those who could hallucinate readily and vividly and those who could not.

In the multivariate equation using the four subscales as simultaneous predictors, the coefficient for the Perceptual–Cognitive subscale was highly significant ($B = 1.41$, $p < .001$), supporting our prediction, whereas the coefficients for the other three subscales did not approach significance. Effects in a logistic regression can also be quantified as proportions of variation by using the Nagelkerke R^2, which is akin to R^2 in multiple regression. Removing the Perceptual–Cognitive subscale from the equation reduces the Nagelkerke R^2 from .42 to .26, indicating that 16% of the variation is uniquely explained by this subscale, $\chi^2(1, N = 146) = 20.44$, $p < .001$. In contrast, removing the three other subscales from the equation reduces the R^2 by less than 1%, indicating that these subscales do not uniquely explain any of the variation, $\chi^2(3, N = 146) = 0.99$, $p = .80$. Given our earlier demonstration that the shared variance among the four subscales is due to a single higher order source, we can subtract the unique proportions from the total of proportion of .42 to infer that about 25% of the variance is attributable to general hypnotizability.[7]

These results strongly confirm both our univariate and multivariate predictions. As a result of the shared influence of general hypnotizability, each subscale considered separately was related to passing the suggestion for color blindness. However, beyond general hypnotizability, only the Perceptual–Cognitive subscale was associated with the ability to pass this suggestion. This unique multivariate relationship supports the idea that the Perceptual–Cognitive subscale measures a basic building block that is crucial to the ability to enact hypnotic suggestions of hallucination-like phenomena.

Posthypnotic amnesia for autobiographical memory. In Barnier et al.'s (2004) experiment, 95 participants were asked to recall two autobiographical memories. They were then hypnotized and given a posthypnotic amnesia suggestion for one of these episodes. After hypnosis, they were asked to recall the episodes again. Because it is rare for even very highly hypnotizable people to completely forget an entire autobiographical episode (Barnier,

[7] Our method—regression using the four subscale scores—is readily understandable and has straightforward practical implications (e.g., to predict suggested color blindness, one should use the Perceptual–Cognitive subscale; the other three subscales would add nothing to the prediction). Another approach to distinguish the contributions of the general and primary factor subscales would be to use a method such as that of Hendrickson and White (1966) to obtain factor-analytically derived scores on the higher order factor of general hypnotizability, and then to use hierarchical multiple regression to distinguish this contribution from the contributions of the four primary factors; we believe this would yield findings very similar to those we obtained.

2002; Cox & Barnier, 2003, Barnier et al. (2004) used multiple measures to index both memory accessibility and memory quality. Accordingly, we reanalyzed data for (a) *memory inaccessibility,* failing to produce the targeted memory (coded 1) versus producing the memory (coded 0), as rated by an observer; (b) *lack of narrative,* no identifiable narrative in the memory (coded 1) versus an identifiable narrative (coded 0), as rated by an observer; and (c) *lack of clarity,* a self-rating on a 7-point scale, where high scores meant that the participant experienced their recall as "dim" and low scores meant that they experienced their recall as "sharp and clear." Note that for these analyses, we recoded the original variables so that higher scores represent a stronger response to the amnesia suggestion. Also, we focused on the memory targeted by posthypnotic amnesia; the data for the nontargeted memory yielded similar but weaker findings.

We expected that each of the four subscales would show a significant univariate relationship to the memory indices, reflecting the common influence of general hypnotizability. Moreover, we expected that when the four subscales serve as simultaneous multivariate predictors of experiencing amnesia, the Posthypnotic Amnesia subscale would emerge as a significant unique predictor, reflecting its measurement of the specific building block relevant to hypnotic amnesia; however, we noted that the distinction among the memory indices raised the possibility that the patterns of unique prediction might differ.

Focusing first on *memory inaccessibility,* we found that inability to produce the targeted autobiographical memory was rare. Of the 95 participants, only 8 (8.4%) did not generate the memory, whereas 87 (91.6%) did. Despite this, the univariate relationship of each subscale to inaccessibility was significant ($p < .05$). However, at the multivariate level, it was the Perceptual–Cognitive subscale ($B = 1.37$, $p < .05$), rather than the expected Posthypnotic Amnesia subscale, that emerged as uniquely predictive. Removing the Perceptual–Cognitive subscale from the equation reduced the Nagelkerke R^2 from .35 to .23, indicating that 12% of the variation is uniquely explained by this subscale, $\chi^2(1, N = 95) = 5.66$, $p < .05$. In contrast, removing the three other subscales from the equation reduced the R^2 by an insignificant 3%, indicating that these subscales do not uniquely explain any of the variation, $\chi^2(3, N = 95) = 1.52$, $p = .68$. We can infer that about 20% of the variance is attributable to general hypnotizability.

