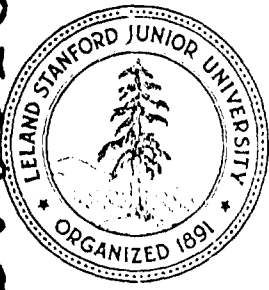


AD705132



STANFORD UNIVERSITY  
Department of Psychology

*Handwritten signature*

THE "TRUTH" ABOUT FALSE CONFESSIONS  
  
Christina Maslach

DDC  
RECEIVED  
MAY 12 1970  
B

Reproduced by the  
CLEARINGHOUSE  
for Federal Scientific & Technical  
Information Springfield Va. 22151

This document has been approved  
for public release and since its  
distribution is unlimited.

THE "TRUTH" ABOUT FALSE CONFESSIONS

Christina Maslach

Reproduction in whole or in part  
is permitted for any purpose of  
the United States Government.

ONR Technical Report: Z-01

April, 1970

## THE "TRUTH" ABOUT FALSE CONFESSIONS

Christina Maslach

Stanford University

### Abstract

One of the few experimental studies to focus upon variables which might play a role in police interrogation is Bem's "false confession" research. He showed that subjects come to believe that their false statements are true when emitted in the presence of a discriminative truth stimulus. In an attempted replication, the present study found evidence to support an alternative explanation of this finding, based upon decreased vigilance induced by the truth stimulus. In contrast, the lie stimulus was shown to promote more cautious responding as evidenced by better recall, greater confidence, slower reaction time, and a different pattern of physiological responsivity.

## THE "TRUTH" ABOUT FALSE CONFESSIONS<sup>1</sup>

Christina Maslach

Stanford University

One of the clearest examples of behavior control is found in the police interrogation room, where a suspect can often be induced to confess to a crime which he may or may not have committed. The magnitude of this control becomes particularly impressive when we realize that the consequences of such behavior are usually imprisonment or death. The manner in which confessions are obtained has important legal implications, since the presence of psychological coercion could influence their admissibility as trial evidence. Zimbardo's (1967) analysis of the various psychological techniques used to elicit confessions highlights the problems such methods pose for our system of justice.

In spite of the importance of this phenomenon, only one experimental study (Bem, 1966) has focused on it. Bem found that under certain conditions subjects come to believe in the false confessions that they have made. According to his self-attribution theory, an individual bases his subsequent attitudes and beliefs on the behavior that he has observed himself performing. Thus, if he sees himself making statements under circumstances previously associated with telling the truth, he will believe that the statements are true (even if they are, in fact, false). Because of both the practical and theoretical implications of this phenomenon, as well as the small sample employed, Bem's experiment deserves to be replicated and extended further. The present study attempts to duplicate Bem's paradigm for studying this self persuasion process,

which may be a critical feature of police interrogation.

Each subject first performs a word task in which he crosses out some words but not others. He is then trained to make true statements in the presence of one colored light (the "truth light") and to make false statements in the presence of a second light (the "lie light"). Finally, he is required to make statements about words which he previously did or did not cross out. Half of these "confessions" are false and half are true. Each of them is made in the presence of one of the two lights, and the subject is then asked to recall whether he actually had or had not crossed out the word. Following Bem, the prediction is that there will be an interaction between the lights and the confessions. False confessions made in the presence of the truth light will produce more recall errors than false confessions made when the lie light is on. On the other hand, true confessions emitted in the presence of the lie light will produce more recall errors than true confessions made in the presence of the truth light.

### Method<sup>2</sup>

#### Subjects

Fifty-five Stanford college students (thirty-six males and nineteen females) were asked to participate in a study on lie detection. Half were recruited from campus dormitories and were paid for their participation, while the others took part in the study to satisfy a requirement of the introductory psychology course.

#### Procedure

The experimental paradigm consists of four successive phases.

Phase I -- The subject engages in some activity which can be later used as the content of his "confessions."

Phase II -- The experimenter obtains information from the subject which will be used in the training procedure of the next phase.

Phase III -- Two discriminative stimuli are established for truth telling and lying.

Phase IV -- These stimuli are paired with "confessions" which the subject is required to make about his previous activity in Phase I. The major dependent measures are the subject's recall of what he actually did in Phase I, as well as his confidence in the accuracy of his recall.

