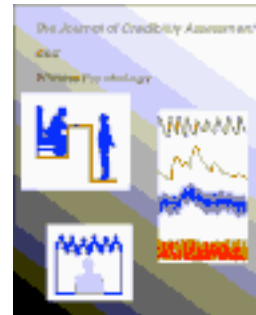


# *The Journal of Credibility Assessment and Witness Psychology*

2003, Vol. 4, No. 1, pp. 1-36

Published by the Department of Psychology of  
Boise State University



## **WHEN DID YOU CONCLUDE SHE WAS LYING? THE IMPACT OF THE MOMENT THE DECISION ABOUT THE SENDER'S VERACITY IS MADE AND THE SENDER'S FACIAL APPEARANCE ON POLICE OFFICERS' CREDIBILITY JUDGMENTS.**

**Jaume Masip, Eugenio Garrido, and Carmen Herrero**

**Department of Social Psychology and Anthropology,**

**University of Salamanca, Spain**

Correspondence concerning this article should be sent to Jaume Masip (jmasip@usal.es) or Eugenio Garrido (garrido@usal.es), Department of Social Psychology and Anthropology, University of Salamanca, Facultad de Psicología, Avda. de la Merced, 109-131, 37005 Salamanca (Spain).

The research reported here was supported by the Junta de Castilla y León, Programa de Apoyo a Proyectos de Investigación, Ref. 30/98.

---

Copyright 2003 by the Department of Psychology of Boise State University and the Authors. Permission for non-profit electronic dissemination of this article is granted. Reproduction in hardcopy/print format for educational purposes or by non-profit organizations such as libraries and schools is permitted. For all other uses of this article, prior advance written permission is required. Send inquiries by hardcopy to: Charles R. Honts, Ph. D., Editor, *The Journal of Credibility Assessment and Witness Psychology*, Department of Psychology, Boise State University, 1910 University Drive, Boise, Idaho 83725, USA.

---

**ABSTRACT: Two experiments were conducted to explore how the moment observers make their decision about the senders' veracity affects their judgment and detection accuracy. In Experiment 1 police officers and undergraduates judged the credibility of video-recorded statements. Contrary to our expectation, officers did not judge the statements earlier than the students. An initial lie bias became evident. In Experiment 2 a still face, which could be of the same witness as in Experiment 1, or of two other witnesses, was shown to officers as they listened to truthful or deceptive accounts taken from Experiment 1. There was no effect of the sender's facial appearance on the lie bias found in the first experiment, which emerged here as well. Accuracy for detecting deceptive accounts decreased across time in both studies, while accuracy for truthful accounts increased only in Experiment 2. How visual and verbal information contributed to these effects is discussed.**

---

## **WHEN DID YOU CONCLUDE SHE WAS LYING? THE IMPACT OF THE MOMENT THE DECISION ABOUT THE SENDER'S VERACITY IS MADE AND THE SENDER'S FACIAL APPEARANCE ON POLICE OFFICERS' CREDIBILITY JUDGMENTS**

### **Introduction**

DePaulo and Rosenthal (1979) identified three main areas of inquiry in the field of nonverbal detection of deception: (a) people's ability to lie successfully and to accurately detect deception, (b) channel or modality effects on accuracy, i.e., what kind of information (visual, vocal, verbal, transmitted by the face, transmitted by the body, etc.) is most useful for untrained observers to detect deception, and (c) the study of the behavioral indicators of deception (real deception cues, perceived deception cues, and behaviors believed by people to be useful to detect deception). Within the first area pointed out by DePaulo and Rosenthal (1979), attention has been paid to sender and/or receiver variables that may affect their ability to deceive or detect deception (variables such as gender, age, experience, personality traits, etc.), as well as to certain situational variables such as motivation to lie successfully, familiarity between sender and receiver, time to create a deceptive story, perceived consequences of being detected, etc. (see reviews by DePaulo, DePaulo, Tang, & Swaim, 1989; DePaulo, Stone, & Lassiter, 1985; Ekman, 1992; Ekman & O'Sullivan, 1989; Ford, 1996; Kalbfleisch, 1992; Köhnken, 1989; Kraut, 1980; Masip & Garrido, 2000, 2001a; Miller & Burgoon, 1982; Miller & Stiff, 1992, 1993; Vrij, 1998, 2000; Zuckerman, DePaulo, & Rosenthal, 1981; Zuckerman & Driver, 1985). In general, meta-analyses on the results obtained from this approach show that detection accuracy (i.e., accuracy at detecting both truths and lies, Miller & Stiff, 1993) by untrained detectors usually falls between 45 % and 60 % correct classifications, where 50 % is the chance level. In addition, it has been found that police officers are no more accurate than lay people in their credibility judgments (e.g., DePaulo & Pfeiffer, 1986; Ekman & O'Sullivan, 1991; Garrido, Masip, Herrero, & Tabernerero, 1997; Garrido, Masip, & Herrero, 2003; Henderson & Hess, 1982; Köhnken, 1987; Kraut & Poe, 1980; Sanderson, 1978, cited by Bull, 1989; Vrij, 1992; Vrij & Graham, 1997; see reviews by Bull, 1989, Garrido & Masip, 1999, and Vrij, 2000). Instead, there is some evidence that police officers may even be less precise than non-officers, due to a lie bias they may display when making their judgments (Garrido et al., 1997; Sanderson, 1978, cited by Bull, 1989).

In view of that poor accuracy level among observers trying to discern whether someone is lying or telling the truth, Miller and Stiff (1993) suggested that that issue should be considered from an alternative perspective: instead of investigating detection accuracy, researchers should identify explanations for observers' errors in their credibility judgments. In line with that

suggestion, in the two experiments reported here we try to identify some factors that may have an effect upon observers' accuracy at judging credibility; specifically we are trying to discern what processes underlie the poor performance attained by police officers in Garrido et al.'s (1997, 2003; see also Masip, 2002) study. In that experiment, officers' detection accuracy did not differ significantly from chance level, while students' accuracy was significantly above chance. The poor accuracy among officers was due to their tendency to judge all statements as false. Officers' accuracy at judging deceptive statements was as high as that of undergraduates, while their accuracy at judging honest statements was poorer. In fact, the tendency of police officers to judge statements as deceptive was the same regardless of the real quality (truthful or deceptive) of those statements, while students were somewhat more sensitive to the real truth value of the stories.

In an attempt to have a closer look at that lie bias among officers, Garrido and Masip (2001) explored whether it was due to a reduced capacity among them to perceive a general expressive pattern as defined by Becerra, Sánchez, and Carrera (1989). These authors suggested that the accurate detection of deceit is based on observers' perception of a general expressive pattern in the sender's behavior, a pattern that changes as the statement quality (value of truth) varies. If so, it could be the case that officers did not perceive that pattern. There may be several reasons for that. For instance, police officers may have a stronger Generalized Communicative Suspicion (GCS) than non-officers. Levine and McCornack (1991) differentiated between GCS and situationally-aroused suspicion or "state" suspicion. The former would be a "predisposition toward believing that the messages produced by others are deceptive" (Levine & McCornack, 1991, p. 328), and is described as a relatively enduring and cross situational cognitive construct. On the other hand, situationally-aroused or state suspicion is prompted by certain contextual cues. It was defined by Levine and McCornack (1991) as "a belief that communication within a specific setting and at a particular time may be deceptive" (p. 328). Unlike GCS, state suspicion is transitory and is based upon certain situational variables.

Detecting deceit is an important task for police officers. During their daily work, they are often involved in social interactions where mistrust and lack of confidence are normal, and where they must question the interviewee's assertions. That is to say, situations where a state suspicion is aroused. Yet, this suspicion, given its frequency in police work, could become chronic, arousing among officers a belief that the interviewee is probably not being truthful. This process would end up generating a kind of suspicion that would no longer be a response to contextual cues nor would it be transitory anymore. Rather, that suspicion would be a GCS. Research has shown that high GCS ratings are associated with a tendency to make judgments of deceptiveness (Levine & McCornack, 1991). With regard to our police officers, it could be the case that their generalized suspicion prevented them from scrutinizing the witness's behavioral displays, thus not being able to perceive his or her general expressive pattern. If this were actually the case, then perhaps officers made only a biased "guess" based on their initial suspicion. Alternatively, police officers' generalized suspicion may have given rise to a confirmation bias, making them attentive to only those behaviors supporting their view that the sender was lying. In either case, police officers would be unable to perceive the general pattern described by Becerra et al. (1989). However, Garrido and Masip's (2001) results showed that not only officers, but also non-officers, were unable to perceive any general expressive pattern in the sender's behavior. Thus, that factor cannot account for the differences between the police and lay people.

In this paper we describe some further explorations of the processes underlying officers' lie bias in Garrido et al.'s (1997, 2003) study. Experiment 1 looks at whether police officers and students came to their conclusion about the sender's veracity at different times, and whether this can account for the judgmental differences between these groups which were detected by Garrido and his colleagues. Also, an interesting question is how the moment observers come to

a conclusion about whether the sender is lying or telling the truth affects detection accuracy, that is, is there any point in time where accuracy is higher? Experiment 2 is a follow-up study to answer some questions raised by the results obtained in Experiment 1. Thus, the contribution of the sender's facial appearance and that of her dynamic nonverbal behavior to the profile found in Experiment 1 is explored.

## **EXPERIMENT 1**

The availability of information (both useful and misleading behavioral cues) depends upon the moment observers make their decision about the truthfulness of the senders' account. If receivers decide at the very beginning of a sender's performance, the amount of available verbal and nonverbal information from that sender will be very limited. Conversely, if observers decide after the sender's performance has concluded, they will be able to take into account all the verbal and nonverbal behavior displayed by that sender throughout his or her performance. Thus, if information gathered by observers paying attention to the senders' behavior is used as a basis for making veracity judgments, accuracy will probably be influenced by the moment observers conclude that the sender is lying or telling the truth, at the beginning, middle, or end of his or her performance.

A first interesting question is whether police officers and non-officers (undergraduate psychology students) tend to decide at different moments in time. Such a difference might account for the differences between those groups found by Garrido et al. (1997, 2003). Our prediction concerning the moment variable is based on the contributions of Levine and McCornack when conceptualizing their GCS, as well as on the work of Stiff, Kim and Ramesh (1992). Thus, it could be the case, for instance, that officers have a strong generalized communication suspicion as mentioned above, so that they enter the situation with the a priori belief that the sender is lying, while lay observers are more attentive to the sender's behavioral displays. In that case, officers would tend to decide quickly at the beginning of the statement, because they "would be certain of it" and would see no need to pay attention to the witness's behavior, while students would tend to decide later in the sender's performance, after paying close attention to that witness's behavior and after having processed the information so gathered. Also, Stiff et al.'s (1992) paper permits drawing an alternative process, which would lead to the same prediction. Those authors justify the development and existence of a cognitive heuristic which would lead relational partners –among which mutual confidence and trust are the norm, as well as necessary to maintain the relationship– to judge the other member's performance as truthful, without even processing the information conveyed by that other member which could potentially be relevant to judge his or her credibility. Probably the same rationale could be used to account for officers' judgmental tendencies, but in the opposite: instead of those cooperative interactions characterized by relational intimacy and trust that relational partners are involved in, police officers often get into interactions where distrust and suspicion are usual. This could create among officers a belief that the interviewee is not being truthful, in the same way that a belief that the other person is being truthful is aroused among relational partners. And, in the same way that a potential lie detector involved in a close relationship bases his or her credibility judgments on the a priori belief that his or her partner is honest, thus making heuristic judgments of truthfulness without even processing the incoming information, police officers could do something similar to conclude that the witness is being deceptive. On the other hand, lay observers judging the credibility of strangers' statements would use a rather different strategy. They would be less biased than officers concerning the sender's honesty, they would be less confident than officers in their skills to assess other people's credibility (Garrido et al., 2003), and, therefore, they would be more willing than officers to attend to and to take into account the behaviors displayed by the witness during his or her statement.

Both processes, the one based on a GCS among officers and the one derived from Stiff et al.'s (1992) findings concerning lie detection amongst relational partners, suggest that police officers will display a lie bias such as that found by Garrido et al. (1997, 2003) while lay observers will be able to take into account information drawn from the sender's behavior to make their judgment, which will make them more accurate at assessing credibility. Thus, our first hypothesis predicts that police officers will make their decision about the sender's veracity earlier than non-officers, because, unlike these, officers will tend not to pay attention to the incoming information which would help them make an accurate judgment.

