

SELF-REPORT SCALES OF HYPNOTIC DEPTH¹

CHARLES T. TART

University of California, Davis

Abstract: Hypnotic "depth" is a complex theoretical concept for which no direct measures exist at present, but which is reflected in various experiential and behavioral responses to hypnosis. Thirty-five Ss were asked to scale their depth of hypnosis on a 10-point scale after each suggestibility test item on the Stanford Hypnotic Susceptibility Scale, Form C (Weitzenhoffer & Hilgard, 1962). These self-reports were highly correlated with measures of hypnotic behavior and experience. Instructions to report depth immediately and without thinking produced reports which correlated somewhat better with the other measures than instructions to consciously make a best estimate. This self-report scale promises to be highly useful in studies of hypnosis.

Since the beginnings of the scientific study of hypnosis, it has been recognized that there are degrees of hypnosis, that some Ss respond more profoundly than others, and that some phenomena are possible only in "better" Ss. Exactly what was meant by the concept of "depth" thus implied for hypnosis has been unclear in the hypnosis literature. Particularly, hypnotic *susceptibility* and hypnotic *depth* have frequently been confused. Hypnotic susceptibility refers to the degree of responsiveness to hypnotic suggestions administered under standardized conditions (Hilgard, 1965), and in comparing Ss' degrees of hypnotic susceptibility we are comparing their overall responsiveness to the assessment procedure. Hypnotic depth, on the other hand, refers to a momentary state of *S* along some dimension of "profundity" of the hypnotic state. Throughout a hypnotic session, depth can thus vary anywhere between zero and the maximum attainable for a given *S*. While hypnotic depth has been conceptualized in a variety of ways in the literature, it is sufficient here to note that it has always been seen as affecting *S*'s responses and experiences while hypnotized. A standard assessment procedure, such as Weitzenhoffer and Hilgard's Stanford Hypnotic Susceptibility Scale: Forms A and B (SHSS: A and B, 1959) and Form C (SHSS:C, 1962), may be confounded by changes in depth occurring throughout the assessment of an individual *S*. More important, in studies using the presence or absence of the

Manuscript submitted July 18, 1969.

¹This study was partially supported by grant MH-13312 from the National Institute of Mental Health. I also wish to thank Mr. David Perry and Mrs. Lois Dick for assistance in running subjects and analyzing data, respectively.

hypnotic state as an independent variable, variations in depth in a given *S* throughout the course of the experiment, as well as such variations across *Ss*, could seriously confound results, even to the point of increasing variance sufficiently to obscure genuine effects.

LeCron (1953) reported that he was able to measure hypnotic depth as frequently as he desired, and with little effort, by simply instructing his *Ss* that they would be able to scale their hypnotic depth on a 100-point scale. He defined this scale to them somewhat vaguely (the values are shown in Table 1), but told them that their "subconscious minds" always knew how deeply hypnotized they were, and that whenever he asked them what their depth was a number would flash into consciousness from their subconscious minds and they would report it immediately. Although he reported few details of his studies, LeCron stated that these self-reports of hypnotic depth correlated extremely well with his clinical estimates of *Ss*' depth, and could, for most purposes, replace other measures of depth, such as particular hypnotic phenomena.

LeCron's report led to a number of other *Es* developing various methods for self-report of hypnotic depth (or hypnotic "state" in some studies). Names have been given to these scales according to where the work was done, and their main characteristics are summarized in Table 1. All these scales require *S* to report the depth of his hypnotic state when questioned about it, except the Hypnotic Depth Indicator (HDI), which requires *S* to adjust a dial or move his hands continuously along a ruled scale as his hypnotic depth increases or decreases.²

Since these six scales differ along many important dimensions, it is difficult to compare them. The first such dimension is the concept of hypnotic depth held by *E*, and the degree to which this was communicated to *Ss* explicitly (the formal instructions) and implicitly (demand characteristics). The second is the degree of definition of the scale. At one extreme (the HDI and Discrete scales) only awake and deep are defined, while at the other (the North Carolina Scale, Tart,

² This scale was first used by Field in an unpublished report (1963), and required *S* to move a dial along a linear sliding scale. It was modified by Evans so that *S* could move the dial continuously, when appropriate, around a 12-inch diameter "clockface" with his eyes closed. Except for two early studies by Field (1965, 1966), a later, slightly modified version has been used in all studies (Evans, 1970; Evans & Orne, 1965; Field, Evans, & Orne, 1965; Orne & Evans, 1966), and is called in these reports the Hypnotic Depth Indicator (HDI). The Discrete scale, discussed below, is defined in the same fashion, but is usually presented to *S* only once after each termination of hypnosis. It asks *S* to rate the deepest level achieved (Evans & Orne, 1965).

TABLE 1
CHARACTERISTICS OF SELF-REPORT SCALES OF DEPTH

Scale name	Defined values	Type of answer	References
LeCron	0 = awake 1-20 = light 20-40 = medium 40-60 = deep 60-80 = plenary 80+ = stuporous	Instant	Hatfield, 1961; LeCron, 1953
North Carolina	0 = waking 1-12 = relaxed, detached, ideomotor movements 20 = analgesia 25 = dreams 30 = amnesia, mental quiet, very high suggestibility 40 = all effects completely real 50+ = mind sluggish	Instant	Tart, 1962; 1963; 1967
Brief Stanford	0 = wide awake 1 = borderline 2 = light 3 = medium 4 = deep	Instant	Hilgard & Tart, 1966; Tart, 1966a; Tart & Hilgard, 1966
Long Stanford	0 = wide awake 1 = borderline 2 = light 5 = deep 10 = very deep, very high suggestibility	Instant & Deliberate	Larsen, 1965; Tart, 1966b Present paper
Harvard Discrete	1 = awake 10 = as deep as possible	Deliberate	O'Connell, 1964
Harvard Continuous	1-10 = awake to deep as possible	Deliberate, Continuous	Field, 1966; Orne & Evans, 1966

1962) particular hypnotic phenomena that Ss can expect to experience at various scale points are mentioned in the instructions. The third is the amount of previous experience with hypnosis Ss may have had before being required to use the scale: if they have not had such experience, they are scaling on the basis both of hypnotic effects and of transitional effects of adapting to a novel situation. Fourth, the scales differ in whether Ss are asked to make conscious, deliberate estimates of their hypnotic depth or whether they are told that the answers will come automatically, instantly, in response to E's query. The fifth dimension is the number of state reports obtained during the hypnotic session, varying from a low of one (the Discrete scale) to obtaining a state report following every particular hypnotic phenomenon suggested (the North Carolina Scale) to continuous self-monitoring of subjective depth (HDI).