In the first phase of the experiment, the subject was given a list of common nouns and an alphabetical guide which contained fifty of these words. He was then asked to cross out each word on the list that also appeared in the alphabetical guide. In the second part of the study, the subject completed a 50-item self-information form, which asked such questions as "What is your favorite music?" and "What did you eat for dinner last night?"

For the remainder of the study, the experimenter communicated with the subject (who was in a sound-proof chamber) via an intercom. The following procedure was then used in Phase III to establish two colored lights as the discriminative stimuli for truth telling and lying. The experimenter asked questions one at a time from the subject's self-information form. After each question, one of two colored lights was illuminated in the subject's room (according to the "lie detection" cover story, the lights were connected to "voice recording equipment"). If the red light came on, the subject had to answer with a truthful statement; if the

white light came on, he had to give an untrue answer to the question (the colors were reversed for half the subjects). Half of the fifty questions were paired with the "truth light" and the remainder with the "lie light."

The final phase of the study tested the effect of this discrimination training on the subject's recall performance. The subject was required to make forty statements about the words he did or did not cross out in the Phase I word task. Half of these statements were true ones and half were false. While the subject was making each of these forced responses, or "confessions," one of the two colored lights was illuminated. There were thus four experimental conditions of ten words each: false confession - truth light, false confession - lie light, true confession - truth light, true confession - lie light. This within-subjects design allowed each subject to serve as his own control. After making each "confession," the subject indicated whether he thought he actually had or had not crossed out the word. He also indicated his confidence in the accuracy of his recall on a five-point scale (ranging from 5 = "absolutely sure" to 1 = "not sure at all"). At the end of the experiment, the subject was questioned about his reactions to the lights and the "confessions," and was then debriefed.

### Results<sup>3</sup>

A preliminary analysis of the data showed that Bem's predicted interaction was not occurring. Rather, subjects were more accurate in their recall under both of the lie light conditions than under the truth light conditions. This finding suggested that the lie light may have

acquired the properties of a "vigilance" cue, signalling the subject to take his time and think carefully about his response. To test this alternative notion, the subject's response time (i.e. the interval between the "confession" and the recall response) was measured, the prediction being that it would be longer in both lie light conditions. In addition, the subject's physiological responsivity (heart rate and GSR) was recorded on an Offner Type R Dynograph. It was predicted that the greater vigilance induced by the lie light would be reflected in a greater probability of GSR responses. Following from the Lacey, Kagan, Lacey, & Moss (1963) findings of heart deceleration with sustained attentiveness, it was also predicted that heart rate would be lower for the lie light conditions. The use of recording electrodes did not appear to disrupt the experiment in any way; in fact, it actually enhanced its validity by making the cover story of "lie detection" even more plausible.

Overall, the data for 55 subjects strongly support the vigilance hypothesis (main effect of lights) and definitely do not confirm Bem's theory of self-observation and persuasion (interaction of lights and confessions). As shown in Table 1, subjects made fewer errors in the

-----  
Insert Table 1 about here  
-----

lie light conditions than in the truth light conditions. This comparison is significant at the .001 level in the false confession condition and thus replicates Bem's original finding. However, the same comparison in the true confession condition ( $p < .02$ ) is opposite in direction to Bem's result. The subjects' confidence ratings generally reflect their recall



performance; confidence was higher for lie light than for truth light conditions ( $t = 2.10$ , 54  $df$ ,  $p < .05$ ).

A better index of the subjects' recall performance combines recall accuracy and confidence into a single score. An error which is made with a confidence rating of 5 is a "bigger" error than one made with a confidence rating of 1, for example, and should be weighted differently. Accordingly, each of the subjects' responses was given a score from 1 to 10, where 1 = error with a confidence rating of 5, 2 = error with a confidence rating of 4 . . . 6 = correct response with a confidence rating of 1 . . . 10 = correct response with a confidence rating of 5. Table 2 presents the means of these combined scores for each of the four

-----  
Insert Table 2 about here  
-----

conditions. The "non-conflict" groups (truth light - true confession, lie light - false confession) can be considered as a baseline against which to compare the opposing "conflict" groups. The "conflict" condition with the lie light has a slightly higher score than this baseline, while the truth light "conflict" group has a lower one.