An important moderating factor on the effects of the moment observers decide on judgmental accuracy may be the value of truth of the statement to be judged. For instance, it could be the case that, for truthful accounts, deciding at the end is beneficial, given the greater amount of accurate information available at that later point. However, the prediction is different when it comes to assessing deceptive statements. Liars may monitor their behavior in order to give a plausible false account (information management) as well as an honest impression (image management and behavior management) (see Buller & Burgoon, 1996, 1998; DePaulo, 1991, 1992; DePaulo & Kirkendol, 1989; Greene, O'Hair, Cody, & Yen, 1985; Masip & Garrido, 1999, 2000, 2001a; Vrij, 1998, 2000; Zuckerman et al., 1981). Therefore, the later observers come to a conclusion about the deceiver's truthfulness, the more misleading cues that deceiver will have had a chance to display. Not all of his or her cues will be misleading, since some behaviors are hardly controllable (e.g., Ekman, 1992; Ekman & Friesen, 1969, 1974), but in any case, later in the sender's performance, the amount of misleading information will be greater in false accounts than in honest ones, while the amount of truthful information will be relatively smaller. However, earlier in the account these differences will be less pronounced. Therefore, an interaction between the moment observers decide and the value of truth of the statements could be expected. Thus, our second hypothesis predicts that, as observers decide later in time, there will be a relative increase in accuracy at detecting truthful accounts and a relative decrease in accuracy at detecting deceptive accounts.

It is important to stress that this study was designed to test hypothesis one. Since only one sender was used, either supportive or non-confirmatory evidence for hypothesis two must be taken only as preliminary and suggestive evidence until replications with a large number of senders be conducted.

## Method

**Participants:** The sender was a female undergraduate student of psychology at a Spanish University. Observers were 121 police officers studying to become police inspectors at the Police Academy of Ávila (Spain), and 147 undergraduate students of psychology at a Spanish University<sup>1</sup>.

---

<sup>1</sup> In the Spanish National Police Force there are officers (two degrees: *policías* and *oficiales de policía*), subinspectors, inspectors, chief inspectors (*inspectores jefe*), superintendents (*comisarios*), and chief superintendents (*comisarios jefe*). Normally, a superintendent is in charge of a police station, and an inspector is in charge of a group of police officers. To become an officer it is necessary to study the Basic Level (*Escala Básica*) at the Police Academy. Applicants are required to have completed their primary education, as well as to pass a competitive examination. To become an inspector it is necessary to study the Executive Level (*Escala Ejecutiva*) of the Police Academy. To enter the Executive Level applicants (lay people, that is, non-officers) must be 28 or younger, must have studied at the University, and must pass a competitive examination. There is, however, another way to access the Executive Level: Police officers

**Procedure:** In order to increase the ecological validity of this study, we addressed some of the concerns expressed by various authors in this area (e.g., Köhnken, 1987, 1989; Miller & Stiff, 1993) by (a) motivating our senders to be convincing, (b) making the content of the statements relevant to police interrogation settings: the topic was the reporting of criminal actions (factual descriptions), (c) by giving senders a few minutes to prepare before giving their statements, (d) by having observers make a dichotomous decision (“true” or “false”) instead of rating the degree of truthfulness or deceptiveness, and (e) by showing observers only two statements of some length (no less than two minutes). Normally, in laboratory research on nonverbal detection of deception a large number of small behavioral samples are shown to observers. However, in the real world officers rarely have to judge the credibility of dozens of statements that are only a few seconds long. We addressed this issue by showing observers only two statements of some length, although this prevented us from using a large sample of senders.

In order to motivate our sender, we offered all psychology students at our University who were taking a social psychology module a substantial academic reward if they participated as witnesses in a lie detection study and were the most convincing of all senders. Four undergraduate females volunteered. Each of them was shown two film sequences depicting criminal actions (S1 and S2). After watching each of these sequences, senders were instructed to work out a deceptive version (D) and a truthful one (T) of the sequence. They were left ten minutes to create each version, and were video recorded as they made their statements [a free narrative account no less than two minutes long. Thus, each sender produced four statements: a deceptive account of the first sequence (S1D), a truthful account of that same sequence (S1T), a deceptive account of the second sequence (S2D) and a truthful account of that second sequence (S2T). A pilot study was conducted with a few undergraduates in order to choose the most convincing liar for the main study<sup>2</sup>. All four candidates received the advertised reward for their participation.

The four performances of the sender who was chosen were edited and shown to 121 police officers and 146 psychology students. Each participant watched two statements: one based on S1 and the other based on S2. These statements could be both truthful (31 police officers and 38 undergraduates were allocated to this condition), both deceptive (29 officers and 40 undergraduates), truthful the first to be shown and deceptive the second to be shown (31 officers and 36 students) or deceptive the first and truthful the second (30 officers and 32 students). All police officers allocated to the same experimental condition were in the same class in the police academy at the moment the experimental session was carried out; allocation of officers to their classes is based on an alphabetical criterion. Allocation of undergraduate

---

with a given number of years of on-the-job experience may apply for promotion to police inspectors; if their application is successful, they are then sent to the Police Academy and enter the Executive Level. These students are normally older than the former, and have long experience as officers. In either case, completing the Executive Level takes two years at the Academy plus a practical-training year at a police station. Our “police officers” were novice (i.e., young and inexperienced) students of the Executive Level in their second year at the Academy. Larger differences would be expected using very experienced officers (which were unavailable at the time data were collected), but Garrido et al. (1997, 2002) found that the tendency to judge statements as deceptive was stronger among the same novice officers used in this study than among the undergraduate students. The military-like kind of life officers have at the Police Academy may be responsible for the effectiveness of such a brief socialization process, thus accounting for the differences between them and the students which were found by Garrido and his colleagues.

<sup>2</sup> The main study, reported by Garrido, Masip, Herrero and Taberner (1997) (conference presentation) and Garrido, Masip, and Herrero (2003) (paper under review), compared police officers’ and lay people’s ability to detect truthful and deceptive statements. In order to ensure that credibility judgments were not obvious, so that differences between more skilled and less skilled groups could emerge, a liar was chosen who, according to the ratings by the participants in the pilot study, was relatively good at deceiving.

students to the experimental conditions was made randomly. Since the number of officers per classroom was not the same across all classrooms, and some students failed to attend their sessions and/or came to a session different from the one they had been assigned to, there were some small variations in size across the experimental groups.

After watching each of the two performances of the sender, observers were given a few minutes to complete a questionnaire. One of the items asked them whether they thought the sender had lied or told the truth. Another item asked observers whether they had reached their conclusion early, as they started to see the sender's performance (Moment 1), at the middle part of that performance (Moment 2), or at the final moment (Moment 3).

## Results

### *Hypothesis Testing*

Data were analyzed separately for S1 and S2. Two stepwise backward hierarchical loglinear analyses were performed using SPSS 9.0. The variables introduced were value of truth of the statement (truthful / deceptive), observers' occupation (police officer / undergraduate), the hit / miss variable, and the moment observers made their decision (Moment 1 or 2 / Moment 3)<sup>3</sup>. Both a significant association among occupation and moment –as predicted in hypothesis 1– and a value of truth X moment X hit / miss –in the way predicted in Hypothesis 2– would be expected to emerge for both statements. In addition, concerning the first three variables, we expected to find results similar to those reported by Garrido et al. (2003; see also Masip, 2002), which were based on these data.

Concerning S1, k-way effect tests showed the fourth-order interaction was of no relevance, likelihood-ratio chi-square:  $\chi^2_{(1)} = 0.04, p = .846$ , but there were substantial third-, second-, and first-order effects, respectively:  $\chi^2_{(4)} = 18.05, p = .001$ ;  $\chi^2_{(6)} = 69.54, p = .000$ ;  $\chi^2_{(4)} = 9.68, p = .046$ . Something similar was found for S2, respectively:  $\chi^2_{(1)} = 1.59, p = .220$ ;  $\chi^2_{(4)} = 22.80, p = .000$ ;  $\chi^2_{(6)} = 22.76, p = .001$ ; and  $\chi^2_{(4)} = 28.83, p = .000$ . The best model for S1 comprised three interactions: Occupation X Value of Truth X Hit/Miss (police officers made more errors when judging truthful statements than when judging the deceptive; this effect was presented and discussed by Garrido et al., 2003), Value of Truth X Moment X Hit/Miss, and Occupation X Moment (these interactions are discussed briefly). This model had an excellent goodness of fit: its likelihood-ratio chi-square was  $\chi^2_{(3)} = 0.24, p = .971$ , and the greatest standardized residual had an absolute value of 0.29. The best model for S2 was somewhat simpler, comprising only the two third-order interactions also included in the S1 model, that is, Occupation X Value of Truth X Hit/Miss and Value of Truth X Moment X Hit/Miss. This model had a likelihood-ratio chi-square of  $\chi^2_{(4)} = 4.93, p = .294$ , and the greatest standardized residual had an absolute value of 0.85.

---

<sup>3</sup> Few judgmental decisions were made at Moment 1: only 77 (28 of which came from the first statement [S1] and 49 from the second [S2]). At moments 2 and 3 the number of judgments was virtually the same (N = 231 at Moment 2, N = 223 at Moment 3); this was so both in the first statement (122 at Moment 2, 117 at Moment 3) and in the second statement (109 at each moment), and these judgments were much more numerous than those of the first moment. This small frequency of initial judgments turned out to be a limitation: by looking at the expected frequencies in the contingency tables where all the variables to be introduced in the loglinear analysis were crossed, it became evident that, in moment one, these were too small to conduct the analysis. Therefore, in order to calculate the statistics moments 1 and 2 were grouped and taken together. Thus, this variable had two categories: Moments 1 and 2 v. Moment 3.

In order to examine the specific contribution of each effect to the fit of the model, attention was paid to partial association tests and parameter estimates. In Table 1 this information is

**TABLE 1. Partial Association Tests And Parameter Estimates (In Absolute Values) Of The Effects That Had A Relevant Contribution To The Fit Of The Model, Either In S1 Or In S2.**

Effects	Partial association				Parameter estimates*			
	S1		S2		S1		S2	
	$F^2_{(1)}$	<i>p</i>	$F^2_{(1)}$	<i>p</i>	$ \beta $	$ z $	$ \beta $	$ z $
<i>Third-order effects</i>								
Occupation X Value of Truth X	8.13	.004	7.04	.008	.20	2.66	.17	2.44
Hit/Miss								
Value of Truth X Moment X	11.40	.001	12.65	.000	.24	3.12	.24	3.58
Hit/Miss								
<i>Second-order effects</i>								
Occupation X Moment	3.24	.072	0.45	.832	.15	1.92	.02	0.25
Value of Truth X Moment	1.39	.239	12.18	.001	.03	0.42	.19	2.82
Value of Truth X Hit/Miss	49.63	.000	3.87	.049	.47	6.20	.10	1.47
Moment X Hit/Miss	12.25	.001	0.99	.321	.21	2.80	.09	1.32
<i>First-order effects</i>								
Moment	4.09	.043	9.88	.002	.04	.55	.13	1.96
Hit/Miss	3.16	.076	16.11	.000	.16	2.18	.19	2.79

\* In absolute values. To examine the direction of effects see Appendix 1.

summarized for all those effects which either approached significance or were significant, whether in S1 or in S2. The direction of effects can be observed in Appendix 1, where the presentation model suggested by Tabachnick and Fidell (1996) was used.

Concerning the occupation, value of truth, and hit / miss variables, Garrido et al.'s (2003) results were here replicated with some minor nuances. In particular, the most relevant effect



(the three-way interaction) was found again and, furthermore, it did not interact with the moment observers decided their judgment. Thus, the introduction of that variable in the analyses did not substantially alter the former results. Since they were already discussed elsewhere (Garrido et al., 2003) and are not the main focus of the present report, they will not be discussed here again.