In attempting to assess the validity of self-report scales of hypnotic depth, all published studies to date have compared self-report measures for their Ss against some behavioral measure of hypnotic *susceptibility*, usually some version of the standardized Stanford scales (Weitzenhoffer & Hilgard, 1959, 1962), or against particular hypnotic phenomena which are traditionally believed to require a certain minimum depth before they can be experienced well. The procedure of almost all experimental studies of hypnosis apparently assumes that depth stays constant after induction. If, however, depth does fluctuate rapidly for at least some Ss, then obtaining only a few self-reports of depth might well result in atypical values. Consequently, correlations of self-reported depth with overall susceptibility could not reach very high figures.

This consideration is particularly important in evaluating the published reports. Two studies, one of the original LeCron scale (Hatfield, 1961) and one of the HDI (O'Connell, 1964), obtained only two reports and one report per S, respectively, during the hypnotic state. Such designs are extremely susceptible to atypical reports from the S. Both studies reported *statistically* significant correlations between the self-report measure and the susceptibility measure (.32 for Hatfield, .55 for O'Connell), but neither of these was high enough to uphold LeCron's original contention that self-report measures could replace ordinary measures of depth or be practically useful.

The North Carolina (Tart, 1962, 1963, 1967), Brief Stanford (Hilgard & Tart, 1966; Tart, 1966a; Tart & Hilgard, 1966), and Long Stanford Scales (Larsen, 1965; Tart, 1966b) were designed to come much closer to LeCron's original technique, viz., to require frequent reports and instructions to Ss that the reports would just flash into

their minds instantly, rather than being something they had to deliberately estimate. With these scales, state reports were obtained immediately following the induction procedure and following every hypnotic test item.

One study with the North Carolina Scale (Tart, 1963) allows statistical comparison of a shortened version of the scale (only the first 30 points defined, rather than 50) and hypnotic phenomena: self-reports by Ss obtained just before item administration correctly predicted hypnotic dreaming 100% of the time and posthypnotic amnesia 82% of the time.

There have been three studies with the Brief Stanford Scale. In the most extensive one (Hilgard & Tart, 1966), Ss were tested in a variety of conditions, all of which included administration of the SHSS:C. There were a total of 220 Ss tested during the various hypnotic conditions. The depth reports on this 0-4 scale were obtained between all suggestibility test items, as well as immediately after the induction. Correlations from the various conditions between the SHSS:C score and the *mean* self-report score were .67, .68, .65, .75, and, quite surprisingly, in one group, $-.01$. Thus, the scale was generally quite significantly related to hypnotic behavior. Many of these Ss were also hypnotized on 2 different days, and the correlation between their mean self-reported depth from day 1 to day 2 ranged from .81 to .99 over the various groups, indicating that the self-report scaling of hypnotic depth has the same degree of reliability for this scale as does SHSS:C. Further, the *initial* state report, obtained just after the end of induction, was found to be highly predictive of subsequent hypnotic behavior: if S reported 0, indicating he felt not hypnotized at all, it was extremely rare for him to pass more than one or two of the easier suggestibility test items (mean score of 1.3 on the SHSS:C vs. a mean score of 5.7 for those who reported a state of 2 or more, "definitely hypnotized").

The Brief Stanford Scale has also been found to correlate between .40 and .47 in various groups with the type of hypnotic dream S has when the type of dream is also scaled in terms of the profundity of the experience (Tart, 1966a). This finding has been confirmed by Honorton³ and by Parker.⁴

This paper will report the results of a study of the Long Stanford Scale. This study was designed to: (a) examine the way Ss use state (depth) reports *per se*; (b) compare conscious, deliberate answers with answers which S was told would come from the "subconscious"

³ C. Honorton, personal communication, 1968.

⁴ A. Parker, personal communication, 1969.

part of his mind; (c) investigate the relationship between state reports and measures of hypnotic *behavior*; (d) examine the relationship between state reports and various aspects of the *experience* of hypnosis. Frequent reports were obtained from Ss; they all had previous hypnotic experience, having had the Harvard Group Scale of Hypnotic Susceptibility, Form A (HGSHS:A) of Shor and E. Orne (1962) before they were run in this study; there was a fair degree of scale definition. These conditions are much closer to LeCron's original conditions than those of other reported studies and thus constitute a more adequate evaluation of the potentials of self-report scaling of hypnotic depth.

METHOD

Subjects

Thirty-five undergraduate males were tested with the procedures of this study. All had the following background with *E*: (a) *E* spoke briefly about hypnosis in their undergraduate psychology courses; (b) *E* administered the HGSHS:A to them; and (c) *E* subsequently asked for volunteers for an individual hypnosis experiment. The Ss were told that their level of response to the HGSHS:A was not important in volunteering for the individual experiments.

Procedure

When each *S* reported for the experimental session, he was told that the purpose of the experiment was "... to find out things about the nature of *individual patterns of experience* in a widely used and standardized hypnotic procedure," and "... to investigate how a self-report scale of how hypnotized you feel relates to your individual experiences and responses." The *S* was then given the instructions for using the state report scale: after he indicated he understood them, he put on headphones and a modified version (Hilgard & Tart, 1966) of SHSS:C was administered with tape-recorded procedures. State reports were requested following the induction procedure and following every item on the tape-recorded SHSS:C.⁵ At the end of the SHSS:C testing, a detailed inquiry about the nature of *S*'s hypnotic *experiences* was carried out. Then *S* filled out Field's Inventory

⁵ There were three SHSS items (4, 7, and 9) that called for two state reports each. These were the Taste Hallucination which asks the *S* to hallucinate two tastes sequentially, the Age Regression to two separate ages, and the Anosmia to Ammonia with two sniffs of ammonia, with and without suggested anosmia. The state reports following each of these parts are designated by dividing the items into 4A and 4B, 7A and 7B, 9A and 9B, in subsequent graphs.