Focusing next on *lack of narrative,* which was a more prevalent effect of the amnesia suggestion, we found that 21 (22.1%) participants lacked an identifiable narrative, whereas the memories of 74 (77.9%) had a clear narrative. Once again, the univariate relationship of each subscale to this measure was significant ($p < .01$). At the multivariate level, however, only the expected Posthypnotic Amnesia subscale emerged as uniquely predictive ($B = 1.06$, $p < .01$). Removing the Posthypnotic Amnesia subscale from the equation reduced the Nagelkerke R^2 from .34 to .24, indicating that 10% of the variation is uniquely explained by this subscale, $\chi^2(1, N = 95) = 7.51$, $p < .01$. In contrast, removing the three other subscales from the equation reduced the R^2 by less than 1%, indicating that these subscales do not uniquely explain any of the variation, $\chi^2(3, N = 95) = 2.14$, $p = .54$. We can infer that about 23% of the variance is attributable to general hypnotizability.

Focusing finally on *lack of clarity,* we found that this variable was well distributed, with a mean of 3.84 ($SD = 1.67$) near the midpoint of the 7-point Likert scale. The univariate relationship (correlation) of each subscale with this variable was significant (rs $= .41$ to $.43$, $p < .001$). Yet in a multiple regression, none of the subscales was uniquely predictive, each contributing an insignificant 1% to the explained variance. The implication is that virtually all of the 25% of the variance explained by the subscales in *lack of clarity* is attributable to general hypnotizability.

In summary, for each memory index, the four subscales were significant predictors at the univariate level. However, the multivariate findings were different for each of the memory variables. With the four subscales serving as simultaneous multivariate predictors, the Perceptual–Cognitive subscale showed an unexpected unique relationship with *memory inaccessibility,* the Posthypnotic Amnesia subscale showed an expected unique relationship with *lack of narrative,* and no subscale showed a unique relationship with *lack of clarity*. Although *lack of narrative* and *lack of clarity* are both indices of memory quality, they differ in that the former was objectively scored on the basis of what the participant actually recalled, whereas the latter was a self-rated impression. It seems that the Posthypnotic Amnesia subscale measures a specific trait relevant to memory quality as objectively rated, whereas only general hypnotizability is relevant to qualitative self-impressions of memory quality.[8]

Verbal posthypnotic suggestion. In Barnier and McConkey's (1996) Experiment 2, 126 participants were hypnotized and given the posthypnotic suggestion that when asked, "Well, what did you think of that?" they would say "Psych 1." The experimenter also suggested that they would forget that they had been told to do so. The data we reanalyzed focused on participants' responses to a formal test of the posthypnotic suggestion immediately after awakening. One variable was whether participants made any verbal response to the posthypnotic cue. The other variable was participants' self-rating on a 7-point Likert scale of the extent to which they felt the urge to respond.

We expected that each subscale would show a significant univariate relationship to these variables, reflecting the common influence of general hypnotizability. However, when the four subscales serve as simultaneous multivariate predictors of responding to the posthypnotic suggestion, the Perceptual–Cognitive and Posthypnotic Amnesia subscales may be uniquely predictive. The former subscale is implicated because the HGSHS:A posthypnotic suggestion item loaded on this factor (albeit not strongly), and the latter subscale is implicated because of the amnesia component of the posthypnotic suggestion.

Focusing first on the behavioral response, we found that 81 (64.3%) participants responded to the posthypnotic suggestion and 45 (35.7%) did not. The univariate relationships were less consistent than in the previous analyses, suggesting less of a role for general hypnotizability: Although the Perceptual–Cognitive subscale was significant ($p < .001$), as was the Direct Motor subscale

[8] Other objectively scored versus self-rated measures of memory quality, which we analyzed but do not present, showed the same difference in their relationship to the four subscales. Like *lack of narrative,* an objectively scored measure of *lack of specificity* showed significant univariate relationships with each of the subscales, but a unique relationship with only the Posthypnotic Amnesia subscale, as expected. Like *lack of clarity,* a self-rated measure of *lack of thoughts and feelings* showed significant univariate relationships with each of the subscales, but a unique relationship with none of them.