Instead, our focus in the present paper centers on those effects involving the moment variable. A certain tendency was found in S1 to make the decisions at moments 1 and 2 (56.18 % of judgments) instead of making them at Moment 3 (43.82 %). Something similar, although the trend was clearer, happened in S2 (percentages were, respectively, 59.19 % and 40.82 %). This effect could be due to having added the number of decisions made at Moment 1 to those made at Moment 2, since in both statements the latter had a frequency that was quite similar to that of Moment 3 decisions.

The *first hypothesis* predicted that police officers would hurry and make their judgment earlier than the undergraduates. Therefore, an association between being an officer and deciding at moments 1 and 2, and between being a student and deciding at Moment 3 would be expected. However, the Occupation X Moment association was not significant in S2. In S1 it did not reach statistical significance either, as indicated by the two measures used to explore the individual effects (although it was close to significance:  $\chi^2_{(1)} = 3.24$ ,  $p = .072$ ;  $z = -1.92$ ), but the program retained the effect while searching for the best model during the stepwise procedure (the associated change likelihood-ratio chi-square was  $\chi^2_{(1)} = 5.48$ ,  $p = .019$ ). In any case, the effect was the opposite to what was expected: police officers did not tend to make their decisions earlier than the undergraduate students, but later (see Appendix 1): While in S1 48.76 % of police officers decided at moments 1 and 2 in comparison with the remaining 51.24 %, who decided at Moment 3, 62.33 % of undergraduates decided during the early moments and only 37.67 % of them did so at Moment 3. In S2 the associations failed to reach significance, but they pointed in the same direction. In summary, our first prediction did not receive empirical support. If there was any occupational group which acted hastily in making their credibility judgments it was not the officers, but rather the undergraduate students.

The *second hypothesis* predicted an interaction between the value of truth of the statement, the decision-making moment, and the correctness of the credibility judgment (hit or miss), in the sense that deceptive statements would be more accurately detected earlier in the statement than later on, while truthful statements would be judged with higher accuracy at the final moment rather than at the beginning of the statement. To begin with, it should be mentioned that some second-order effects were substantial. In S2, the interaction Moment X Value of Truth indicates that, when judging the false account (S2D), the decision was made basically at the beginning (69.85 % of cases); this did not happen when judging S2T (48.09 %). In S1 it was the Moment X Hit/Miss interaction that was relevant: the decisions made at the beginning of the statement were accurate more often than those made at the final moment. But both of these effects were qualified by the higher-order value of Truth X Moment X Hit/Miss interaction, which lent support to our second hypothesis. When judging the deceptive statements an association was found both in S1 and S2 between making the decision early (moments 1 and 2) and guessing right, as well as between deciding at the final moment and judging wrongly. An opposite tendency became apparent when judging the truthful statements (see Appendix 1). As stated before, this effect was one of the components of the final model in both analyses: the one concerning S1 and the one concerning S2.

Although this interaction was significant it would be interesting to analyze whether, in an absolute sense, there was a significant decrease in accuracy when judging deceptive statements at Moment 3 in comparison with the early moments, as well as whether the increase in accuracy for the truthful statements was significant too. In order to examine these effects, individual chi-square analyses were performed to examine the associations among the

hit/miss and the moment variables separately for truthful and deceptive accounts. The results of these analyses are summarized in Table 2. It is apparent that the predicted decrease for the deceptive statements was significant. However, the increase in accuracy for truthful statements across time was not found, although a marginally significant trend in the predicted direction was apparent in S2T.

**TABLE 2. Moment X Value Of Truth X Hit/Miss Contingency Tables, And Chi-Square Analyses For Truthful And Deceptive Statements.**

Statements	Moment		$\chi^2$ (1)	<i>p</i>
	Moments 1 and 2	Moment 3		
Sequence 1 (S1)				
Deceptive (S1D)				
Hit	62 ( 1.8)	24 (-2.1)	22.53	.000
Miss	13 (-2.5)	32 ( 2.9)		
Truthful (S1T)				
Hit	19 ( 0.2)	14 (-0.2)	0.10	.747
Miss	56 (-0.1)	47 ( 0.1)		
Sequence 2 (S2)				
Deceptive (S2D)				
Hit	74 ( 1.1)	19 (-1.6)	12.13	.000
Miss	21 (-1.6)	21 ( 2.4)		
Truthful (S2T)				
Hit	30 (-0.8)	42 ( 0.8)	2.98	.084
Miss	33 ( 0.9)	25 (-0.9)		

In conclusion, our second hypothesis was supported by the data. There was a relative decrease in accuracy over time when judging deceptive statements, and a relative increase when judging the truthful ones. However, in an absolute sense, although the decrease when judging deceptive statements was significant, the increase of judgmental accuracy for truthful statements did not reach statistical significance.

#### *Early Lie Bias*

It is worth noticing that, aside from the idiosyncrasies of each statement, the data reported here show that, in general, a strong initial bias toward making judgments of deceptiveness was apparent. There were large differences between observed frequencies of hits and misses at moments 1 and 2, both when the statements were deceptive (many more hits than misses) and when they were truthful (more misses than hits) (see Table 2). At Moment 3 these differences were severely reduced or reversed.

Although the increased number of initial judgments of deceptiveness seemed to be more evident when the statements were deceptive, the differences between the number of truth and lie judgments were statistically significant not only in that case,  $\chi^2(1) = 58.82, p = .000$ , but also when statements were truthful,  $\chi^2(1) = 11.59, p = .001$ . On the contrary, at Moment 3 there were no significant differences between judgments of truthfulness and judgments of deceptiveness, either when statements were deceptive,  $\chi^2(1) = 0.67, p = .414$ , or when they were truthful,  $\chi^2(1) = 1.53, p = .212$ . All these results indicate that lie judgments were more numerous when deciding at the beginning of statements than when the decision was made at the end. Indeed, a Moment (moments 1 and 2 / Moment 3) X Judgment (judgment of truthfulness/ of deceptiveness) *Chi-square* analysis was significant,  $\chi^2(1) = 25.66, p = .000$ .

These tendencies cannot be accounted for by the differences between police officers and undergraduates in terms of the kind of judgment each group tended to make. As reported elsewhere (Garrido et al., 1997, 2003), police officers' tendency to deem statements as deceptive was stronger than non-officers',  $\chi^2(1) = 9.57, p = .002$ . This tendency was significant not only at moments 1 and 2,  $\chi^2(1) = 8.81, p = .003$ , but also at Moment 3,  $\chi^2(1) = 4.05, p = .044$ , and the difference between truth and lie judgments at moments 1 and 2 was significant not only among officers,  $\chi^2(1) = 51.72, p = .000$ , but also among the students,  $\chi^2(1) = 18.90, p = .000$ , while at Moment 3 the difference in frequency of truth and lie judgments was only marginally significant among police officers,  $\chi^2(1) = 2.95, p = .086$ , and was completely non-significant among the undergraduates,  $\chi^2(1) = 1.26, p = .261$ . Therefore, the differences between the officers and the students cannot account for that trend towards judging statements as deceptive fundamentally at the beginning of the sender's performance. In addition, in this regard it is important to keep in mind that the officers, whose tendency to judge the statements as deceptive was stronger than undergraduates', displayed a certain propensity to make their judgments later than the students. Also, a backward stepwise hierarchical loglinear analysis was performed to examine the relation between Occupation, Moment, and Judgment (truth / lie judgment). If the association between moment and judgment were moderated by the observers' occupation, then k-way effect tests would yield significant results for  $k = 3$ , and the analysis would not continue beyond the saturated model (Occupation X Moment X Judgment). However, the null hypothesis that third-order effects were zero was supported, likelihood-ratio chi-square:  $\chi^2(1) = 0.48, p = .488$ , and the best model comprised the Occupation X Judgment interaction, partial  $\chi^2(1) = 12.65, p = .000, |\lambda| = .17, |z| = 3.47$  (which reflects the police's tendency to judge the statements as deceptive), the Occupation X Moment interaction, partial  $\chi^2(1) = 5.49, p = .019, |\lambda| = .12, |z| = 2.41$  (which reflects the aforementioned tendency among police officers to make their judgment later than the undergraduates, which in this analysis was clearly significant), and the judgment X moment interaction, partial  $\chi^2(1) = 28.60, p = .000, |\lambda| = .25, |z| = 5.27$  (which reflects the tendency we are discussing to make lie judgments at the beginning of the statements but not at the end of them). This model had an adequate goodness of fit, likelihood-ratio chi-square:  $\chi^2(1) = 0.48, p = .488$ ; the greater standardized residual had an absolute value of 0.36.

In summary, regardless of whether statements were truthful or deceptive, early judgments were primarily lie judgments. Later on, at Moment 3, truth and lie judgments were more balanced. Police officers' tendency to make lie judgments cannot account for this effect.

## Discussion

*Hypothesis 1: The moment officers and non-officers made their decision*

Hypothesis one, which predicted that police officers would make their decision early in the statement and that students would decide later, was not supported. Differences were contrary to what was expected (i.e., there was a tendency among officers to decide at the final moment

and among the students to do so at the beginning), and the effect was retained in some of the loglinear analyses. Maybe officers, due to their awareness that information gathered from witnesses is important, were more attentive and did not decide until they had collected the information, while students were less thorough and hastened their decision. Nevertheless, officers' bias to judge truthful statements as deceptive (Garrido et al., 1997, 2003) suggests that they were either incapable of gathering the relevant information in order to make their veracity judgments, or they did not make correct use of the information they collected.

#### *Hypothesis 2: The influence of moment on accuracy*

Hypothesis 2 was partially supported. Deceptive statements tended to be judged accurately at the beginning and inaccurately at the end. Probably, the greater amount of misleading information at Moment 3 in comparison to moments 1 and 2 made observers less accurate at detecting deception, since that information serves to (a) give an impression of being honest, and (b) make the lie plausible and, hence, credible (Buller & Burgoon, 1996, 1998; Leekam, 1992). This leads to truth judgments. Clues to deceit will increase later in time as well, but research shows that untrained observers are not good at using such clues; instead, they often base their judgments on invalid indicators (DePaulo et al., 1985; DePaulo, Zuckerman, & Rosenthal, 1980b; P. DePaulo et al., 1989; Ekman, 1989; Vrij, 1998, 2000; Vrij & Winkel, 1993).

However, for truthful accounts the frequency of observed hits and misses at moments 1 and 2 in comparison with Moment 3 did not depart significantly from what was expected. In other words: observers' judgments were more or less equally accurate at any moment in time. This is at odds with our second hypothesis, which predicted an increase in accuracy across time. In any case, as stated above, hypothesis two was just an exploratory hypothesis, and the present findings are only preliminary. A study with a large number of senders is currently being conducted to replicate these findings.

#### *Early Lie Bias*

Our results show a strong initial lie bias for our observers. This was so among both students and officers, despite the finding reported by Garrido et al. (1997, 2003) that, overall, officers were more prone to make judgments of deceptiveness than non-officers. Thus, while an initial accuracy level common for both truthful and deceptive statements and close to chance probability could be expected, since at the beginning of the sender's performance the amount of both accurate and misleading information was similar for all statements and similarly scarce, a strong tendency to judge statements as false was found at that point in time. Thus, when observers decided early in time they tended to make judgments of deceptiveness. A possible explanation for this initial bias may lie on our sender's physical appearance. Research indicates that people's facial appearance may influence social perceivers' impressions of sincerity (e.g., Berry & Brownlow, 1989; Berry & McArthur, 1985; Zebrowitz & Montepare, 1992; Zebrowitz, Voinescu, & Collins, 1996). The only available information at the beginning of the statements was the sender's appearance. Thus, it is possible that initial credibility judgments were influenced by by our sole sender's appearance. If our sole sender had a facial appearance that fitted the stereotype of a liars' face, the intriguing initial lie bias could be due to that factor. This possibility was investigated in Experiment 2.

#### *Disproportion Of The Number Of Decisions Made At Different Moments.*

The small number of decisions made at Moment 1 in comparison with those made at moments 2 and 3 may be due to the subjective nature of the distinction between Moment 1 and Moment 2. There were no "markers" of the boundaries between the different moments.