A more precise metric than mean combined scores is provided by the memory operating curve (Atkinson, Bower, & Crothers, 1965). This curve takes into account not only the size (i.e. weighting) of the response, but the type of correct or incorrect response that is made. In other words, there are two types of errors (saying "yes" when the true answer is "no," or "no" when the true answer is "yes") and, similarly, two types of correct responses. To determine the shape of the memory operating curve, a 10-point scale of the subjects' possible responses, ranging

from "no" with confidence of 5 to "yes" with confidence of 5, is used. The frequency of subjects' responses at each of these scale positions is recorded, thus forming a response frequency distribution. One such distribution is calculated for the 20 test items which had a correct answer of "yes," and another for the 20 items which were correctly answered by "no." These two distributions are converted to cumulative probability distributions which then form the axes of the memory operating curve. Two such curves were plotted for the present study -- one for the lie light conditions and one for the truth light conditions.

The shape of the curves describes how well the subjects remembered what words they did or did not cross out in Phase I. If their recall had been no better than chance, the resulting curve would be the diagonal shown in Figure 1. That is, subjects would be just as likely

-----  
Insert Figure 1 about here  
-----

to give a particular response (for example, "no" with confidence of 4) to a "yes" item as to a "no" item. However, if their recall was absolutely perfect (i.e. they always gave the correct answer with confidence of 5), the curve would be a right angle along the upper lefthand corner of the graph. Thus, the better the subjects's memory, the further away the curve is from the chance performance diagonal and the closer it is to the perfect performance right angle. Figure 1 presents the memory operating curves for the present study and shows graphically that subjects had better memory in the lie light conditions than in the truth light conditions.

The measures of response time and physiological responsivity offer

some weak support for the vigilance hypothesis, since the differences were in the predicted direction but did not reach significance. The lie light conditions showed a slightly longer response time, lower heart rate, and larger numbers of GSR responses. Some further support for the vigilance hypothesis comes from subjects who voluntarily reported that they became more anxious and attentive when the lie light was present, but relaxed and didn't try so hard when the "friendly" truth light came on.

#### Discussion

The findings of this study suggest that the presence of the lie light causes the subjects to think harder and to be more cautious, with the result that they make fewer errors and are more confident in their recall accuracy. The lie light appears to be a cue for a generalized state of vigilance, and thus has a different cognitive significance for the subject than the truth light. This cue property could result from the training procedure in Phase II. While it is fairly easy for the subject to respond with a true answer to a question, he has to think harder and make a "double response" in order to lie. That is, he has to first think of the true answer and then think of one different from it. Such responses, which involve more complex reasoning, have been shown by Clark (1969) to be associated with a longer reaction time.

An analysis of each subject's pattern of responding reveals that the majority (51%) of the subjects showed the vigilance main effect, while 20% of them showed Bem's predicted interaction. Most of the remaining subjects showed response patterns that were the reverse of

these two hypotheses. Eighteen percent had interaction patterns that were the opposite of Bem's predicted interaction, while only 7% showed a reversed main effect that was contrary to the vigilance hypothesis. The remaining 4% of the subjects showed no effect of the discriminative light training on their recall. These results suggest that subjects were responding differently to the experimental situation, perhaps as a function of their pre-experimental training histories. It may be that such variables as incentive or motivation to do well are responsible for these individual differences, and should be systematically varied in future research.

Although Bem speaks of inducing belief in false confessions, this phenomenon is not actually demonstrated by the results of either his original study or the present replication. If a person had actually come to believe his false statements, his confidence in the accuracy of his recall should be as high as the confidence in his true statements. However, the experimental findings show that confidence ratings drop as the subject makes more errors. In other words, he is, with good reason, becoming more unsure of his recall accuracy. Such confusion and lack of confidence may set the stage for belief in false confessions (by making the person more vulnerable to subsequent interrogation tactics), but they are definitely not evidence of the belief itself. Future research could better establish the possible link between confusion and susceptibility to attitude change. The findings of the present study suggest that subjects would be most likely to change their attitudes after being exposed to the truth light - false confession condition, but would be most resistant to attitude change techniques in the lie light - true

confession condition. In contrast, Bem's model and experimental findings would predict that both of these "conflict" conditions would be equally susceptible to attitude change techniques, and more so than either of the "non-conflict" conditions.