($p < .05$), the other two subscales were not. At the multivariate level, the Perceptual–Cognitive subscale emerged strongly in the expected way as uniquely predictive ($B = 1.28$, $p < .001$), although the Posthypnotic Amnesia subscale did not. In addition, the Motor Challenge subscale showed up as a significant suppressor variable ($B = -0.67$, $p < .05$). Removing the Motor Challenge and Perceptual–Cognitive subscales from the equation reduced the Nagelkerke R^2 from .27 to .06, indicating that 21% of the variation was uniquely explained by these two subscales, $\chi^2(2, N = 126) = 21.86$, $p < .001$. In contrast, removing the Direct Motor and Posthypnotic Amnesia subscales from the equation reduced the R^2 by less than 1%, indicating that these subscales did not uniquely explain any of the variation, $\chi^2(2, N = 126) = 0.78$, $p = .68$. In this regression, unlike all the previous ones, the contribution of general hypnotizability appears to be essentially negligible, at no more than about 5%.

Focusing next on the urge to respond, we found that this variable was well distributed, with a mean of 3.88 ($SD = 2.05$) near the midpoint of the 7-point Likert scale. The results were fully consistent with our predictions. The four subscales each showed significant univariate relationships (correlations) with *urge to respond* (rs = .29 to .48, $p < .01$). Moreover, in a multiple regression, the Perceptual–Cognitive subscale ($\beta = .38$, $p < .001$) and the Posthypnotic Amnesia subscale ($\beta = .27$, $p < .01$) emerged as the significant unique predictors. Of the 30% variance explained for *urge to respond,* 15% was attributable to specific building blocks measured by the Perceptual–Cognitive and the Posthypnotic Amnesia subscales, and the remaining half was attributable to general hypnotizability.

Motor posthypnotic suggestions. In Barnier and McConkey's (1999b) experiment, 108 participants were hypnotized and given the posthypnotic suggestion that upon awakening they would respond when they heard the word "experiment." Participants in the simple-suggestion condition were told that they would rub their right ear lobe; those in the complex-suggestion condition were told that they would hold their left arm out at shoulder height with the palm up, imagine feeling something heavy, and feel their hand and arm getting heavier and heavier and moving down until it reached the arm of the chair. Half the participants were given an accompanying suggestion for amnesia and the other half were not. As in Barnier and McConkey (1996), one response measure was whether participants made any behavioral response to the posthypnotic cue, and the other measure was participants' self-rating on a 7-point Likert scale of the extent to which they felt the urge to respond. Preliminary analysis showed that neither the complexity of the suggestion nor accompanying amnesia (which were randomly assigned, manipulated variables) interacted significantly with the four subscales in predicting either response measure. Therefore, the following regression models focus on only the subscales as predictors. Our predictions were the same as for the data set of Barnier and McConkey (1996).

Focusing first on the dichotomous behavioral response, we found that 41 (38.0%) participants responded to the posthypnotic suggestion and that 67 (62.0%) did not. The univariate relationships were inconsistent, as in the previous analysis of behavioral response to a verbal posthypnotic suggestion, again suggesting little role for general hypnotizability. At the multivariate level, the logistic regression analysis showed that the Perceptual–Cognitive subscale ($B = 0.69$, $p < .05$) and Posthypnotic Amnesia subscale ($B = 0.57$, $p < .05$) were the significant unique predictors, confirming our hypothesis that these specific components are particularly relevant for enacting posthypnotic suggestions. It is also interesting that the Direct Motor subscale played no role at either the univariate or multivariate levels, despite the motor nature of the response in this experiment. Removing the Perceptual–Cognitive and Posthypnotic Amnesia subscales from the equation reduced the Nagelkerke R^2 from .17 to .02, indicating that 15% of the variation was uniquely explained by these two subscales, $\chi^2(2, N = 108) = 12.79$, $p < .01$. In contrast, removing the Direct Motor and Motor Challenge subscales from the equation reduced the R^2 by an insignificant 3%, indicating that these subscales did not uniquely explain any of the variation, $\chi^2(2, N = 108) = 2.48$, $p = .29$. In this regression, general hypnotizability contributed virtually nothing.