Observers who decided at Moment 3 were probably those who waited until the video presentation was over to judge whether the sender lied or told the truth. However, boundaries between Moment 1 and 2 were not so clear. What for some observers was Moment 1, may have been Moment 2 for others. In addition, the small number of Moment 1 decisions may indicate that, unless observers decided at the very beginning of the video clip (and not at any point within the first third), they generally said they decided at Moment 2. Replications using clear “markers” to separate the time periods of interest, such as questions by an interviewer (answer to the first question: Moment 1, answer to the second question: Moment 2, and so on), an acoustic signal (e.g., before the first beep: Moment 1, first to second beep: Moment 2, etc.), or a time display on the TV screen (e.g., first minute: Moment 1, second minute: Moment 2, and so on) would help us clarify that issue. Work in progress is addressing this point.

## **EXPERIMENT 2**

As noted above, in the experiment just described a strong lie bias was found for early judgments. An interesting question is why observers showed that bias. Actually, the information they had at that early point in time was scarce, apart from the sender’s physical appearance. Is it possible that a person’s appearance influences credibility judgments? There is evidence indicating that this could be the case. For instance, one’s facial appearance has been found to influence a series of attitudes, behaviors, attributions, and judgments made by others (see reviews by Alley, 1988; Berry & Zebrowitz, 1986; Bruce & Young, 1998; Bull, 1982; Bull & McAlpine, 1998; Bull & Rumsey, 1988; Shepherd, 1989; Zebrowitz, 1997). Thus, might it be possible that there is a social stereotype of the appearance of a liar’s face, so that initial credibility judgments are influenced by the extent to which the sender looks like an honest individual or a deceptive one. If so, someone with a liar-looking facial appearance would be judged as deceptive early in his or her statement, but perhaps the availability of information provided by the sender as time goes by can reduce that initial tendency.

Zuckerman, DeFrank, Hall, Larrace, and Rosenthal (1979), found what they termed a demeanor bias in their senders: some were consistently judged as honest and some as deceptive, regardless of whether they lied or told the truth. The existence of a demeanor bias has been confirmed by later research conducted by Bond, Kahler, and Paolicelli (1985). As conceptualized by Zuckerman et al. (1979), that bias would depend on some internal characteristics influencing the sender’s perceptible demeanor, which, in turn, would determine observers’ ratings. Indeed, some authors have tried to see the influence of some personality traits and social skills of the sender upon observers’ credibility judgments (e.g., Geis & Moon, 1981; Miller, deTurck, & Kalbfleisch, 1983; Riggio & Friedman, 1983; Riggio, Tucker, & Widaman, 1987; Riggio, Tucker, & Throckmorton, 1987; Vrij, 1992; Vrij & Winkel, 1993), assuming that these traits and skills influence in some way the behavior displayed by the communicator (for empirical tests of this assumption see Riggio, Tucker, & Widaman, 1987; Vrij, Akehurst, & Morris, 1997). However, as Bond and Robinson (1988) suggest, it may be the case that “these biases originate in fixed features of the mien, an innocent- or guilty-looking visage” (p. 304). If this were the case, then the biased judgments of credibility would depend directly upon the sender’s appearance, instead of depending on some personality traits or social skills that influence behavior. Remember that in Experiment 1 we used only one sender. If that sender had a face that fits the social stereotype for the face of a liar, then her appearance could have been responsible for observers’ initial lie-bias. Later in time, however, behavioral information drawn from the sender’s behavior may have reduced that bias. Specifically, the misleading information provided in the false stories reduced judgments of deceptiveness. Unlike us, Zuckerman et al. (1979) used series of 15-second videotaped segments, too brief a time period to find a reduction in the demeanor bias. That is, all judgments in Zuckerman et al.’s study were made at what in our experiment was Moment 1 (or, at best, what our observers regarded as Moment 2); that is probably why they found such a strong demeanor bias. One of

the aims of the present, follow-up experiment was to check whether an initial lie bias is found again if the sender's face is different from that of study one.

Here we took two of the statements used in Experiment 1 (S1T and S1D). Those statements were presented via audio, while a still face of a young woman, supposedly the one making the statement, appeared on a TV screen. That face could be of the same sender as in Experiment 1, or of two other senders. Comparing the initial accuracy for truthful and deceptive statements for the several purported senders enabled us to check whether the initial lie bias found in Experiment 1 was due to the facial appearance of the sender used in that study.

## Method

**Participants:** The sender was one female undergraduate student of psychology at a Spanish University. Still images of two other senders, also females of similar ages, were used. Observers were 224 police officers studying to become police inspectors at the Police Academy.

**Procedure:** The procedure used to create the statements is described in the method section of Experiment 1. Here in Experiment 2, audio recordings of S1T and S1D were presented. The decision to select the two versions of only one original sequence was prompted by the need to use a limited number of participants<sup>4</sup>. Sequence 1 was chosen because S1T and S1D were entirely different from one another, while S2D was a variation of S2T where central details were changed to make it deceptive. Since, as we shall see later, each participant would have to judge both statements, these had to be entirely different from one another. If we had used S2, judgments for the second statement would not have been independent from those for the first.

Video clips were edited where the audio recordings of S1T and S1D were coupled with a still image of the witness who purportedly had made the statement. This image was of the same sender shown in Experiment 1, or one of two other senders who initially volunteered to participate in our study and had made their statements. These pictures were taken from the tapes of their statements. All senders, as shown in the still images employed, faced the camera and displayed a neutral facial expression.

---

<sup>4</sup> Notice that in Experiment 1 four groups of observers were used (each of these was in turn comprised of a subgroup of undergraduates and another one of police officers). In Experiment 2 three different still faces had to be shown while the same words as in Experiment 1 were heard. This makes 12 groups, too large a number of samples. Therefore, only the statements based on one of the original video sequences were taken for this experiment. This was not a problem, since the strong tendency found in Experiment 1 to make judgments of deceit at moments 1 and 2 but not at Moment 3 was evident for both S1,  $F(1) = 10.28$ ,  $p = .001$ , and S2,  $F(1) = 17.51$ ,  $p = .000$ . Also, police officers were taken as observers in Experiment 2 not only because they were available at the time data were to be collected for that experiment, but also because it increases the external validity of the findings when it comes to generalizing them to real criminal cases. In addition, using only officers as observers was not problematic since, as reported above, in Experiment 1 the tendency to judge statements as deceptive at the beginning of the sender's performance was statistically significant among officers, while at moment three that trend had at best a marginal significance. In fact, chi-square analyses made on the data of Experiment 1 to examine the relation between moment (1 and 2 v. 3) and judgment (truthfulness / deceptiveness judgment) were significant for both undergraduate students,  $F(1) = 13.24$ ,  $p = .000$ , and police officers,  $F(1) = 15.73$ ,  $p = .000$ , and this was so not only for S2, students:  $F(1) = 5.42$ ,  $p = .020$ , officers:  $F(1) = 14.11$ ,  $p = .000$ , but also for S1 which, as mentioned above, was the statement chosen to be used in Experiment 2, students:  $F(1) = 7.78$ ,  $p = .005$ , officers:  $F(1) = 5.09$ ,  $p = .024$ .

Groups and number of observers per group are shown in Table 3. Again, all those police officers who attended their lectures in a given classroom were allocated to the same experimental group. As mentioned in Experiment 1, allocation of officers to their classrooms is based on an alphabetical criterion.

**TABLE 3. Groups And Number Of Observers Per Group In Experiment 2.**

Pairs of statements	Faces		
	A	B	C (same as in Exp. 1)
S1T □ S1D	38	42	36
S1D – S1T	40	42	26

The judgmental sessions were similar to those described in Experiment 1. Statements were presented to observers via a videotape connected with a TV monitor. Observers completed the same questionnaires as in Experiment 1, although some additional questions were added. One asked observers how attractive they found the sender. Answers were collected on a continuous scale from 1 (very unattractive) to 7 (very attractive). Another question asked observers how old they thought the sender was. These two questions were at the end of the questionnaire.

## Results

### *Manipulation Checks*

If we are to analyze how facial appearance influences credibility judgments we must first make sure that our facial stimuli are different from each other in some characteristics likely to influence social judgments. Two such characteristics are age (e.g., Montepare & Zebrowitz, 1998) and attractiveness (e.g., Alley & Hildebrandt, 1988; Zebrowitz, 1997).

*Age.* Observer's ratings of targets' age were 23.47 years for Face A, 24.36 for Face B, and 22.76 for Face C (which was the face of the sender we used in Experiment 1),  $F_{(2,221)} = 6.75$ ,  $p = .001$ . Post-hoc Fisher's LSD tests showed that Face B was judged as significantly older than Face A,  $p = .034$ , and Face C,  $p = .000$ , but Face A was not perceived as significantly older than Face C,  $p = .111$ .

*Attractiveness.* Observers also rated the degree of attractiveness of the faces. Ratings were 4.76 for Face A, 3.98 for Face B, and 4.59 for Face C,  $F_{(2,221)} = 21.79$ ,  $p = .000$ . Fisher's LSD tests showed that Faces A and C were perceived as similarly attractive,  $p = .209$ , but Face B was judged as less attractive than faces A and C, both  $ps = .000$ .

### *Credibility Judgments*

In Experiment 1 a lie bias, which was greater among police officers than among undergraduates (see Garrido et al., 1997, 2003), was found. That bias decreased as observers made their decision about the witness's veracity later in time. Unlike Experiment 1, in this study visible dynamic cues displayed by the sender (e.g., her gestures and body movements) were absent, her facial appearance was manipulated, and all the observers were members of the Spanish National Police Force. In Experiment 2 we addressed the following questions: First, whether in these circumstances a lie bias also appears; second, whether this bias decreases across time; and third, whether this is dependent upon the sender's facial appearance, that is

to say: (a) whether the lie bias is apparent for any of the faces (i.e., that of Experiment 1) but not for the others, and (b) whether the decrease of that bias over time happens when any of the still faces is presented but not when the others are presented.

In order to find an answer to these questions we conducted two backward stepwise hierarchical loglinear analyses, one for the truthful statement and another one for the deceptive one (participants who had judged each statement were exactly the same; the order of the truthful and the deceptive statement was counter-balanced, as shown in Table 3). The variables which were introduced in the analyses were observers' veracity judgment (truthfulness / deceptiveness judgment), the face (A / B / C), and the moment observers said they made their decision concerning the witness's credibility (Moments 1 and 2 / Moment 3)<sup>5</sup>.

For the *truthful statement*, k-way effect tests supported the null hypothesis that third-order effects were equal to zero, likelihood-ratio chi-square:  $\chi^2_{(2)} = 0.77, p = .681$ , and rejected the hypotheses that second-,  $\chi^2_{(5)} = 24.32, p = .000$ , and first-order effects,  $\chi^2_{(4)} = 44.70, p = .000$ , were zero. The best model comprised two second-order interactions: Judgment X Moment, partial  $\chi^2_{(1)} = 15.45, p = .000$ , and Moment X Face, partial  $\chi^2_{(2)} = 8.12, p = .017$ . In addition, the main effect of judgment had a significant partial chi-square,  $\chi^2_{(1)} = 40.38, p = .000$ . The model had an adequate goodness of fit: likelihood-ratio  $\chi^2_{(4)} = 4.11, p = .392$ ; the standardized residual that had a larger absolute value was 0.89. Parameter estimates and their corresponding z values are shown in Appendix 2. In addition, Table 4 shows the observed frequencies and the standardized residuals with reference to the independence model corresponding to the two interactions of the final hierarchical model. The judgment effect indicates that, just as in Experiment 1, the frequency of lie judgments (70.72 %) was larger than the frequency of judgments of truthfulness (29.28 %). This was so regardless of the face that was presented, because the Face X Judgment effect was not significant. However, credibility judgments were actually affected by the moment the decision was made, as indicated by the Judgment X Moment interaction: there was an association between making the decision at Moment 1 or 2 and judging the statement as deceptive, and making it at Moment 3 and judging the statement as truthful (see Table 4 and Appendix 2). Thus, a decrease over time of the lie bias was found here as well. This happened regardless of the face that was presented. However, the face had an effect, not upon whether statements were judged as truthful or deceptive, but on the moment the decision was made: those who watched face A tended strongly to make their decision at the beginning of the statement, while those who watched face B tended moderately to decide at the end. The tendency for face C was not significant<sup>6</sup> (see Table 4 and Appendix 2).