(1965) with respect to what he had just experienced. Finally *E* answered any questions *S* had about the experiment, but requested him not to speak to others about it until after the study was completed. These steps are described in more detail below.

State scale and instant vs. deliberate conditions. The instructions for using the state (self-report depth) scale were given to *S* on a typed sheet of paper which *S* read to himself. There were separate sheets for the Instant and for the Deliberate conditions, so *S* would not be aware that there was more than one condition. In order to prevent *E* from behaving differently in handing this set of instructions to *S*, an assistant randomized the instruction sheets beforehand and *E* simply handed *S* whatever instruction sheet was on top in a folder, without looking at it. This resulted in an *N* of 20 in the Instant condition and 15 in the Deliberate condition.

The first paragraph of the instructions was identical for each group and read as follows:

During your experience of hypnosis, I will be interested in knowing just how hypnotized you are. You will be able to tell me this by calling out a number from zero to ten, depending on how hypnotized you feel yourself to be. *Zero* will mean that you are awake and alert, as you normally are. *One* will mean a kind of borderline state, between sleeping and waking. *Two* will mean that you are lightly hypnotized. If you call out the number *five*, it will mean that you feel quite strongly and deeply hypnotized. If you feel really very hypnotized, you would call out an *eight* or *nine*. *Ten* will mean that you are very deeply hypnotized and you can do just about anything I suggest to you. Naturally, hypnosis can increase and decrease in depth from time to time, and that is the kind of thing I'll be interested in finding out from you.

[The instructions for the Instant condition then went on:]

Let me explain *how* you will report your state of hypnosis. When I ask "State?", you are to tell me the *first* number that pops into your mind, and this will represent your state at that time. We've found that this first impression is more accurate than if you stop to think about just what the number should be. This may seem a little hard at first, but it will get easy as you go along. Just call out the first number that pops into your mind when I ask, "State?" Remember the number zero means your normal waking state, five means quite strongly hypnotized, and ten means you are deep enough to experience just about anything I suggest. Just say the first number from zero to ten that comes into your mind when I ask, "State?" Let's try it now. State? (All *Ss* called out a zero at this time.) At various times during your experience I'll ask for your state, and you'll call out the first number that pops into your mind.

[The instruction for the Deliberate condition read:]

Let me explain *how* you'll report your state of hypnosis. When I ask "State?", I want you to estimate how deeply hypnotized you feel at that moment and call out this number to me. We've found that college subjects can give rather accurate estimates if they give it a few seconds thought

each time. This may seem a little hard at first, but it will get easy as you go along. Just call out your best estimate whenever I ask, "State?"

Remember, the number zero means your normal waking state, five means quite strongly hypnotized, and ten means you are deep enough to experience just about anything I suggest. Just call out your estimate between zero and ten whenever I ask, "State?" Let's try it now. State? (All Ss called out a zero at this time.) At various times during your experience I'll ask for your state, and you'll call out your estimate. You may take a maximum of fifteen seconds to make your estimate.

The Ss in both groups, then, were told that hypnosis varies in depth, that they can accurately scale it and report it to *E* when asked, and that the deeper they feel, the more hypnotic phenomena they can expect to experience.

SHSS:C. A modified version of the SHSS:C was used. Two items (~~3~~, mosquito hallucination, and ~~10~~, hallucinated voice) were dropped to save time. The entire induction and testing procedure was tape-recorded to be certain that Ss would be treated alike regardless of condition. State reports were requested on the tape, with a 15-second interval of silence to allow *S* to respond following each request. (In retrospect, this interval could have been reduced to about 8 seconds without cutting off any Ss' responses.) Thirteen state reports per *S* were obtained. Data on the shortened and tape-recorded version of the SHSS:C have been presented elsewhere (Hilgard & Tart, 1966).

Inquiry about hypnotic experiences. Following the termination of hypnosis, a long inquiry about the intensity and degree of involuntariness of each major suggested effect was conducted. This method for scoring *S*'s experience in response to the SHSS:C yields a score ranging from 0 to 116 and is called the SHSS:C-experiential score. Complete data on it will be presented elsewhere.⁶ The inquiry continued with questions about how well *S* thought his state reports reflected his hypnotic state, whether they were deliberate or automatic, and ended with administration of the Field's Inventory and debriefing. The *S* was reminded at several points during the inquiry that *he* was the expert on his experience so the outcome of the study depended on his reporting as accurately as possible.

Demand characteristics. The instructions given to Ss attempted to create expectations that while hypnotic depth might vary frequently, it might also remain relatively constant over long time periods, and that it was this degree of constancy or variation that was under study. It was hoped that this would maximize accuracy of reporting. The *E* was biased in expecting better results from the Instant condition, but, as he was blind to the existing condition, he could not systematically affect Ss' behavior.

⁶Tart, C. T. Measuring the experience of hypnosis. In preparation.

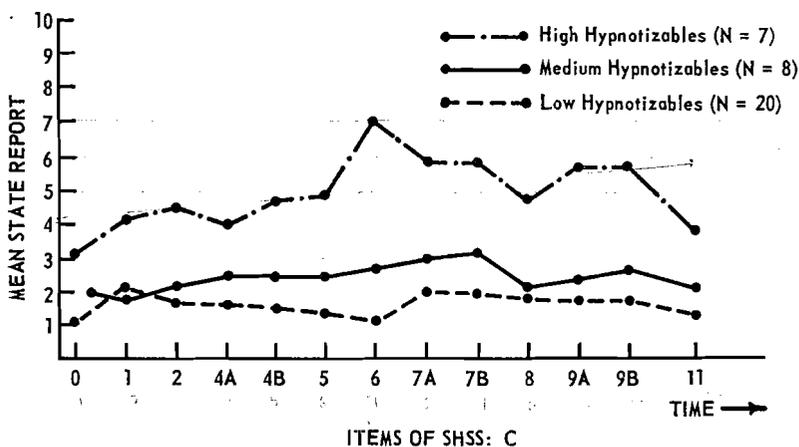


FIG. 1. Variation in mean state report over the course of testing for Ss of high, medium, and low hypnotic susceptibility.