Focusing next on the urge to respond, we found that this variable had a mean of 2.57 ($SD = 1.98$), below the midpoint of the 7-point Likert scale. With the exception of the Direct Motor subscale, univariate relationships (correlations) of the subscales with this variable were significant ($p < .05$). Of more importance, in the multiple regression the Perceptual–Cognitive subscale ($\beta = .38$, $p < .001$) and the Posthypnotic Amnesia subscale ($\beta = .27$, $p < .01$) emerged as the significant unique predictors. Of the 20% variance explained for *urge to respond,* 12% was attributable to specific building blocks measured by the Perceptual–Cognitive and the Posthypnotic Amnesia subscales, and the remaining 8% was attributable to general hypnotizability. These results were in accord with our predictions for posthypnotic suggestions. We found it surprising that the Direct Motor subscale again played no role at either the univariate or multivariate levels.

Discussion

Although these reanalyses based on our four subscales of hypnotic ability illustrate an interesting range of findings, the pervasive pattern was the expected one in which general hypnotizability and the specific abilities tapped by the subscales played complementary roles. At the multivariate level, we repeatedly found significant prediction by one or two subscales that we hypothesized would tap skills particularly relevant to the hypnotic suggestions. This uniquely explained variation is consistent with the hypothesis that specific skills, in addition to general hypnotizability, influence hypnotic responses in important ways.

The Perceptual–Cognitive subscale was the most consistently important unique predictor, closely followed by the Posthypnotic Amnesia subscale. In contrast, the Direct Motor and Motor Challenge subscales offered almost no unique predictive capacity. The SHSS:C and the HGSHS:A differ markedly in the profile of specific components that they tap, with the HGSHS:A chiefly tapping the direct motor and motor challenge components, and the SHSS:C tapping the perceptual–cognitive component. Hypnosis experimenters tend to find most interesting the complex perceptual–cognitive phenomena, even though hypnotic participants often report more compelling experiences in response to direct motor and motor challenge items (McConkey, Glisky, & Kihlstrom, 1989). Hence, although the SHSS:C and the HGSHS:A are both excellent measures of general hypnotizability, the SHSS:C may be a substantially better predictor of many hypnotic responses in experiments because of its much better measurement of the perceptual–cognitive component. For example, as predictors of color blindness, the total score on the HGSHS:A accounted for 15% of the variation, but the total score on the SHSS:C accounted for 40%. Accordingly, it would stand to reason that failure to replicate experimental effects in a sample screened on the HGSHS:A alone could readily stem from its lack of good measurement of the perceptual–cognitive component.

Although the Direct Motor and Motor Challenge subscales were not important unique predictors in the present set of reanalyses, they could emerge as specifically predictive in experiments investigating less complex hypnotic phenomena, such as suggested motor phenomena. Thus, an important goal for future research would be to demonstrate whether these specific motor components, which are more characteristic of the HGSHS:A than of the SHSS:C, are especially relevant in the experimental investigation of hypnotic effects other than those examined here.

In their study of amnesia for autobiographical memory, Barnier et al. (2004) emphasized the value of a multidimensional approach to indexing individuals' responses to hypnotic suggestions (see also Woody & McConkey, 2003). The present results using our subscales as predictors support and extend this idea of the multidimensionality of response, because they show different patterns of predictors for the three different types of memory measures used in Barnier et al's (2004) study. In particular, memory inaccessibility seems to be a difficult perceptual–cognitive enactment. The essence of this underlying subskill is the ability to create a perceived state of affairs that has the feeling of reality but is inconsistent with actual reality. Hence, individuals who find the targeted memory inaccessible may have approached the task by temporarily creating a reality in which the previously remembered event never happened. In contrast, lack of narrative thread in recounting the autobiographical memory is related to classic posthypnotic amnesia. Thus, poor narrative may stem from some aspect of the individual's recall strategy, such as premature termination of retrieval (Barnier et al., 2004). Finally, self-rated impressions of lack of clarity are related to only general hypnotizability, indicating that these ratings may reflect participants' sense of their overall responsiveness. In this fashion, the use of the subscales may help to diagnose the underlying nature of relatively complex hypnotic responses.