For the *deceptive statement* the results were quite similar. K-way effect tests failed to yield significant results with regard to the third-order effect,  $\chi^2_{(2)} = 2.76, p = .252$ , but not with regard to second-,  $\chi^2_{(5)} = 35.65, p = .000$ , and first-order effects,  $\chi^2_{(4)} = 57.35, p = .000$ . The final model comprised exactly the same interactions as for the truthful statement: Judgment x Moment, partial  $\chi^2_{(1)} = 11.55, p = .001$ , and Moment X Face, partial  $\chi^2_{(2)} = 18.70, p = .000$ . The first-order effect of judgment was significant also in this case,  $\chi^2_{(1)} = 53.52, p = .000$ , indicating

---

<sup>5</sup> Once again, when differentiating between moments 1, 2, and 3, expected frequencies were too small at Moment 1, particularly when judgments of truthfulness were made. Therefore, for both truthful and deceptive statements, moment-1 and moment-2 judgments were taken together and compared with moment-3 judgments.

<sup>6</sup> Two variables in a contingency table (such as Table 4) are related in a cell if the standardized residual in that cell has an absolute value equal to or higher than 1 (Martín, Cabero, & Ardanuy, 1997). Also, two variables are related in a cell such as those of Appendix 2 if the associated z value has an absolute value equal to or higher than 1.96 (e.g., Tabachnick & Fidell, 1996).



that judgments of deceptiveness (73.99 %) exceeded judgments of truthfulness (26.01 %). The fit of the model was good, with a likelihood-ratio chi-square of  $\chi^2_{(4)} = 3.16, p = .531$ , and the greater standardized residual had an absolute value of 0.78. As shown in Table 4 and Appendix 2, regardless of the face which was presented an association between making the decision at moments 1 or 2 and judging the statement as deceptive was found, as well as an association between deciding at Moment 3 and judging the statement as truthful. Also, just as in the former case, those who watched face A tended to make their judgment at moments 1 and 2, those who watched face B tended to make it at Moment 3, and the tendency for face C was not significant.

**TABLE 4. Observed Frequencies And Standardized Residuals With Reference To The Independence Model Corresponding To The Judgment X Moment And The Moment X Face Interactions For The Truthful Statement And The Deceptive Statement.**

Judgment	JUDGMENT X MOMENT					
	Moments 1 and 2			Moment 3		
	Observed	Standardized Residual	Observed	Standardized Residual	Observed	Standardized Residual
<i>Truthful Statement</i>						
Judgments of truthfulness	21	(-2.2)	43	( 2.3)		
Judgments of deceptiveness	95	( 1.4)	62	(-1.5)		
<i>Deceptive Statement</i>						
Judgments of truthfulness	18	(-2.2)	40	( 2.3)		
Judgments of deceptiveness	98	( 1.3)	67	(-1.4)		

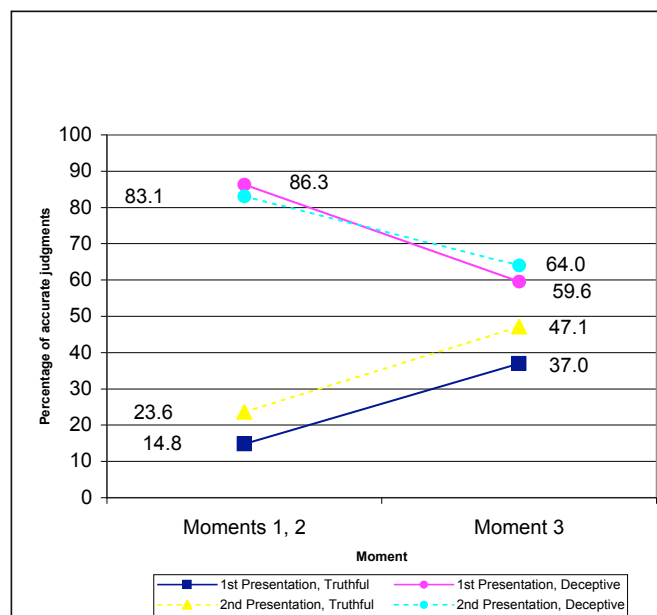
Moment	MOMENT X FACE					
	Face A		Face B		Face C	
	Observed	Standardized Residual	Observed	Standardized Residual	Observed	Standardized Residual
<i>Truthful Statement</i>						
Moments 1 and 2	50	( 1.4)	38	(-0.9)	29	(-0.5)
Moment 3	28	(-1.5)	46	( 1.0)	32	( 0.6)
<i>Deceptive Statement</i>						
Moments 1 and 2	56	( 2.4)	30	(-2.1)	31	(-0.2)
Moment 3	22	(-2.5)	54	( 2.2)	31	( 0.3)

In summary, as was the case in Experiment 1, Experiment 2, for both the truthful (S1T) and the deceptive (S1D) statements found: (a) a lie bias, (b) this bias decreased over time, and (c) neither of these effects was influenced by the purported witness' facial appearance. An influence of the facial appearance upon the moment decisions were made was found as well: both when the statement was truthful and when it was deceptive, face A judgments were made at the beginning of the statement, and face B judgments were made at the end.

*Detection Accuracy*

In the preceding paragraph the conclusion was drawn that there was a strong tendency to say the sender was lying. This should have an influence on accuracy, so that judgments of the truthful statement should be wrong more often than judgments of the deceptive statement, and the latter should be accurate more often than the former. In addition, we have seen that this lie bias tended to decrease over time. Therefore, the trends towards judging incorrectly of the truthful statement and guessing correctly those of the deceptive statement should decrease over time as well.

To examine those questions two backward stepwise hierarchical loglinear analyses were calculated, one for the first statement that was presented (first presentation) and the other for the second (second presentation) (some observers watched S1T first, and then S1D; other watched the clips in the reverse order; see Table 3). The variables which were introduced in the analyses were Moment (1 and 2 *v.* 3), Value of Truth, and Hit / Miss. In both cases the k-way effect test indicated that the third order interaction was significant: likelihood-ratio chi-squares were:  $\chi^2_{(1)} = 17.54, p = .000$ , for the first presentation, and  $\chi^2_{(1)} = 11.83, p = .001$ , for the second. Consistent with these results, in neither case did the process continue beyond the saturated model. As shown in Appendix 3, the Value of Truth X Hit/Miss interaction was substantial. Both for the first presentation, partial  $\chi^2_{(1)} = 51.29, p = .000$ , and for the second presentation, partial  $\chi^2_{(1)} = 36.57, p = .000$ , judgments of the truthful statement tended to be wrong, and those of the deceptive statement tended to be accurate. In fact, the percentage of accurate judgments of the truthful statement was only 25.22 % in the first presentation, and 34.91 % in the second; the corresponding values for the deceptive statement were 72.22 % and 74.78 % (see Figure 1). Therefore, the lie bias had a strong effect upon accuracy. However, this effect was influenced by the moment the decision was made, as indicated by the Moment X Value of Truth X Hit/Miss effect in both analyses (see Appendix 2): it was stronger at the beginning of the statement than at the final moment, as shown in Figure 1. In fact, when the decision was made at the initial moments, there was a larger proportion of accurate judgments for the deceptive account than for the truthful, for the first presentation:  $\chi^2_{(1)} = 23.11, p = .000$ ; for the second presentation:  $\chi^2_{(1)} = 25.09, p = .000$ . However, when the decision was made at the end, although the proportion of accurate judgments of



**Figure 1. Accuracy rates on judging the credibility of truthful and deceptive statements at moments 1 and 2, as well as at Moment 3, for the first and the second presentations of Experiment 2.**

the deceptive account was still somewhat larger than that of the truthful, one this difference was not statistically significant, for the first presentation:  $\chi^2_{(1)} = 3.63, p = .057$ ; for the second presentation:  $\chi^2_{(1)} = 1.14, p = .285$ . These effects are clearly shown in Figure 1.

Similarly, in the early moments the proportion of errors upon judging the truthful statement was significantly larger than the proportion of accurate judgments for both the first presentation:  $\chi^2_{(1)} = 30.31, p = .000$ , and for the second presentation:  $\chi^2_{(1)} = 15.29, p = .000$ . On the contrary, the proportion of errors at judging the false statement was smaller than the proportion of correct judgments for both the first presentation:  $\chi^2_{(1)} = 26.84, p = .000$ , and for the second presentation:  $\chi^2_{(1)} = 28.45, p = .000$ . In the final moment the differences were in the same direction, but the residuals and, consequently, the significance, decreased with respect to the initial moments: truthful statement in first presentation:  $\chi^2_{(1)} = 3.63, p = .057$ , in second presentation:  $\chi^2_{(1)} = 0.18, p = .674$ ; deceptive statement in first presentation:  $\chi^2_{(1)} = 2.12, p = .145$  and in second presentation:  $\chi^2_{(1)} = 3.92, p = .048$ .

In conclusion: the overall lie bias made accuracy for the deceptive statement higher than accuracy for the truthful one. As this bias decreased over time, so did the tendency to be more accurate when judging the deceptive statement than when judging the truthful one.

## Discussion

The second experiment was planned as a follow-up study after the first one. Some experimental conditions of Experiment 1 were changed to check whether Experiment 1 results concerning the moment changed. Thus we hoped to identify the factors determining such results. More specifically, the questions addressed were: (a) Does the lie bias found in the first experiment hold true when dynamic visible cues (gestures and body movements) are suppressed from the videotaped statement?, (b) Does this bias decrease as observers decide later in time, just as happened in the previous experiment?, and (c) Does the existence of the lie bias depend upon the sender's facial appearance, so that a demeanor bias is in operation? To find an answer to those questions a truthful and a deceptive statement of Experiment 1 were presented to observers in their audio format, accompanied by a still image of the person who, supposedly, had enacted the statements. This image could be of the sender who had been used in Experiment 1, or of one of two other young women. Observers had to indicate whether each statement was truthful or deceptive and at what moment they had come to their conclusion on the veracity of the statement, at the beginning (Moment 1), middle (Moment 2) or end (Moment 3) of the videotaped statement.

Analyses were performed to explore the relationships between credibility judgments, the moment at which they were made, and the still face being shown. Results indicate that, overall, the lie bias found in Experiment 1 appeared in Experiment 2 as well: both when judging the truthful statement and when judging the deceptive one, the number of deception judgments was substantially larger than the number of truthfulness judgments. Consequently, there was an association between judging the truthful statement and doing so incorrectly, and between judging the deceptive statement and guessing it correctly. Now, did this effect hold true for the various faces, or only for some of them? And, did it depend in any way on the moment when the judgment was made? Our data indicate that: (a) the decision-making moment had an influence on judgments: the lie bias decreased as time went by, and (b) the witness' facial appearance did not affect credibility judgments.

### *The Influence Of Moment Upon Veracity Judgments And Accuracy*

When the decision was made at moments 1 or 2, a tendency to say statements were deceptive was found; this tendency decreased when the decision was made at Moment 3. This

made early accuracy rather high for the deceptive statement and rather poor for the truthful one, but at Moment 3 these differences had lost significance, although deceptive-statement judgments continued being slightly more accurate than truthful-statement judgments<sup>7</sup>. It is interesting that the Judgment X Moment interaction was significant not only in the loglinear analysis calculated for the false statement, but also in the one calculated for the truthful. Remember that in Experiment 1 the predicted increase in accuracy for truthful statements failed to reach significance, that is, the frequency of truthfulness judgments at Moment 3 was not higher than its frequency at moments 1 and 2. However, in the present experiment, regardless of the still face being shown, accuracy for truthful statements increased significantly over time<sup>8</sup>. Also, the effect already detected in Experiment 1 consisting of a decrease when judging deceptive statements was found here as well.

What reason may account for the fact that the formerly predicted increase in accuracy for the truthful statements did not emerge in Experiment 1 but did appear in Experiment 2? This prediction was based on the assumption that, at the end of the statement, there would be a maximum amount of available accurate information when truthful statements were being presented, while at this same moment the misleading information would reach its maximum when deceptive statements were being presented. This would result in a progressive increase in accuracy over time when judging the truthful statements, coupled with a decrease when judging the deceptive ones.