RESULTS

Use of the state scale. State reports (13 per S), ranged from 0 to 8, but most of them (67%) were 1, 2, or 3, indicating borderline to moderately hypnotized.

One of the rationales for investigating state reports is the hypothesis that the depth of hypnosis may vary fairly rapidly with time. Time is here conveniently represented by the sequence of state reports obtained following each test item. For Ss as a whole, graphs of mean state reports over time are essentially linear for both Instant and Deliberate conditions. In a group of unselected Ss, however, there will always be several who are not hypnotized, or who are only very lightly hypnotized, and such Ss would serve as a stabilizing influence on such a graph. In Figure 1, state report for successive items has been plotted with Ss divided into high, medium, and low susceptibles on the basis of their SHSS:C-behavioral score (high Ss scoring 7 or more; medium Ss scoring 4, 5, or 6; and low Ss scoring 3 or less). Instant and Deliberate conditions are combined. The low susceptible Ss show a mean state report for the whole session of 1.2 (range: 0-4.5 per S); medium Ss, 2.5 (range: 1.7-4.6 per S); and high Ss, 5.0 (range: 2.6-6.6 per S). The differences between means of low and high Ss are significant ($t = 5.3, df = 25, p < .005, 1\text{-tailed}$), as are the differences between medium Ss and high Ss ($t = 3.9, df = 13, p < .005, 1\text{-tailed}$). The differences are only suggestive, however, between low Ss and medium Ss ($t = 1.65, df = 26, p < .10, 1\text{-tailed}$). Considering mean state reports over the course of the suggestibility testing, the curve for the low susceptibles is essentially flat, while there

Handwritten notes on the right margin:
 11.20
 7. 1.05
 5.94 = 1.14
 1.25
 1.14
 1.3
 1.05

is a suggestion of a rise for medium *Ss* and high *Ss*. Variability of the curves increases with items #6 (the dream), #7B (age regression), and #9B (sniff of ammonia without suggested anosmia). Statistical assessment of differences between the curves is not appropriate because of the confounding effects of: (a) time *per se*, (b) unequal increases in item difficulty over time, and (c) variations in term wording such that some items contain suggestions which might serve to deepen the hypnotic state while others have suggestions which might lighten it.⁷ This source of confusion, inherent in the use of SHSS:C, is regrettable, and the important possibility that suggestibility *testing* serves as a deepening or lightening procedure should be carefully noted for future investigations.

This analysis is sensitive only to long-term variability in state reports. Short-term variability, from one state report to the next, is also important. Considering all state reports from all *Ss*, most (50%) of the time there was no change from one report to the next, and 35% of the time the change was only 1 point. These figures can be misleading in the impression of stability they create, however, for, if state is important, a single state change of large magnitude might drastically affect the outcome of an experiment. If one examines the largest *absolute* (i.e., without regard to sign) change from one state report to the next for each *S*, one finds only 9% showed 0 as their largest change, 34% showed 1, 43% showed 2, 11% showed 3, and 3% showed a change of 4 points. As *Ss* only used 8 points of the defined scale, changes of 2 points or more seem of considerable magnitude, and such changes were shown at least once by 57% of *Ss*.

Another way of examining variability in the use of state reports is to consider the number of times each *S* showed a change from one state report to the next, regardless of the size of such change. Figure 2 shows the frequency of such changes: it was common for *Ss* to have as many as eight or nine changes in state report during the session.

Thus, as originally hypothesized, *Ss*' reports of the depth of their hypnotic state do vary significantly in the course of a hypnotic session when examined over short time intervals, but are fairly stable when *Ss*' ratings are averaged together over the entire session.

At the end of the hypnotic session, all *Ss* were asked what criteria they had used in making their state reports or, if they reported their responses had been automatic, what sorts of perceived changes in their experience seemed to go along with changes in state report. One *S* could report nothing; some *Ss* reported two or more criteria. Descrip-

⁷Because there are several reports per *S* in the graphed material, statistical comparison would not be legitimate.

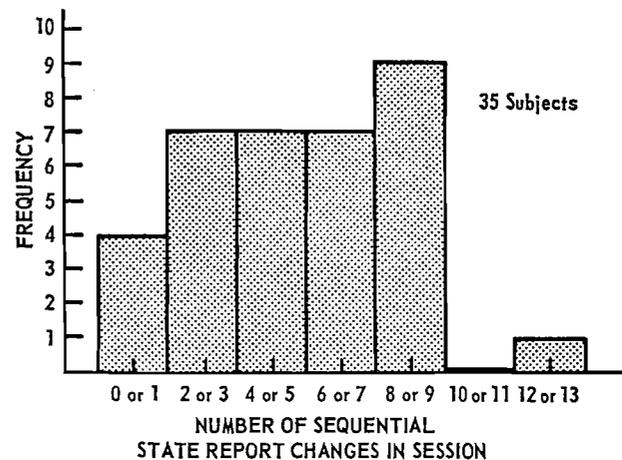


FIG. 2. Frequency of changes of sequential state reports.

tive categories for grouping answers, staying as close to Ss' descriptions as possible, were evolved. The 34 Ss who reported their criteria mentioned 56 criteria among them. Because of the small number of cases which result if these categories are subdivided according to condition, and because of the statistical difficulties caused by varying numbers of criteria per S, the following description is only suggestive.

The following sorts of criteria, all positively correlated with increasing state report, were mentioned in at least 10% of Ss' reports: (a) intensity of reaction to previous suggestibility test item, 21%; (b) feelings of drowsiness, 20%; (c) fading of the environment, 14%; (d) changes in body image or perceived body position, 12%; (e) relaxation, 11%; and (f) feelings of compulsiveness of responses, 11%. If categories (d) and (e) are combined with several other infrequent categories under the general category of bodily changes, 32% of the reports are accounted for.