General Discussion

Our work advances previous factor analytic research with hypnosis scales in three important ways. First, it provides one novel solution to the difficulty confounds that have often challenged factor analyses of hypnosis scales (Balthazard & Woody, 1985), by using a sophisticated statistical approach to the analysis of dichotomous items (also recently applied to the HGSHS:A; Sadler & Woody, 2004). This approach, which overcomes difficulty artifacts (Bock et al., 1988), extends important earlier work by Tellegen and Atkinson (1976) and Kihlstrom (2001) in demonstrating that the multidimensionality of hypnosis scales is not an artifact; instead, multiple differentiable skills underlie hypnotic performance.

Second, compared with previous research, our work provides a more representative sampling of hypnotic item content, because we analyzed the combined items of the SHSS:C and the HGSHS:A, rather than the items of either scale alone. This approach has clear advantages. For example, posthypnotic amnesia could emerge as a separable factor because the analysis included measures of it from each scale. Similarly, although the HGSHS:A alone has too few perceptual–cognitive items to define such a factor clearly, the SHSS:C provides the additional items. Likewise, although the SHSS:C alone has too few direct motor and motor challenge items to define these factors clearly, the HGSHS:A provides the additional items. Thus, because the item content of the two scales is complementary, the combined item set overcomes the shortcomings of either scale alone, allowing us to derive four reasonably distinct subscales that represent special hypnotic abilities: Direct Motor, Motor Challenge, Perceptual–Cognitive, and Posthypnotic Amnesia.

Third, our reanalyses of experimental data with the subscales provide evidence that these specific dimensions are uniquely predictive, in patterns that differ sensibly for different hypnotic responses and experimental-dependent variables. Perry et al. (1992) noted about Weitzenhoffer and Hilgard's (1967) Stanford Profile Scales of Hypnotic Susceptibility, "It may be a sad commentary on the present state of research on hypnotizability that they have rarely been utilized" (p. 487). Until now, what may have been missing was concrete evidence that specific aspects of the profile of hypnotic subskills can be uniquely predictive of experimental findings and thereby potentially illuminating. We provide this evidence.

What Are the Implications of Multiple Hypnotizabilities?

Theory and research on intelligence serve as an interesting parallel with theory and research on hypnotizability. In work on intelligence, there has been a creative tension between ideas about the importance of a unifying, general construct of intelligence and ideas about the importance of multiple, specific types of intelligence. For example, Gottfredson (1998) argued that a global dimension of cognitive competence has proven to be "the single most effective predictor known of individual performance at school and on the job" (p. 24). In contrast, Gardner (1983) attracted attention with his speculation that people possess relatively independent, specific modules of intelligence. One reason that the notion of *multiple intelligences* has evoked interest is its apparent implications for learning and education. Gardner (1983; Gardner, Kornhaber, & Wake, 1996) argued that teaching should be specifically tailored toward the individual student's profile of intellectual strengths. Likewise, among clinicians there is interest in the possibility of tailoring the use of hypnosis in therapy to the client's specific profile of subskills—what Brown (1992) called a "specificity theory of hypnotic responsiveness" (p. 456) and Barber (1991) labeled the "locksmith model" (p. 241).

Nonetheless, in both education and clinical work, this idea of tailoring is limited because of psychometrically problematic assessments of the multiple subskills involved. For example, Gardner (1983) suggested that multiple intelligences require no formal assessment but can be assessed informally by observant teachers, parents, and so forth. Likewise, O'Connell, Orne, and Shor (1966) proposed a brief, flexible clinical measure of hypnotizability, the Diagnostic Rating Scale, which consisted of just a single item for each of four different types of skill, with the content of each item chosen by the hypnotist. In contrast, we would argue that if special abilities exist, then they must be measured with psychometrically sound, standardized multiple items.

Our perspective on hypnotizability is that each distinguishable ability involves the combination of general hypnotizability with a more specific, unique component (see also Barnier & McConkey, 2004; McConkey & Barnier, 2004; Woody & McConkey, 2003). What might be the underlying nature of these specific components? Woody and Szechtman (2000) advanced the idea that hypnosis modifies underlying "feelings of knowing," or subjective convictions, and this view is consistent with a long tradition that has emphasized the subjective underpinnings of hypnotic responses (e.g., Sutcliffe, 1960, 1961). It is possible that the four different components and associated subscales that we have identified arise because the feelings of knowing (or subjective phenomena) that underlie them are qualitatively distinct from one another. For example, in hypnotic responses of the perceptual–cognitive type, the crucial subjective experience appears to be one of a feeling of external reality in the face of an inconsistent actual reality (see also Tellegen, 1978/1979)—a hallucination is the prototype of this kind of experience. In contrast, the crucial subjective experience in hypnotic responses of the motor challenge type appears to be a feeling of involuntariness or ineffectiveness of one's will (Woody & Sadler, 1998)—trying to do something but failing is the prototype of this kind of experience. An individual experiencing a hallucination does not have the conviction of involuntariness, nor does an individual who tries to move but cannot move necessarily have the conviction of external reality. Given these differences, further analysis of the underlying nature of the specific building blocks of hypnotic performance is a critical and promising area for future research (for a relevant conceptual analysis, see Woody & McConkey, 2003).