It is somewhat difficult to extrapolate from the two present studies to the "real world" because the experimental situation was a rather artificial and uninvolved one. Moreover, since the obtained differences were fairly small, it is questionable whether they are of practical significance. Further experiments on false confessions should use different paradigms which build in a greater degree of mundane realism, and in which the act of "confessing" has greater consequences for the subject. While the "false confession" of an experimental subject does nothing more than lower his test score, the false confession of the suspect in the squad room can result in the loss of his freedom or life.

## References

- Atkinson, R. C., Bower, G. H., & Crothers, E. J. An introduction to mathematical learning theory. New York: John Wiley, 1965.
- Bem, D. J. Inducing belief in false confessions. Journal of Personality and Social Psychology, 1966, 3, 707-710.
- Clark, H. H. Linguistic processes in deductive reasoning. Psychological Review, 1969, 76, 387-404.
- Lacey, J. I., Kagan, J., Lacey, B. C., & Moss, H. A. The visceral level. In P. H. Knapp (Ed.), Expression of the emotions in man. New York: International Universities Press, 1963.
- Lane, H. L., & Bem, D. J. A laboratory manual for the control and analysis of behavior. Belmont, Calif.: Wadsworth, 1965.
- Zimbardo, P. G. The psychology of police confessions. Psychology Today, June-July, 1967.

Footnotes

<sup>1</sup>This research was supported by ONR Contract N000 14-67-A-0112-0041 to Philip G. Zimbardo, who also provided the laboratory facilities and valuable advice. Thanks are extended to Robert Matschullat, John Weeks, Edward Oliver, and Jan Wattles for their help.

<sup>2</sup>A more detailed procedural description of this experiment, as well as the stimulus materials, can be found in the laboratory manual by Lane & Bem (1965).

<sup>3</sup>All of the following data were analyzed by two-tailed t-tests based on difference scores for each subject (in contrast to the one-tailed t-tests employed by Bem).

TABLE 1

Mean Number of Recall Errors for Ten Trials

(Comparable Data from Bem in Brackets)

Stimulus light	Confession	
	True	False
Truth	3.04 [2.36]	3.76 [3.82]
Lie	2.49 [3.82]	2.91 [1.82]

Lie vs. truth light ( $t = 3.68$ , 54 df,  $p < .001$ )True vs. false confession ( $t = 1.92$ , 54 df,  $p < .10$ )