The Ss were also asked how well they felt their state reports correlated with the depth of their hypnotic state. They were given fixed response categories of Very Well, Fairly Well, or Poorly. Almost all Ss (97%) reported they thought the correlation was fair to good.⁸

⁸I had attempted to run another group of Ss from advanced, rather than introductory, psychology courses in this study. Only two Ss made appointments for testing, so the results were discarded. It is interesting to note, however, that one S in this group described clearly varying experiential correlates of his state reports, but stated that he had little confidence in their accurately reflecting hypnotic depth because he knew a lot about psychology and knew that subjective reports were of no value!

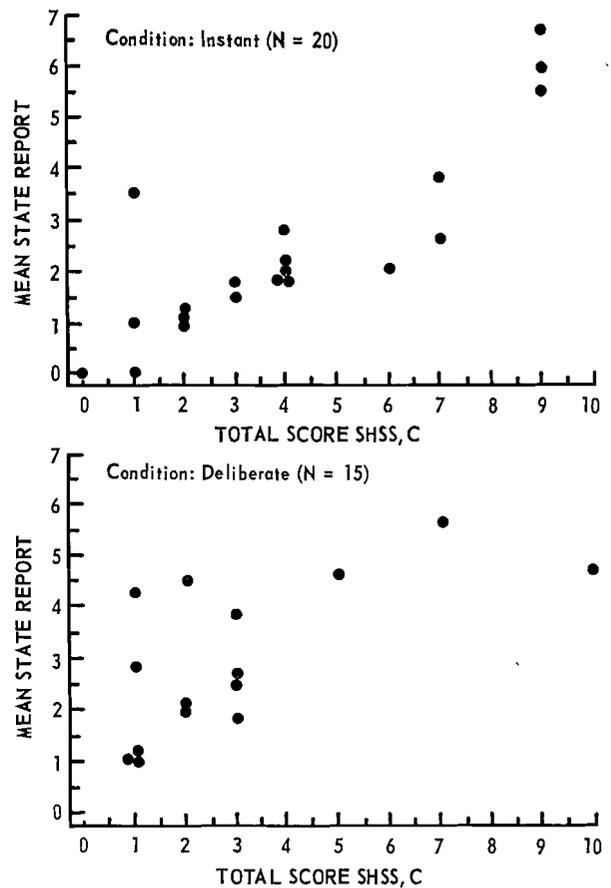


FIG. 3. Relationship between mean state reports and SHSS:C-behavioral scores for Instant and Deliberate conditions.

Relation of state reports to behavioral indices of hypnosis. The relation of state report to a behavioral index of hypnotic responsiveness was ascertained by correlating each S's mean state report with his total SHSS:C-behavioral score. Figure 3 shows scatter plots for these two measures for the Instant and Deliberate conditions separately. For the combined conditions, $r = .74$ ($p < .0005$, 1-tailed).

Because total scores on SHSS:C and mean state report are long-term measures, correlations between them are not an ideal way of assessing a relationship that may vary rapidly with time. A more sensitive measure is to compare the state report obtained just before each suggestibility test item was administered with whether or not that

item was passed or failed by the standard SHSS:C-behavioral scoring criteria. Item by item, the mean state report across all *Ss* was always greater for subsequently passed than for subsequently failed items: differences ranged from a low value of 1.12 to a high value of 2.76. Practically all differences for any given item⁹ were statistically significant by *t* test ($p < .005$, 1-tailed). Further, the more difficult an item, the greater was the average difference in preceding mean state report for passing and failing. The difficulty of each item, given by the percentage of *Ss* failing the item in the standardization data for SHSS:C-behavioral (Weitzenhoffer & Hilgard, 1962), correlates with the size of the mean state report difference ($p < .10$, 1-tailed).

Relation of state reports to experiential qualities of hypnosis. There is, of course, a significant relationship between the experience of the hypnotized *S* and his overt behavior. In the present sample, the correlation between the SHSS:C-experiential score and the SHSS:C-behavioral score is .77 ($p < .0005$, 1-tailed).

Figure 4 presents scatter plots of the relationship between the mean state report of each *S* and his SHSS:C-experiential score. The correlation is .76 ($p < .0005$, 1-tailed).

Another measure of *Ss*' experience of hypnosis is Field's Inventory, on which *Ss* answered 38 true-false questions about various phenomena they might have experienced during the course of the hypnosis session.¹⁰ As Field's Inventory was developed on the basis of what people generally report about being hypnotized, it is a more general measure of the hypnotic experience than the SHSS:C-experiential score, which is based on *Ss*' experiences in response to the specific suggestibility test items of the SHSS:C. The correlation is .66 ($p < .005$, 1-tailed) between mean state report and Field's Inventory.

A finer examination of this relationship can be made. An examination of the content of Field's Inventory questions indicates that, of the 38 questions, only 10 appear to be directly or implicitly suggested by the procedures of SHSS:C or of the earlier HGSHS:A.¹¹ Thus Field's Inventory score can be broken into a nonspecific score as well as a total score. Mean state report correlates .63 ($p < .005$, 1-tailed) with the nonspecific scores derived from the Inventory.

⁹ A test could not be carried out for item #2, as there was only one *S* failing this item.

¹⁰ *Ss* were obtained from three different psychology classes at various times: the third time *Ss* were run, the experimenter neglected to administer Field's Inventory, so the *N* for those analyses using Field's Inventory is 15 in the Instant condition and 10 in the Deliberate condition.

¹¹ These 10 items are numbers 4, 6, 9, 10, 14, 16, 24, 29, 30, and 34.