With this in mind, we should point out that research has shown that verbal cues are the most useful when it comes to making credibility judgments, while nonverbal indicators (gestures and movements) are in general the most misleading (see meta-analyses by DePaulo, Zuckerman, & Rosenthal, 1980a; DePaulo et al., 1985; Kalbfleisch, 1985; Zuckerman et al., 1981). If both kinds of information (i.e., visual and verbal indicators) are presented at the same time, observers probably do not pay much attention to verbal cues, which are the most useful, attending instead to the visual information, which is the most misleading. This would be consistent with the distraction hypothesis, the information overload hypothesis, or the situational familiarity hypothesis. The *distraction hypothesis* posits that visual cues would distract observers from processing verbal and vocal information (Maier & Thurber, 1968; Miller, Bauchner, Hocking, Fontes, Kaminski, & Brandt, 1981; Miller & Stiff, 1993). The *information overload hypothesis* maintains that processing all incoming information would cause a cognitive overload in observers, who therefore would block out or overlook important cues (Bauchner, Kaplan, & Miller, 1980; Miller et al., 1981; Stiff, Miller, Sleight, Mongeau, Garlick, & Rogan, 1989). Both of these hypotheses, posited to account for the poorer accuracy rates attained when visual cues are present as compared to those situations where they are absent, predict that observers do not process the verbal information. However, Stiff et al. (1989) found that verbal information was processed by observers, although it was not used to make credibility judgments. The authors found partial support for an alternative hypothesis: the

---

<sup>7</sup> Overall accuracy, that is, accuracy collapsing across the truthful and the deceptive statement, was always close to chance probability. Chi-square analyses on the hit / miss frequencies in neither case yielded statistically significant results; for the First Presentation:  $\chi^2(1) = 0.32, p = .571$ , for Moment-1-and-2 Judgments;  $\chi^2(1) = 0.08, p = .776$ , for Moment-3 Judgments; for the Second Presentation:  $\chi^2(1) = 1.63, p = .201$ , for Moment-1-and-2 Judgments;  $\chi^2(1) = 1.20, p = .274$ , for Moment-3 Judgments. Similarly, in neither of the loglinear analyses that examined the relations among Moment, Value of Truth, and the Hit/Miss variable, was the Moment X Hit/Miss second-order effect significant. These results lend further support to Levine, Park, and McCornack's (1999) arguments in favour of examining, in the field of the detection of deception, the separate accuracy for truthful and deceptive communications instead of focusing on the overall accuracy rate.

<sup>8</sup> It is important to keep in mind that this lack of significance in Experiment 1 was also apparent for S1T, which was the truthful account we used in the present study.

*situational familiarity hypothesis*, which maintains that observers in familiar situations use verbal cues, since they can “visualize” the situation and assess the validity of verbal information (systematic processing), while observers in unfamiliar situations use, to some extent, nonverbal information, because there is little basis for evaluating the verbal content (heuristic processing). The observers we used were unfamiliar with the situation. In Experiment 1 the visual information was available; thus, they may have relied too much on that kind of information as a basis for their judgments. Results from further explorations on the data of Experiment 1 support this explanation: when asked to indicate the cues they had used to make their judgments, observers reported significantly more nonverbal indicators, especially visual ones, than verbal cues (Masip, Garrido, & Rojas-Díaz, 2001; see also Garrido, Masip, Herrero, & Rojas-Díaz, 2000). This attention devoted exclusively to visual indicators may have caused accuracy for deceptive statements to decrease over time (as an increasing number of misleading visual indicators were being shown), while accuracy at judging the truthful statements did not rise in Experiment 1, since the most revealing cues are verbal, that is, just those cues observers did not attend to in that experiment. However, dynamic visual information, more misleading than the verbal one, was absent in Experiment 2. This may have led observers to pay close attention to the verbal cues, as well as to process those cues. This, in turn, may have contributed to the increased accuracy at judging the truthful statement at Moment 3.

This explanation should nevertheless be taken with caution. First, since verbal information is less misleading than the visible, hiding the latter should not only have resulted in an increase in accuracy over time for the truthful statement, but in addition it should have restricted the decrease in accuracy for the deceptive account. However, that decrease was significant not only in Experiment 1, but in Experiment 2 as well. A possible reason for that is that, after all, our sender was able to successfully control her verbal behavior. Second, despite experimental results showing the superiority of verbal information, as compared with the nonverbal, when it comes to making veracity judgments, our own research shows verbal cues may be processed in a biased manner. This, in turn, may have an effect upon the credibility judgments. For instance, police officers who participated as observers in Experiment 1 said the statements were implausible and contained verbal contradictions, while undergraduate students said they were plausible and verbally consistent. That is to say, each group of observers mentioned verbal indicators that were opposite to those mentioned by the other group, despite the fact that they all had been shown exactly the same videotapes. As a result of these perceptions the officers’ tendency to judge the statements as deceptive was stronger than the undergraduates’, and the latter’s tendency to judge them as truthful was stronger than among officers. Similar results were found for a few nonverbal indicators (Garrido et al., 2000; Masip et al., 2001). Third, it would be inadequate to generalize from this second experiment (which was quite modest –its only pretension was to clarify some results found in Experiment 1–, where only two statements, both of them based on the same sequence, both of which were enacted by the same sender, were used) to other statements, witnesses, and situations. Caution is therefore strongly warranted when interpreting the results reported here, at least until further research replicates them.

### *Witnesses’ Facial Appearance*

In the discussion of Experiment 1 it was suggested that the lie bias, which was particularly strong at the beginning of the statement, could be caused by the sender’s facial appearance. Therefore, in the present experiment several different faces were shown, to examine whether the lie bias of Experiment 1 or its time variation were influenced by the witness’s appearance. However, contrary to our predictions, the senders’ facial appearance had no influence either upon the overall lie bias, or upon its reduction over time. Therefore, these effects do not depend on the witness’s facial appearance, at least, not for the range of faces used in this study. They are not influenced by the witness’ visible behavior (gestures and body movements) either,

because that behavior was not shown in this experiment. Thus, they must be caused by verbal and paralinguistic cues, which were available in both studies.

The lack of influence of the face may nevertheless be due to several reasons. First, this was a competitive situation where observers were somewhat challenged to spot senders' lies. This differs from the cooperative interactions the average citizen is involved in his/her daily life, where truth is taken for granted and there is no motive to suspect the other is being deceptive. The nature of the task (detecting deception) may have raised observers' "state" suspicion (Levine & McCornack, 1991), making lie judgments more likely (Burgoon, Buller, Ebesu, & Rockwell, 1994; Stiff et al., 1992; Toris & DePaulo, 1984; Zuckerman, Driver, & Guadagno, 1985, p. 165), regardless of other factors such as the witnesses' facial appearance. Second, observers in this experiment were police officers. Garrido et al. (1997, 2003) and Sanderson (1978, cited by Bull, 1989) found a lie bias in officers' credibility judgments. Officers were also more accurate at judging lies than truths in recent studies conducted by Ekman, O'Sullivan and Frank (1999) and Porter, Woodworth and Birt (2000). Burgoon et al.'s (1994) military experts displayed a similar bias. It was suggested earlier that experts may hold a generalized communication suspicion (Levine & McCornack, 1991; see also Burgoon et al., 1994), which could increase their lie judgments. For instance, O'Sullivan, Ekman, and Friesen (1988) stated that "observers with a deception bias, because of their professional experience, for example as police officers or lawyers, may be more likely to view all behavior as deceptive and therefore have a heuristic which will permit them to classify deceptive behavior correctly, but which will be misleading in evaluating honest behavior" (p. 214). In addition, as suggested above, it appears reasonable that a lie bias heuristic could emerge for police officers, in the same way a truth bias cognitive heuristic emerges, according to Stiff et al. (1992), for relational partners. It may be that the strong lie bias displayed by our officers prevented them from being influenced by the subtle differences that existed between our senders. Perhaps students, whose lie bias was weaker, would have been sensitive to changes in the senders' facial appearance. Third, it may be that the differences in facial appearance between our senders were too small to have an influence upon credibility judgments. Despite the fact that observers' perceptions of their ages and physical attractiveness differed from one face to one another, all faces were perceived as in their twenties. Perhaps if a child's face, the face of a young person, that of a mature one, and an elderly person's face had been used very different results would have emerged. Also, the attractiveness of all three senders was close to average, ranging from 3.98 (face B) to 4.76 (face A), a rather small range on a 1-to-7-point scale. And, in addition, it is not only physical attractiveness that influences social judgments, but also mistaken identities, animal analogies, sickness similarities, babyfacedness, etc. (Zebrowitz, 1997). For example, recent research shows that, controlling for attractiveness, age and babyfacedness influence attributions of a series of traits and behavioral tendencies, including truthfulness / deceptiveness (Masip, Garrido, & Herrero, 2003a). Also, these facial characteristics have been found to influence the credibility judgments of written statements (Masip, Garrido, & Herrero, 2003b).

Finally, it could be argued that maybe participants did not pay any attention to the faces being shown, perhaps because they were fully aware that a still face with a neutral expression provides little information on whether the sender is lying or telling the truth. However it is unlikely that participants did not attend to the faces, because although facial appearance had no effect upon credibility judgments, they influenced the moment at which the decision was made. Regardless of the statement being judged (the truthful one or the deceptive one), when face A was being shown decisions were made at moments 1 or 2, and when face B was being shown decisions were made at the final moment. No clear tendencies emerged for Face C. Faces A and B differed from each other both in terms of the age observers perceived in them and in attractiveness. Therefore, it appears that any of these two tendencies could account for the moment differences between Faces A and B. However, the perceived age of Face C did not differ significantly from that of Face A, and, just as Face A, it was perceived to be younger than Face B. Something similar happened with reference to attractiveness: Faces A and C did not

differ from each other in this characteristic, and they both were judged as significantly more attractive than Face B. Therefore, if differences found between Faces A and B were due to age or attractiveness, Face C would have lined up with Face A, and this did not happen (actually, its non-significant tendencies were in the same direction as those of Face B). Therefore, the differences between the faces in terms of the moment the decision was made were not due to age or attractiveness, but, rather, to some other facial characteristic that was not taken into account in the present experiment.

## **General Discussion**

### *Officers Versus Non-Officers*

Contrary to our first prediction, police officers tended to make their decision as to whether the sender was truthful or deceptive later than non-officers. Although this effect was non-significant, it is possible that officers, knowing that collecting information from witnesses is important, paid attention to the sender's behavior for a longer time than undergraduates, thus deciding later than the students. However, their pronounced lie bias (Garrido et al., 1997, 2003), suggests that they were incapable of either collecting the right information or using it correctly to make accurate credibility judgments. In any case, these results do not discard the hypothesis that an a priori belief that the sender is deceptive can have an effect on officers' judgments. Indeed, it may be the case that police officers, instead of failing to process the incoming behavioral information, as they would do if, as suggested before, a cognitive heuristic such as that identified by Stiff et al. (1992) were in operation, actually do process that information, but in a biased manner aimed at finding support for their prior conceptions that the sender is being untruthful (conceptions based, for instance, on a GCS). In that case, officers would be unwittingly subjected to confirmation bias: "the tendency to interpret, seek, and create information in ways that verify existing beliefs" (Brehm & Kassin, 1993, p. 129). Recent, still unpublished research lends support to this idea (Masip, 2002; Masip et al., 2001).

### *Accuracy Over Time*

In the studies reported here, a decrease in observers' accuracy at detecting deceptive statements was found as time went by, coupled with a similar increase in detecting truthful accounts, particularly when dynamic visual cues were not available to observers. This is probably due to the greater amount of misleading information in false performances as time goes by, and the greater amount of accurate information in the truthful accounts. Those time variations question the generalizability of findings of previous research, since most experiments on nonverbal detection of deception have been conducted using very small behavioral samples. It is apparent that receivers' detection accuracy depends on the moment in a long statement at which they make their decision, and it interacts with the value of truth of the statement: the moment truthful accounts are best detected is the same moment at which deceptive accounts are least detected, while overall accuracy is close to chance probability at any point in time.