TABLE 2

Mean Scores Combining Accuracy and Confidence

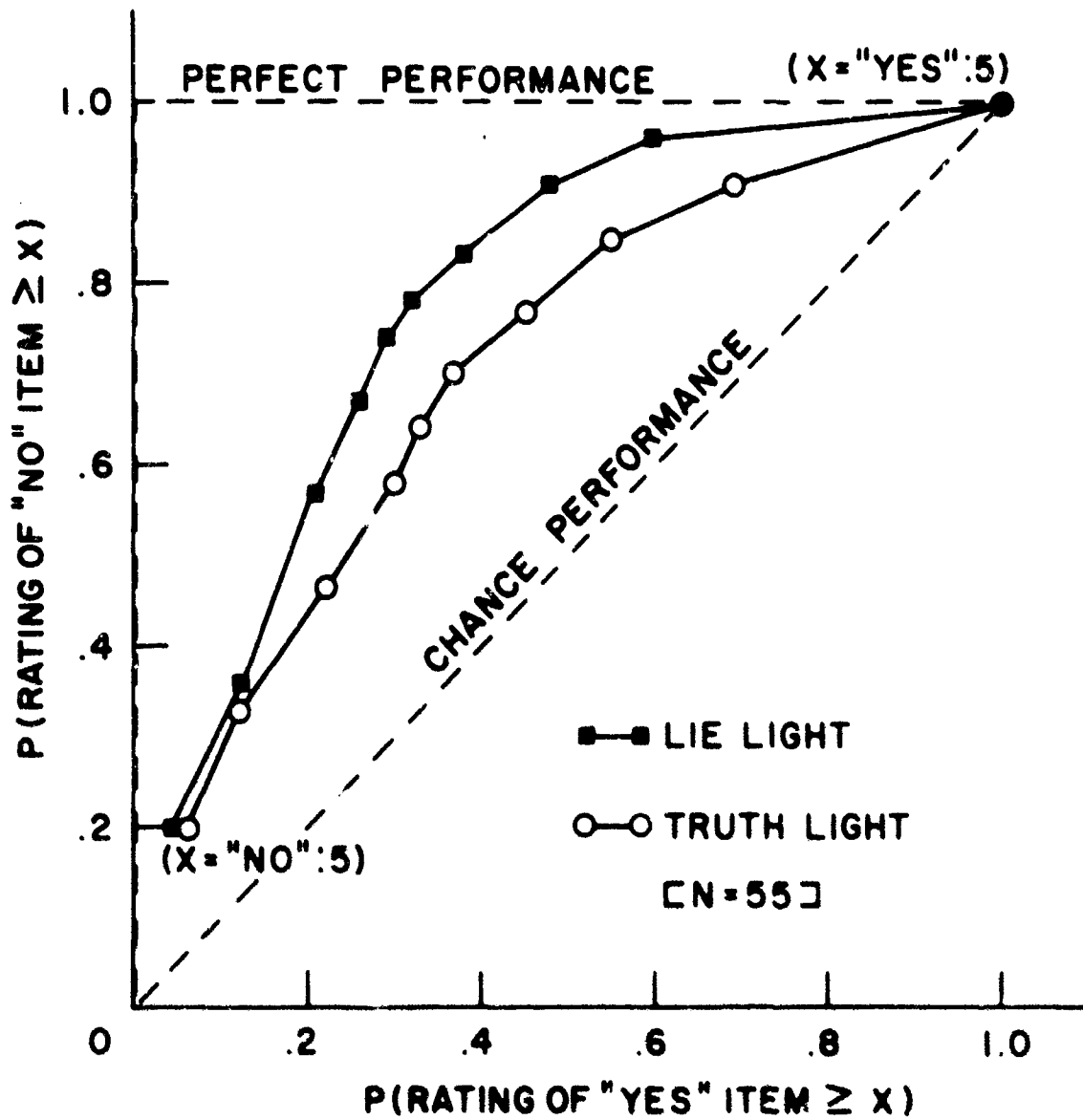
(Range: 1 - 10)

Stimulus light	Confession	
	True	False
Truth	6.90	6.45
Lie	7.27	7.07

Lie vs. truth light ( $t = 4.62$ , 54 df,  $p < .001$ )True vs. false confession ( $t = 1.82$ , 54 df,  $p < .10$ )

Figure Caption

Figure 1: Memory operating curve



## DOCUMENT CONTROL DATA - R &amp; D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)		2a. REPORT SECURITY CLASSIFICATION	
Stanford University		Unclassified	
		2b. GROUP	
		---	
3. REPORT TITLE			
THE 'TRUTH' ABOUT FALSE CONFESSIONS			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)			
Technical report			
5. AUTHOR(S) (First name, middle initial, last name)			
Christina Maslach			
6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS	
April 1970	18	6	
8a. CONTRACT OR GRANT NO.		9a. ORIGINATOR'S REPORT NUMBER(S)	
ONR Contract N000 14-67-A-0112-0041		---	
b. PROJECT NO.			
---			
c. ---		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d. ---		---	
10. DISTRIBUTION STATEMENT			
Distribution of this document is unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY	
Philip G. Zimbardo, Principal Investigator		Office of Naval Research Washington, D. C. 20360	
13. ABSTRACT			
<p>One of the few experimental studies to focus upon variables which might play a role in police interrogation is Bem's "false confession" research. He showed that subjects come to believe that their false statements are true when emitted in the presence of a discriminative truth stimulus. In an attempted replication, the present study found evidence to support an alternative explanation of this finding, based upon decreased vigilance induced by the truth stimulus. In contrast, the lie stimulus was shown to promote more cautious responding as evidenced by better recall, greater confidence, slower reaction time, and a different pattern of physiological responsivity.</p>			