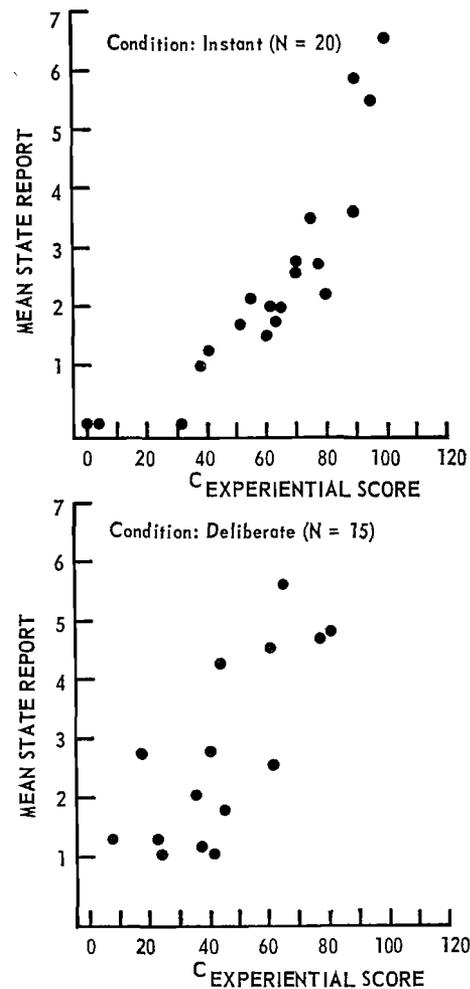


FIG. 4. Relationship between mean state reports and SHSS:C-experiential scores for Instant and Deliberate conditions.

Effects of experience on state reports. In the above analyses, *Ss* gained a fair amount of hypnotic experience in the course of making their state reports. In addition to the initial HGSHS:A, they had the induction of the SHSS:C and experience gained from the various SHSS:C items as the test proceeded. It is of interest to see how well state reports related to other aspects of hypnotizability with less hypnotic experience. This can be done by using only the first state report given by each *S*, i.e., the one immediately following the hypnotic

TABLE 2
RELATION OF INITIAL STATE REPORT TO VARIOUS MEASURES

Measure	Condition	<i>N</i>	<i>r</i>	<i>p</i> < ^a
SHSS:C-behavioral	Instant	20	.61	.005
	Deliberate	15	.58	.02
	Inst. + Dibr.	35	.56	.0005
SHSS:C-experiential	Instant	20	.79	.005
	Deliberate	15	.63	.01
	Inst. + Dibr.	35	.69	.0005
Field's Inventory	Instant	15	.72	.005
	Deliberate	10	.69	.005
	Inst. + Dibr.	25	.69	.0005

^a One-tailed values.

induction section of the SHSS:C, thus eliminating the experience of reactions to the suggestibility test items. Table 2 presents the correlation coefficients of initial state report with SHSS:C-behavioral, SHSS:C-experiential, and Field's Inventory across conditions. Although the correlations are generally reduced compared to those based on the full scale procedure, they are all statistically significant.

Instant vs. deliberate conditions. There was no clear difference in the usage of various values of the state reports in the two conditions. Mean state reports were 2.5 and 2.6 in the Instant and Deliberate conditions, respectively.

A slightly higher proportion of *Ss* in the Instant condition reported their state reports reflected the depth of their hypnotic state Very Well, but the difference was insignificant ($\chi^2 = 1.84$, $df = 2$, $p < .50$, 1-tailed).

The state-scale instructions for the Instant condition emphasized that a number would immediately pop into *S's* mind when he was asked "State?" and that, in the Deliberate condition, he would *consciously estimate* the best number. To check on the effectiveness of the instructions in the two conditions, *E* timed with a stop watch the delay between the request for state and *S's* responses. For the Instant condition the mean delay was 2.0 seconds, for the Deliberate condition, 4.2 seconds, a highly significant difference in the postulated direction ($t = 4.07$, $df = 33$, $p < .001$, 1-tailed).

The relationship between mean state report and SHSS:C-behavioral score was presented in Figure 3. Although the correlation is better for the Instant condition ($r = .85$, $p < .005$, 1-tailed) than for the Deliberate condition ($r = .67$, $p < .005$, 1-tailed), the difference between these two coefficients is not significant ($p = .11$, 1-tailed) by the

method of Johnson and Jackson (1959). The correlation between mean state report and SHSS:C-experiential scores is .84 for the Instant condition ($p < .005$, 1-tailed) and .75 ($p < .005$, 1-tailed) for the Deliberate condition. Again, the difference between the two correlation coefficients does not reach statistical significance ($p = .36$, 1-tailed).

The same pattern of correlation coefficients being higher, but not significantly so, for the Instant condition appears when comparing mean state report and Field's Inventory scores. The correlation for Instant condition is .69 ($p < .005$, 1-tailed) and for the Deliberate condition is .64 ($p < .02$, 1-tailed). The correlation between the Field's nonspecific scores with mean state report is .66 in the Instant condition ($p < .005$, 1-tailed) and .62 ($p < .02$, 1-tailed) in the Deliberate condition. Initial state reports, already presented in Table 2, also show this pattern.

In short, correlations with other measures of hypnotic performance are always higher for state reports from the Instant condition, but not significantly so.

DISCUSSION

The primary question of interest in the present study is the validity of the state scale, i.e., the degree to which the self-reports of hypnotic depth or state accurately reflect the dimension of hypnotic depth. To answer this question completely, one must have a clear definition of the theoretical concept: "depth" of hypnosis. Although such a concept is being developed,¹² this research work on self-report scales began on a more empirical basis. Thus, the question will be considered primarily from this basis, with only minimal theoretical conjecture here.

One must assume that there are dimensions of depth or profundity of the hypnotic state, that a given *S* may move along such dimensions from time to time, and that there are experiential correlates of position on each dimension that *S* can consciously perceive or unconsciously react to and report. That is, for a given *S* at a given position on the depth continuum there must be either steady state experiences or experiences in response to specific suggestions that *S* may use to judge his depth. Further, one must assume some commonality of the depth dimension across *S*s in order to have a scale that is useful for persons other than the individual for whom it was developed, i.e., that there is a great deal in common to the experience we label "hypnosis" across *S*s.