### *Visual Versus Verbal Information*

Our data seem to indicate that visual information prevented observers from properly using the growing verbal information that was presented in truthful statements. This is consistent with the distraction hypothesis, the information overload hypothesis, and the situational familiarity hypothesis, as well as with previous research showing the relative usefulness of verbal cues, compared to nonverbal ones, in judging credibility. Both when dynamic visual information was shown and when it was not available observers' accuracy at detecting deceptive accounts decreased as time went by. This suggests that, although extant research has shown that nonverbal cues are more misleading than verbal ones, it seems that the audio

channel (which conveys verbal and vocal information) can be successfully controlled by the sender in order to create a plausible lie and to appear honest. In fact, Ekman (1981) hypothesized that the verbal content would be very amenable to control, although meta-analyses show that the verbal information is very useful for making accurate credibility judgments: "the power (i.e., the accuracy) of the word, either written or spoken" (Zuckerman et al., 1981, p. 27).

### *Facial Appearance And Credibility Judgments*

An accuracy level close to chance probability was expected for initial judgments in Experiment 1. However, a strong tendency to judge statements as deceptive was apparent for these early judgments. This tendency decreased over time. Since the only information available at the beginning of a statement is the sender's physical appearance, it was suggested that our witness's facial appearance could be responsible for that initial lie bias or its variation over time. However, we found no evidence of a face effect in the form of a demeanor bias in Experiment 2. Neither the existence of a lie bias nor its decrease over time was affected by the sender's facial appearance. However, the three faces that were used in Experiment 2 fell within the same age range and were close to average attractiveness, despite the significant differences that were found in the observers' attractiveness and age ratings. Research on how facial stereotypes may influence credibility judgments must be conducted. Recently, we completed a series of three experiments addressing this issue (Masip & Garrido, 2001b; Masip, et al., 2003a,b).

### *Caveats And Further Research*

It must be acknowledged that the two studies reported here suffer from several methodological disadvantages. Aside from the problem of having used facial stimuli with rather small variations in the relevant facial features (age and babyfacedness indicators), three points must be mentioned here. First: those observers who took their decision at a given moment were not the same as those who decided at any other moment. This raises a question: were the differences found over time due to the influence of the moment variable or were they due to the impact of differences between the respondents who decided at different times? In addition, observers were not randomly assigned to the different moments, but they were free to make their decision at the time they preferred. Then, can we confidently assert that there is a strong initial lie bias or, rather, what happens is that those observers with the strongest lie bias decide at Moment 1 or Moment 2? This is unlikely, since officers, who were the most biased in Garrido et al.'s (1997, 2003) study, did not tend to decide early, but were the most biased at any moment in time. Certainly, that issue deserves further exploration. Second: As pointed out in the discussion of Experiment 1, the distinction between the different moments was a subjective one. Research on the influence of time on credibility judgments should use clear markers to differentiate between the moments of interest. Third: using only one sender is inappropriate. Several faces were used in Experiment 2, but in both experiments the speech was of the same person. The time profile found in both studies could be due to the verbal and/or paralinguistic idiosyncrasies of that sender. Thus, caution is warranted before generalizing these results to other senders.

In view of these problems the authors are about to conclude a study where a relatively large sample of senders (both males and females) watched videotapes depicting a theft. Later on, they were interviewed twice about the facts they had witnessed. In one case they had to tell the truth, in the other case they had to lie. Each interview had three questions. The answer to the first question was regarded as Moment 1, the answer to the second question as Moment 2, and the answer to the last question was taken as Moment 3. Witnesses' videotaped responses were shown to observers who had to judge the credibility of each statement three times: after watching the first answer, after watching the second one, and after watching the third



(definitive judgment). This design overcomes the problems of the two experiments described in this report. Indeed, its results will shed further light on the effect of the moment observers make their decision on credibility judgments and accuracy.

### References

- Alley, T. R. (1988). Physiognomy and social perception. In T. R. Alley (Ed.), *Social and applied aspects of perceiving faces* (pp. 167-186). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Alley, T. R., & Hildebrandt, K. A. (1988). Determinants and consequences of facial aesthetics. In T. R. Alley (Ed.), *Social and applied aspects of perceiving faces* (pp. 167-186). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bauchner, J. E., Kaplan, E. A., & Miller, G. R. (1980). Detecting deception: The relationship of available information to judgmental accuracy in initial encounters. *Human Communication Research*, 6(3), 253-264.
- Becerra, A., Sánchez, F., & Carrera, P. (1989). Indicadores aislados versus patrón general expresivo en la detección de la mentira. *Estudios de Psicología*, 38(1), 21-29.
- Berry, D. S., & Brownlow, S. (1989). Were the physiognomists right? Personality correlates of facial babyishness. *Personality and Social Psychology Bulletin*, 15(2), 266-279.
- Berry, D. S., & McArthur, L. Z. (1985). Some components and consequences of a babyface. *Journal of Personality and Social Psychology*, 48(2), 312-323.
- Berry, D. S., & Zebrowitz, L. A. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, 100(1), 3-18.
- Bond, C. F. Jr., Kahler, K. N., & Paolicelli, L. M. (1985). The miscommunication of deception: An adaptive perspective. *Journal of Experimental Social Psychology*, 21, 331-345.
- Bond, C. F. Jr., & Robinson, M. (1988). The evolution of deception. *Journal of Nonverbal Behavior*, 12(4), 295-307.
- Brehm, S. S., & Kassir, S. M. (1993). *Social psychology. Second edition*. Boston: Houghton Mifflin Company.
- Bruce, V., & Young, A. (1998). *In the eye of the beholder. The science of face perception*. Oxford: Oxford University Press.
- Bull, R. (1982). Physical appearance and criminality. *Current Psychological Reviews*, 2, 269-282.
- Bull, R. (1989). Can training enhance the detection of deception? In J. C. Yuille (Ed.), *Credibility assessment*. (pp. 83-99). Dordrecht: Kluwer Academic Publishers.
- Bull, R., & McAlpine, S. (1998). Facial appearance and criminality. In A. Memon, A. Vrij, & R. Bull (Eds.), *Psychology and law. Truthfulness, accuracy and credibility* (pp. 59-76). London: McGraw-Hill.
- Bull, R., & Rumsey, N. (1988). *The social psychology of facial appearance*. New York: Springer-Verlag.
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory*, 6(3), 203-242.

- Buller, D. B., & Burgoon, J. K. (1998). Emotional expression in the deception process. In P. A. Andersen., & L. Guerrero. (Eds.), *Handbook of communication and emotion. Research, theory, applications and contexts* (pp. 381-402). San Diego.: Academic Press.
- Burgoon, J. K., Buller, D. B., Ebesu, A. S., & Rockwell, P. (1994). Interpersonal deception: V. Accuracy in deception detection. *Communication Monographs*, *61*, 303-325.
- DePaulo, B. M. (1991). Nonverbal behavior and self-presentation: A developmental perspective. In R. S. Feldman, & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 351-397). Cambridge: Camnbridge University Press.
- DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*(2), 203-243.
- DePaulo, B. M., & Kirkendol, S. E. (1989). The motivational impairment effect in the communication of deception. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 51-70). Dordrecht: Kluwer Academic Publishers.
- DePaulo, B. M., & Pfeiffer, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, *16*(3), 249-267.
- DePaulo, B. M., & Rosenthal, R. (1979). Telling lies. *Journal of Personality and Social Psychology*, *37*(10), 1713-1722.
- DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), *The self and social life* (pp. 323-370). New York: McGraw-Hill.
- DePaulo, B. M., Zuckerman, M., & Rosenthal, R. (1980a). Detecting deception. Modality effects. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 1, pp. 125-162). London: Sage.
- DePaulo, B. M., Zuckerman, M., & Rosenthal, R. (1980b). Humans as lie detectors. *Journal of Communication*, *30*, 129-139.
- DePaulo, P. J., DePaulo, B. M., Tang, J., & Swaim, G. W. (1989). Lying and detecting lies in organizations. In R. A. Giacalone, & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 377-393.). Hillsdale, NJ: Lawrence Erlbaum.
- Ekman, P. (1981). Mistakes when deceiving. *Annals of the New York Academy of Sciences*, *364*, 269-278.
- Ekman, P. (1989). Why lies fail and what behaviors betray a lie. In J. C. Yuille (Ed.), *Credibility assessment* (pp. 71-81). Dordrecht: Kluwer Academic Publishers.
- Ekman, P. (1992). *Telling lies. Clues to deceit in the marketplace, politics, and marriage*. (2nd. ed.). New York: W. W. Norton & Company.
- Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*, 88-106.
- Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, *29*(3), 288-298.
- Ekman, P., & O'Sullivan, M. (1989). Hazards in lie detection. In D. C. Raskin (Ed.), *Psychological methods in criminal investigation and evidence* (pp. 253-280). New York: Springer.
- Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, *46*(9), 913-920.

- Ekman, P., O'Sullivan, M., & Frank, M. (1999). A few can catch a liar. *Psychological Science, 10*(3), 263-266.
- Ford, C. V. (1996). *Lies! Lies!! Lies!!! The psychology of deceit*. Washington, DC: American Psychiatric Press.
- Garrido, E., & Masip, J. (1999). How good are police officers at spotting lies? *Forensic Update, 58*, 14-21.
- Garrido, E., & Masip, J. (2001). Previous exposure to the sender's behavior and accuracy at judging credibility. In R. Roesch, R. R. Corrado, & R. Dempster (Eds.), *Psychology in the courts. International advances in knowledge* (pp. 271-287). London: Routledge.
- Garrido, E., Masip, J., & Herrero, C. (2003). *Police officers' credibility judgments: Accuracy and estimated ability*. Submitted for review.
- Garrido, E., Masip, J., Herrero, C., & Rojas-Díaz, M. (2000). La detección del engaño a partir de claves conductuales por agentes de policía. In A. Ovejero, M. V. Moral, & P. Vivas (Eds.), *Aplicaciones en psicología social* (pp. 97-105). Madrid: Biblioteca Nueva.
- Garrido, E., Masip, J., Herrero, C., & Taberner, C. (1997). *Policemen's ability to discern truth from deception of testimony*. Paper presented to the 7<sup>th</sup> European Conference of Psychology and Law, Stockholm, 3-6 September 1997.
- Geis, F. L., & Moon, T. H. (1981). Machiavellianism and deception. *Journal of Personality and Social Psychology, 41*(4), 766-775.
- Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research, 11*, 335-364.
- Henderson, J., & Hess, A. K. (1982). *Detecting deception: The effects of training and socialization levels on verbal and nonverbal cue utilization and detection accuracy*. Unpublished manuscript. Auburn University, Auburn, AL.
- Kalbfleisch, P. J. (1985). *Accuracy in deception detection: A quantitative review*. Doctoral Dissertation: Michigan State University.
- Kalbfleisch, P. J. (1992). Deceit, distrust and the social milieu: Application of deception research in a troubled world. *Journal of Applied Communication Research, 20*(3), 308-334.
- Köhnken, G. (1987). Training police officers to detect deceptive eyewitness statements: Does it work? *Social Behaviour, 2*, 1-17.
- Köhnken, G. (1989). Behavioral correlates of statement credibility: Theories, paradigms, and results. In H. Wegener, F. Lösel, & J. Haisch (Eds.), *Criminal behavior and the justice system* (pp. 271-289). London: Springer-Verlag.
- Kraut, R. (1980). Humans as lie detectors. *Journal of Communication, 30*, 209-216.
- Kraut, R., & Poe, D. (1980). Behavioral roots of person perception: The deception judgments of customs inspectors and laymen. *Journal of Personality and Social Psychology, 39*(5), 784-798.
- Leekam, S. R. (1992). Believing and deceiving: Steps to becoming a good liar. In S. J. Ceci, M. D. Leichtman, & M. Putnick (Eds.), *Cognitive and social factors in early deception* (pp. 47-62). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