An Alternative View

Although we have commented on the multidimensionality of hypnotizability in terms of distinct, stable subtraits, another view is possible. We could reinterpret aspects of the factor analysis we earlier dismissed as anomalous, lending them instead more importance and spelling out a somewhat competing perspective. In particular, it is remarkable that certain pairs of seemingly similar items appeared to function quite differently in the SHSS:C versus the HGSHS:A. The perceptual–cognitive items in the HGSHS:A (fly hallucination and posthypnotic suggestion) behaved like direct motor items, whereas the motor challenge items in the SHSS:C (arm rigidity and arm immobilization) behaved like perceptual–cognitive items. This pattern indicates that specific components tapped by at least some hypnotic suggestions may be malleable and that their malleability may be heavily influenced by the context of the suggestion. That is, when arm rigidity and arm immobilization are preceded by and embedded among perceptual–cognitive items, as they are in the SHSS:C, participants may approach them in a perceptual–cognitive manner, unlike how they approach them in the heavily ideomotor HGSHS:A. Likewise, when insect hallucination and posthypnotic suggestion are preceded by and embedded among motor suggestions, as they are in the HGSHS:A, participants may approach them in a direct motor manner, unlike how they approach them in the predominantly perceptual–cognitive SHSS:C.

It is possible to carry this argument further by noting that in all the reanalyzed studies, the experimental behaviors of interest followed the SHSS:C, in the same session. Thus, the SHSS:C may have established a particular context for participants' approach to the suggestions in the experiment proper. In short, prior suggestions might affect the strategies that participants adopt, serving, in effect, as suggestions for how to respond to subsequent task demands (see also Sheehan & McConkey, 1982). If so, the specific factors isolated in the present study may represent aspects of a contextually dependent, strategic flexibility, overlaid on a more stable foundation of general hypnotizability. Further determination of whether specific factors are better conceptualized in this way or as stable traits requires a research design that temporally, and perhaps psychologically, separates the administration of various standardized hypnotic items from the experimental behaviors to be predicted and explained.

Finally, our work has interesting implications for the construction of scales to assess hypnotizability. Continuing to emphasize the SHSS:C as the gold standard of hypnotizability measurement (e.g., Perry et al., 1992) implies a lack of interest in measuring separable component skills, because the SHSS:C alone cannot resolve these skills from one another well. Moreover, up to the present, hypnosis-scale construction generally has not followed a clear facetlike strategy of sampling from the universe of possible items. The design of future hypnosis scales—especially with regard to the sampling of item content—should be informed by a fuller, more multidimensional understanding of the relevant underlying individual differences. We hope the present work contributes to that goal.

References

Arbuckle, J. L., & Wothke, W. (1999). *Amos 4.0 user's guide.* Chicago: SmallWaters.

Balthazard, C. G. (1993). The Hypnosis Scales at their centenary: Some fundamental issues still unresolved. *International Journal of Clinical and Experimental Hypnosis, 41,* 47–73.

Balthazard, C. G., & Woody, E. Z. (1985). The "stuff" of hypnotic performance: A review of psychometric approaches. *Psychological Bulletin, 98,* 283–296.

Barber, J. (1991). The locksmith model: Accessing hypnotic responsiveness. In S. J. Lynn & J. W. Rhue (Eds.), *Theories of hypnosis: Current models and perspectives* (pp. 241–274). New York: Guilford Press.

Barnier, A. J. (2002). Posthypnotic amnesia for autobiographical episodes: A laboratory model of functional amnesia? *Psychological Science, 13,* 232–237.

Barnier, A. J., & McConkey, K. M. (1996). Action and desire in posthypnotic responding. *International Journal of Clinical and Experimental Hypnosis, 44,* 120–139.