¹²Tart, C. T. On scaling the depth of altered states of consciousness. Paper privately distributed to attendees at the Conference on the Voluntary Control of Internal States, Council Grove, Kansas, 1969.

Given these assumptions about the nature of the self-report scale, we may then consider the question of its validity. Specifically, it is a question of *construct* validity. Hypnotic depth, as defined here, is a complex, poorly understood dimension which has no direct correlate, i.e., there is no particular phenomenon or behavior whose magnitude may be seen as directly reflecting hypnotic depth. We would expect all of our various measures (state reports, SHSS:C-behavioral, SHSS:C-experiential, and Field's Inventory) to be related to hypnotic depth, but none of them to be equivalent to it. Thus state reports should correlate highly with all these other hypnotic measures, but not perfectly. This is what was found: for SHSS:C-behavioral, SHSS:C-experiential, and Field's Inventory, the correlations with mean state report were .74, .77, and .66, respectively. There is no doubt that state reports are reflecting one or more significant dimensions of hypnotic behavior and experience, even if we cannot specify precisely what these dimensions are at present. LeCron (1953) originally suggested that self-reports of hypnotic depth could replace behavioral measures of depth. The truth of this proposal cannot be evaluated without a precise theoretical model of hypnosis and hypnotic depth which predicts how various behavioral and self-report measures will be related to depth: only then could one determine whether self-report measures generate data which fit theoretical predictions better than behavioral data. Lacking such a theory at present, the decision to use behavioral or self-report measures must be made with respect to specific studies only in terms of the empirical value of one or the other measure.

Assuming that reports with the Long Stanford Scale are probably as valid a measure of hypnotic depth as behavioral reports in terms of present knowledge (and assuming replication of these results), there are a number of advantages to adopting self-report measures as *the* measure of depth as opposed to conventional behavioral measures. First, it takes little time to teach *S* how to use the self-report scale. Second, it takes only seconds to get a report, and asking for these reports does not seem to disturb *S*. Third, ascertaining depth by behavioral measures at any given moment involves giving several suggestibility tests of varying degrees of difficulty: judging by present data, *S*'s depth might very well change in the time taken to administer these tests. Further, if one starts with tests that are too easy one risks boring *S* and taking a long time to find his response ceiling, while if one starts with tests that are too hard the consequent failure experiences might discourage *S* and/or decrease his depth.

To put this in its strongest form, in any experimental or therapeutic application where the depth of hypnosis would be important, the Long Stanford Scale is the method of choice for assessing depth.

Let us consider how *Ss* manage to scale their depth of hypnosis, i.e., how do they come up with the numbers they report? The cognitive bases on which they must scale would be: (a) previous knowledge of hypnotic phenomena, which Shor (1964) has shown to be rather good in a college student population; (b) knowledge of hypnosis gained in actual experimental participation; and (c) the instructions, both explicit and implicit, for scaling. Given this, it is interesting to note that with less conscious cognitive activity involved in the scaling and reporting process (Instant), there may have been stronger relationships to the other measures of hypnotic performance, although these did not reach acceptable levels of statistical significance. LeCron's notion that the answers in the Instant condition come directly from the "subconscious" is much too complex to test experimentally at present, but it is clear that the Instant condition instructions favor an intuitive rather than a cognitive type of response.

One might expect that *S's* reactions to suggestibility test items would be the primary determinants of *S's* state reports, such that the better he reacted, the higher a state he would give, regardless of whether the state reports were generated intuitively or cognitively. The data suggest, however, that this may not be too important a factor, for three reasons. First, only 21% of the replies to the question of how *Ss* scaled depth specifically mentioned reaction to the last item. Second, the state report obtained just *before* a suggestibility test item was highly predictive of passing or failing, as described earlier. Third, the state report following the induction procedure, but before *S* could react to any suggestibility test items, was also highly correlated with other measures of hypnosis. Many *Ss* also mentioned experiences, unrelated to specific suggestions, which increased or decreased in intensity with their depth and which were used for scaling. While reactions to passing and failing an item are a factor in scaling depth, the state reports are not merely a trivial reflection of passing or failing items.¹³

Given that state reports reflect the dimension of hypnotic depth, the present findings have an important implication for studies of "waking" suggestibility in comparison to hypnotic suggestibility. The high and medium susceptible *Ss* tended to show increases in their self-reports of hypnotic depth through the course of the suggestibility testing, indicating that the suggestibility test items may

¹³ We might also note that the *S's* criterion for passing and failing test items might not be the same as the *E's*. This was one reason for developing the experiential scoring of Form C. The higher relationships between self-reports and SHSS:C-experiential support this notion.

also function as deepening procedures. If so, it is possible that suggestibility test items given under "waking" conditions could function as hypnotic induction procedures for highly susceptible *Ss*: this finding has been reported earlier (Hilgard & Tart, 1966). Thus the simple procedure of *not* administering an induction in order to define the "waking suggestibility" group is misleading, because the group may well contain a few deeply hypnotized *Ss* before the testing is finished. Similarly, defining a group as "hypnotized" simply because *E* has gone through a traditional hypnotic induction procedure is also fallacious, as many *Ss* will not become hypnotized.

The present data also suggest that a new degree of precision in studies of hypnosis is possible. In many past studies of particular hypnotic phenomena, it is likely that the depth of hypnosis varied from *S* to *S*, or within the same *S* from one repetition to the next, in an unknown manner. This would result in high variability of results that could obscure important relationships. If state reports are used, *S* could be kept at the same relative depth level for each evocation of the phenomenon under study, thus eliminating a major source of variability. Another possible practical use of state reports is in studies attempting to maximize the hypnotic responsiveness of individual *Ss*: if a state report indicated *S* was probably not deep enough to experience a suggested phenomenon well, his state could be deepened before the phenomenon was suggested; thus a failure experience that might have adverse effects could be avoided. Similarly, the state report could be used to tell how well *S* responded to particular deepening techniques. Other uses will undoubtedly be suggested by future work.

All of these practical and theoretical uses are dependent, of course, on replication and expansion of the present study. Worthwhile improvements would include more extensive interviewing of *Ss* about how they scale depth and correlating state reports with individual hypnotic phenomena rather than with overall performance. Current research is being directed toward investigation of longer state scales, such as the North Carolina Scale, for use with *Ss* who can attain very deep hypnotic states.

REFERENCES

- EVANS, F. J. Simultaneous changes in skin potential and subjective estimates of depth of hypnosis. Paper presented at the meeting of the Eastern Psychological Association, Atlantic City, April, 1970.
- EVANS, F. J., & ORNE, M. T. Motivation, performance, and hypnosis. *Int. J. clin. exp. Hypnosis*, 1965, 13, 103-116.

- FIELD, P. B. An inventory scale of hypnotic depth. *Int. J. clin. exp. Hypnosis*, 1965, 13, 238-249.
- FIELD, P. B. Some self-rating measures related to hypnotizability. *Percept. mot. Skills*, 1966, 23, 1179-1187.
- FIELD, P. B., EVANS, F. J., & ORNE, M. T. Order of difficulty of suggestions during hypnosis. *Int. J. clin. exp. Hypnosis*, 1965, 13, 183-192.
- HATFIELD, ELAINE C. The validity of the LeCron method of evaluating hypnotic depth. *Int. J. clin. exp. Hypnosis*, 1961, 9, 215-221.
- HILGARD, E. R. *Hypnotic susceptibility*. New York: Harcourt, Brace & World, 1965.
- HILGARD, E. R., & TART, C. T. Responsiveness to suggestions following waking and imagination instructions and following induction of hypnosis. *J. abnorm. Psychol.*, 1966, 71, 196-208.
- JOHNSON, P. O., & JACKSON, R. W. B. *Modern statistical methods: Descriptive and inductive*. Chicago: Rand McNally, 1959.
- LARSEN, S. Strategies for reducing phobic behavior. Doctoral dissertation, Stanford University, 1965.
- LECRON, L. M. A method of measuring the depth of hypnosis. *J. clin. exp. Hypnosis*, 1953, 1 (2), 4-7.
- O'CONNELL, D. N. An experimental comparison of hypnotic depth measured by self-ratings and by an objective scale. *Int. J. clin. exp. Hypnosis*, 1964, 12, 34-46.
- ORNE, M. T., & EVANS, F. J. Inadvertent termination of hypnosis with hypnotized and simulating subjects. *Int. J. clin. exp. Hypnosis*, 1966, 14, 61-78.
- SHOR, R. E. The accuracy of estimating the relative difficulty of typical hypnotic phenomena. *Int. J. clin. exp. Hypnosis*, 1964, 12, 191-201.
- SHOR, R. E., & ORNE, EMILY C. *Harvard Group Scale of Hypnotic Susceptibility, Form C*. Palo Alto, Calif.: Consulting Psychologists Press, 1962.
- TART, C. T. A comparison of suggested dreams occurring in hypnosis and sleep. Unpublished masters thesis, University of North Carolina, 1962.
- TART, C. T. Hypnotic depth and basal skin resistance. *Int. J. clin. exp. Hypnosis*, 1963, 11, 81-92.
- TART, C. T. Types of hypnotic dreams and their relation to hypnotic depth. *J. abnorm. Psychol.*, 1966, 71, 377-382. (a)
- TART, C. T. Thought and imagery in the hypnotic state: Psychophysiological correlates. Paper presented at the meeting of the Amer. Psychol. Ass., New York, 1966. (b)
- TART, C. T. Psychedelic experiences associated with a novel hypnotic procedure, mutual hypnosis. *Amer. J. clin. Hypnosis*, 1967, 10, 65-78.
- TART, C. T., & HILGARD, E. R. Responsiveness to suggestions under "hypnosis" and "waking-imagination" conditions: A methodological observation. *Int. J. clin. exp. Hypnosis*, 1966, 14, 247-256.
- WEITZENHOFFER, A. M., & HILGARD, E. R. *Stanford Hypnotic Susceptibility Scale: Form C*. Palo Alto, Calif.: Consulting Psychologists Press, 1962.

Escalas de Autoevaluación de la Profundidad Hipnótica

Charles T. Tart

Resumen: La así llamada profundidad hipnótica constituye un complejo problema teórico para la cual no existen hoy día medidas directas sino

indirectas—vivencias y respuestas conductuales. Después de administrarseles a 35 sujetos experimentales la Escala Stanford Forma C de Weitzenhoffer y Hilgard (1962) se les solicitó a dichos sujetos que se autoevaluaran en una escala de 10 puntos las vivencias de profundidad hipnótica de cada uno de los ítems de la Escala. Esta autoevaluación correlacionó altamente con los resultados obtenidos por observación. Cuando se solicitó una autoevaluación inmediata—sin instrucciones previas—se produjeron correlaciones positivas aún mayores. La escala de autoevaluación promete ser bastante útil en futuros estudios sobre hipnosis.

Selbstberichtsskalen der hypnotischen Tiefe

Charles T. Tart

Abstrakt: Hypnotische "Tiefe" ist ein komplexer, theoretischer Begriff, für den zur Zeit keine direkten Masze vorhanden sind, doch spiegelt er sich in verschiedenen empirischen und benehmensmässigen Reaktionen auf Hypnose wider. 35 Vpn. wurden ersucht, die Tiefe ihrer Hypnose nach jedem einzelnen Teil des Stanford Hypnoseempfindlichkeitsmaszstabs, Form C (Weitzenhoffer und Hilgard, 1962), an einer 10-punktigen Skala anzugeben. Die selbstberichteten Masze wiesen ein hohes Korrelat zu Maszen des Benehmens und Erlebens in Hypnose auf. Instruktionen, die darauf gerichtet waren, die Tiefe unverzüglich und ohne Nachdenken zu berichten, erzeugten Masze, die sich etwas besser mit den andern Maszen, die durch Instruktionen erzielt waren, in denen die beste Schätzung gewissenhaft zu berichten war, korrelieren liessen. Dieser Selbstberichtsmaszstab verspricht, in Studien der Hypnose von höchstem Nutzen zu sein.