Barnier, A. J., & McConkey, K. M. (1998a). Posthypnotic responding away from the hypnotic setting. *Psychological Science, 9,* 256–262.

Barnier, A. J., & McConkey, K. M. (1998b). Posthypnotic responding: Knowing when to stop helps to keep it going. *International Journal of Clinical and Experimental Hypnosis, 46,* 204–219.

Barnier, A. J., & McConkey, K. M. (1999a). Hypnotic and posthypnotic suggestion: Finding meaning in the message of the hypnotist. *International Journal of Clinical and Experimental Hypnosis, 47,* 192–208.

Barnier, A. J., & McConkey, K. M. (1999b). Posthypnotic suggestion, response complexity, and amnesia. *Australian Journal of Psychology, 51,* 1–5.

Barnier, A. J., & McConkey, K. M. (2004). Defining and identifying the highly hypnotisable person. In M. Heap, R. Brown, & D. Oakley (Eds.), *High hypnotisability: Theoretical, experimental and clinical issues* (pp. 30–60). London: Brunner-Routledge.

Barnier, A. J., Wright, J., & McConkey, K. M. (2004). Posthypnotic amnesia for autobiographical episodes: Influencing memory accessibility and quality. *International Journal of Clinical and Experimental Hypnosis, 52,* 260–279.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46,* 443–459.

Bock, R. D., Gibbons, R. D., & Muraki, E. (1988). Full-information factor analysis. *Applied Psychological Measurement, 12,* 261–280.

Brown, D. P. (1992). Clinical hypnosis research since 1986. In E. Fromm & M. R. Nash (Eds.), *Contemporary hypnosis research* (pp. 427–458). New York: Guilford Press.

Cox, R., & Barnier, A. J. (2003). Posthypnotic amnesia for a first romantic relationship: Forgetting the entire relationship versus forgetting selected events. *Memory, 11,* 307–318.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Gardner, H., Kornhaber, M. L., & Wake, W. K. (1996). *Intelligence: Multiple perspectives.* Fort Worth, TX: Harcourt Brace College.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Gottfredson, L. S. (1998). The general intelligence factor. *Scientific American Presents—Exploring Intelligence, 9,* 24–29.

Hendrickson, A. E., & White, P. O. (1966). A method for the rotation of higher-order factors. *British Journal of Mathematical and Statistical Psychology, 19,* 97–103.

Hilgard, E. R. (1965). *Hypnotic susceptibility.* New York: Harcourt, Brace & World.

Hilgard, E. R., Crawford, H. J., Bowers, P. G., & Kihlstrom, J. F. (1979). A tailored SHSS:C, permitting user modification for special purposes. *International Journal of Clinical and Experimental Hypnosis, 27,* 125–133.

Hull, C. L. (1933). *Hypnosis and suggestibility: An experimental approach.* New York: Appleton-Century.

Kihlstrom, J. F. (2001, August). *Foxes and hedgehogs and hypnosis, oh my!* Address presented at the 109th Annual Convention of the American Psychological Association, San Francisco.

Kihlstrom, J. F., & Register, P. A. (1984). Optimal scoring of amnesia on the Harvard Group Scale of Hypnotic Susceptibility, Form A. *International Journal of Clinical and Experimental Hypnosis, 32,* 51–57.

Mallard, D., & Bryant, R. A. (2001). Hypnotic color blindness and performance on the Stroop Test. *International Journal of Clinical and Experimental Hypnosis, 49,* 330–338.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology, 70,* 810–819.

McConkey, K. M., & Barnier, A. J. (2004). The highly hypnotisable person: Unity and diversity in behaviour and experience. In M. Heap, R. Brown, & D. Oakley (Eds.), *High hypnotisability: Theoretical, experimental and clinical issues* (pp. 61–84). London: Brunner-Routledge.

McConkey, K. M., Glisky, M. L., & Kihlstrom, J. F. (1989). Individual differences among hypnotic virtuosos: A case comparison. *Australian Journal of Clinical and Experimental Hypnosis, 17,* 131–140.

Muthén, B. (1978). Contributions to factor analysis of dichotomized variables. *Psychometrika, 43,* 551–560.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'Connell, D. N., Orne, M. T., & Shor, R. E. (1966). A comparison of hypnotic susceptibility as assessed by diagnostic ratings and initial standardized test scores. *International Journal of Clinical and Experimental Hypnosis, 14,* 324–332.

Perry, C., Nadon, R., & Button, J. (1992). The measurement of hypnotic ability. In E. Fromm & M. R. Nash (Eds.), *Contemporary hypnosis research* (pp. 459–490). New York: Guilford Press.

Sadler, P., & Woody, E. (2004). Four decades of group hypnosis scales: What does item response theory tell us about what we've been measuring? *International Journal of Clinical and Experimental Hypnosis, 52,* 132–158.

Sheehan, P. W., & McConkey, K. M. (1982). *Hypnosis and experience: The exploration of phenomena and process.* Hillsdale, NJ: Erlbaum.

Shor, R. E., & Orne, E. C. (1962). *Harvard Group Scale of Hypnotic Susceptibility.* Palo Alto, CA: Consulting Psychologists Press.

Shye, S., Elizur, D., & Hoffman, M. (1994). *Introduction to facet theory: Content design and intrinsic data analysis in behavioral research.* Thousand Oaks, CA: Sage.

Sutcliffe, J. P. (1960). "Credulous" and "skeptical" views of hypnotic phenomena: A review of certain evidence and methodology. *International Journal of Clinical and Experimental Hypnosis, 8,* 73–101.

Sutcliffe, J. P. (1961). "Credulous" and "skeptical" views of hypnotic phenomena: Experiments on esthesia, hallucination, and delusion. *Journal of Abnormal and Social Psychology, 62,* 189–200.

Szechtman, H., Woody, E., Bowers, K. S., & Nahmias, C. (1998). Where the imaginal appears real: A positron emission tomography study of auditory hallucinations. *Proceedings of the National Academy of Sciences, USA, 95,* 1956–1960.

Tellegen, A. (1978/1979). On measures and conceptions of hypnosis. *American Journal of Clinical Hypnosis, 21,* 219–237.

Tellegen, A., & Atkinson, G. (1976). Complexity and measurement of hypnotic susceptibility: A comment on Coe and Sarbin's alternative interpretation. *Journal of Personality and Social Psychology, 33,* 142–148.

Thissen, D. (1988). *MULTILOG: Multiple, categorical item analysis and test scoring using item response theory.* Chicago: Scientific Software.

Wainer, H. (1976). Estimating coefficients in linear models: It don't make no nevermind. *Psychological Bulletin, 83,* 213–217.

Waller, N. G., Tellegen, A., McDonald, R. P., & Lykken, D. T. (1996). Exploring nonlinear models in personality assessment: Development and preliminary validation of a negative emotionality scale. *Journal of Personality, 64,* 545–576.

Weitzenhoffer, A. M. (1997). Hypnotic susceptibility: A personal and historical note regarding the development and naming of the Stanford scales. *International Journal of Clinical and Experimental Hypnosis, 45,* 126–143.

Weitzenhoffer, A. M., & Hilgard, E. R. (1962). *Stanford Hypnotic Susceptibility Scale, Form C.* Palo Alto, CA: Consulting Psychologists Press.

Weitzenhoffer, A. M., & Hilgard, E. R. (1963). *Stanford Profile Scales Hypnotic Susceptibility Scales: Forms I and II.* Palo Alto, CA: Consulting Psychologists Press.

Weitzenhoffer, A. M., & Hilgard, E. R. (1967). *Revised Stanford Profile Scales Hypnotic Susceptibility Scales: Forms I and II.* Palo Alto, CA: Consulting Psychologists Press.

Wilson, D. T., Wood, R., & Gibbons, R. (1991). *TESTFACT: Test scoring, item statistics, and item factor analysis.* Chicago: Scientific Software International.

Woody, E. Z., & McConkey, K. M. (2003). What we don't know about the brain and hypnosis, but need to: A view from the Buckhorn Inn. *International Journal of Clinical and Experimental Hypnosis, 51,* 309–338.

Woody, E., & Sadler, P. (1998). On reintegrating dissociated theories: Comment on Kirsch and Lynn (1998). *Psychological Bulletin, 123,* 192–197.

Woody, E., & Szechtman, H. (2000). Hypnotic hallucinations: Toward a biology of epistemology. *Contemporary Hypnosis, 17,* 4–14.

Received February 23, 2004
Revision received September 28, 2004
Accepted October 19, 2004 ■