- Levine, T. R., & McCornack, S. A. (1991). The dark side of trust: Conceptualizing and measuring types of communicative suspicion. *Communication Quarterly*, 39, 325-339.
- Maier, N. R. F., & Thurber, J. A. (1968). Accuracy of judgments of deception when an interview is watched, heard, and read. *Personnel Psychology*, 21, 23-30.
- Martín, Q., Cabero, M. T., & Ardanuy, R. (1997). *Paquetes estadísticos SPSS 8.0. Bases teóricas. Prácticas propuestas, resueltas y comentadas*. Salamanca: Editorial Hespérides.
- Masip, J. (2002). *La evaluación de la credibilidad del testimonio a partir de los indicadores conductuales en el contexto jurídico penal*. Unpublished Doctoral Dissertation. Department of Social Psychology and Anthropology. University of Salamanca.
- Masip, J., & Garrido, E. (1999). Evaluación psicológica de la credibilidad: Contextualización teórica y paradigmas evaluativos. In A. P. Soares, S. Araujo, & S. Caires (Eds), *Avaliação psicológica: Formas e contextos* (Vol. VI, pp. 504-526). Braga: Associação dos Psicólogos Portugueses (APPORT).
- Masip, J., & Garrido, E. (2000). La evaluación de la credibilidad del testimonio en contextos judiciales a partir de indicadores conductuales. *Anuario de Psicología Jurídica*, 10, 93-131.
- Masip, J., & Garrido, E. (2001a). La evaluación psicológica de la credibilidad del testimonio. In F. Jiménez (Ed.), *Evaluación psicológica forense, 1. Fuentes de información, abusos sexuales, testimonio, peligrosidad y reincidencia* (pp. 141-204). Salamanca: Amarú.
- Masip, J., & Garrido, E. (2001b). Is there a kernel of truth in judgements of deceptiveness? *Anales de Psicología*, 17(1), 101-120. (Available online at: [http://www.um.es/facpsi/analesps/v17/v17\\_1/08-17\\_1.pdf](http://www.um.es/facpsi/analesps/v17/v17_1/08-17_1.pdf)).
- Masip, J., Garrido, E., & Herrero, C. (2003a). *Facial appearance and impressions of credibility: The effects of facial babyishness and age on person perception*. Submitted for review.
- Masip, J., Garrido, E., & Herrero, C. (2003b). *Facial appearance and judgments of credibility: The effects of facial babyishness and age on statement credibility*. Submitted for review.
- Masip, J., Garrido, E., & Rojas-Díaz, M. (2001). *Beliefs about indicators of deception and truthfulness in a specific situation*. Paper presented to the 11th European Conference of Psychology and Law, Lisbon, 5-8 June 2001.
- Miller, G. R., Bauchner, J. E., Hocking, J. E., Fontes, N. E., Kaminski, E. P., & Brandt, D. R. (1981). ...and nothing but the truth. How well can observers detect deceptive testimony? In B. D. Sales (Ed.), *Perspectives in law and psychology, vol. II: The trial process*. (pp. 145-179). New York.: Plenum Press.
- Miller, G. R., & Burgoon, J. K. (1982). Factors affecting assessments of witness credibility. In N. Kerr, & R. Bray (Eds.), *The psychology of the courtroom* (pp. 169-194). New York: Academic Press.
- Miller, G. R., deTurck, M. A., & Kalbfleisch, P. J. (1983). Self-monitoring, rehearsal, and deceptive communication. *Human Communication Research*, 10, 97-107.
- Miller, G. R., & Stiff, J. B. (1992). Applied issues in studying deceptive communication. In R. S. Feldman (Ed.), *Applications of nonverbal theories and research* (pp. 217-237). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Miller, G. R., & Stiff, J. B. (1993). *Deceptive communication*. Newbury Park: Sage.
- Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. *Advances in Experimental Social Psychology*, 30, 93-161.

- O'Sullivan, M., Ekman, P., & Friesen, W. V. (1988). The effect of comparisons on detecting deceit. *Journal of Nonverbal Behavior*, *12*(3), 203-215.
- Porter, S., Woodworth, M., & Birt, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior*, *24*(6), 643-658.
- Riggio, R. E., & Friedman, H. S. (1983). Individual differences and cues to deception. *Journal of Personality and Social Psychology*, *45*(4), 899-915.
- Riggio, R. E., Tucker, J., & Throckmorton, D. (1987). Social skills and deception ability. *Personality and Social Psychology Bulletin*, *13*, 568-577.
- Riggio, R. E., Tucker, J., & Widaman, K. F. (1987). Verbal and nonverbal cues as mediators of deception ability. *Journal of Nonverbal Behavior*, *11*(3), 126-145.
- Shepherd, J. (1989). The face and social attribution. In A. W. Young, & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 398-409). Amsterdam: North Holland.
- Stiff, J. B., Kim, H. J., & Ramesh, C. N. (1992). Truth biases and aroused suspicion in relational deception. *Communication Research*, *19*(3), 326-345.
- Stiff, J. B., Miller, G. R., Sleight, C., Mongeau, P., Garlick, R., & Rogan, R. (1989). Explanations for visual cue primacy in judgments of honesty and deceit. *Journal of Personality and Social Psychology*, *56*(4), 555-564.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
- Toris, C., & DePaulo, B. M. (1984). Effects of actual deception and suspiciousness of deception on interpersonal perceptions. *Journal of Personality and Social Psychology*, *47*(5), 1063-1073.
- Vrij, A. (1992). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *The Journal of Social Psychology*, *133*(5), 601-610.
- Vrij, A. (1998). Nonverbal communication and credibility. In A. Memon, A. Vrij, & R. Bull. (Eds.), *Psychology and law. Truthfulness, accuracy and credibility* (pp. 32-58). New York: McGraw-Hill.
- Vrij, A. (2000). *Detecting lies and deceit. The psychology of lying and the implications for professional practice*. Chichester, NJ: Wiley.
- Vrij, A., Akehurst, L., & Morris, P. (1997). Individual differences in hand movements during deception. *Journal of Nonverbal Behavior*, *21*(2), 87-102.
- Vrij, A., & Graham, S. (1997). Individual differences between liars and the ability to detect lies. *Expert Evidence*, *5*(4), 144-148.
- Vrij, A., & Winkel, F. W. (1993). Objective and subjective indicators of deception. *Issues in Criminological and Legal Psychology*, *20*, 51-57.
- Zebrowitz, L. A. (1997). *Reading faces. Window to the soul?* Boulder, Colorado: Westview Press.
- Zebrowitz, L. A., & Montepare, J. M. (1992). Impressions of babyfaced individuals across the life span. *Developmental Psychology*, *28*(6), 1143-1152.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, 22(12), 1258-1269.

Zuckerman, M., DeFrank, R. S., Hall, J. A., Larrace, D. T., & Rosenthal, R. (1979). Facial and vocal cues of deception and honesty. *Journal of Experimental Social Psychology*, 15, 378-396.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, 14, 1-59.

Zuckerman, M., & Driver, R. (1985). Telling lies: Verbal and nonverbal correlates of deception. In A. W. Siegman, & S. Feldstein (Eds.), *Multichannel integrations of nonverbal behaviors* (pp. 129-147). Hillsdale, NJ: Lawrence Erlbaum.

Zuckerman, M., Driver, R., & Guadagno, N. S. (1985). Effects of segmentation patterns on the perception of deception. *Journal of Nonverbal Behavior*, 9(3), 160-168.

### About the Authors

**Jaume Masip** (Ph.D. 2002, University of Salamanca, Spain) is in the Department of Social Psychology and Anthropology of the University of Salamanca since 1996, where he teaches psychology of crime to criminology students and is one of the supervisors of the external practical training of psychology undergraduates. His research interests are nonverbal behavior, verbal and nonverbal detection of deception, and the role of facial appearance on social judgments. He is the creator and webmaster of the Nonverbal Behavior / Nonverbal Communication. Links site:

<http://www3.usal.es/~nonverbal>

**Eugenio Garrido** (Ph.D. 1975, University of Salamanca, Spain) is full professor of Social Psychology and the chair of the Department of Social Psychology and Anthropology of the University of Salamanca, where he teaches legal psychology. His research interests are social cognitive theory and several topics related to psychology and law, such as eyewitness testimony, police research, victimology, the detection of deception and the causes of criminal behavior. He has authored many journal articles and book chapters on the topics he is interested in.

**Carmen Herrero** (Ph.D. 1993, University of Salamanca, Spain) is a social psychology lecturer in the Department of Social Psychology and Anthropology of the University of Salamanca, where she teaches social psychology and victimology. Her research interests involve victimology, judges' attributions concerning their sentencing in sexual abuse cases, the detection of deceit and social-cognitive explanations of deviant behavior.

APPENDIX 1: Partial association tests, parameter estimates ( $\beta$ ) and z statistics corresponding to the significant effects in the Occupation X Value Of Truth X Moment X Accuracy hierarchical loglinear analysis for S1 in Experiment 1.

	Partial association chi-square	p		$\beta$		z				
<i>First-order effect</i>										
Moment	4.09	.043		Moments 1 and 2 .04	Moment 3 -.04	Moments 1 and 2 0.55	Moment 3 -0.55			
<i>Second-order effects</i>										
Moment X Hit/miss	12.25	.001	Mom. 1 & 2 Moment 3	Miss -.47 .47	Hit .47 -.47	Mom. 1 & 2 Moment 3	Miss -2.80 2.80	Hit 2.80 -2.80		
Value of truth X Hit/miss	49.63	.000	Deceptive Truthful	Miss -.47 .47	Hit .47 -.47	Deceptive Truthful	Miss -6.20 6.20	Hit 6.20 -6.20		
Occupation X Moment	3.24	.072	Officers Students	Moments 1 and 2 -.15 .15	Moment 3 .15 -.15	Officers Students	Moments 1 and 2 -1.92 1.92	Moment 3 1.92 -1.92		
<i>Third-order effects</i>										
Value of truth X Moment X Hit/miss	11.40	.001	Deceptive Truthful	Moments 1 and 2 Moment 3	Miss -.24 .24	Hit .24 -.24	Deceptive Truthful	Moments 1 and 2 Moment 3	Miss -3.12 3.12	Hit 3.12 -3.12
Occupation X Value of truth X Hit/miss	8.13	.004	Officers Students	Deceptive Truthful	Miss -.20 .20	Hit .20 -.20	Officers Students	Deceptive Truthful	Miss -2.66 2.66	Hit 2.66 -2.66

Partial association tests, parameter estimates ( $\beta$ ) and z statistics corresponding to the significant effects in the Occupation X Value Of Truth X Moment X Accuracy hierarchical loglinear analysis for S2 in Experiment 1.

	Partial association chi-square	p		$\beta$		z				
<i>First-order effects</i>										
Hit/miss	16.11	.000		Miss -.19	Hit .19	Miss -2.79	Hit 2.79			
Moment	9.88	.002		Moments 1 and 2 .13	Moment 3 -.13	Moments 1 and 2 1.96	Moment 3 -1.96			
<i>Second-order effects</i>										
Value of truth X Hit/miss	3.87	.049	Deceptive Truthful	Miss -.10 .10	Hit .10 -.10	Deceptive Truthful	Miss -1.47 1.47	Hit 1.47 -1.47		
Value of truth X Moment	12.18	.001	Deceptive Truthful	Moments 1 and 2 .19 -.19	Moment 3 -.19 .19	Deceptive Truthful	Moments 1 and 2 2.82 -2.82	Moment 3 -2.82 2.82		
<i>Third-order effects</i>										
Value of truth X Moment X Hit/miss	12.65	.000	Deceptive Truthful	Moments 1 and 2 Moment 3	Miss -.24 .24	Hit .24 -.24	Deceptive Truthful	Moments 1 and 2 Moment 3	Miss -3.58 3.58	Hit 3.58 -3.58
Occupation X Value of truth X Hit/miss	7.04	.008	Officers Students	Deceptive Truthful	Miss -.17 .17	Hit .17 -.17	Officers Students	Deceptive Truthful	Miss -2.44 2.44	Hit 2.44 -2.44



APPENDIX 2: Partial association tests, parameter estimates ( $\beta$ ) and z statistics corresponding to the significant effects in the Judgment X Face X Moment hierarchical loglinear analysis for S1Truthful in Experiment 2.

	Partial association chi-square	p	$\beta$			z		
<i>First-order effect</i>								
Judgment (of truthfulness / of deceptiveness)	40.38	.000	J. of Truthfulness		J. of Deceptiveness	J. of Truthfulness		J. of Deceptiveness
			-			-		
			.44		.44	-5.58		5.58
<i>Second-order effects</i>								
Judgment X Moment	15.45	.000	Moments 1 and 2		Moment 3	Moments 1 and 2		Moment 3
			J.Truthfulness			J.Truthfulness		
			J.Deceptivnss			J.Deceptivnss		
			-			-		
			.29		.29	-3.70		3.70
			.29		-.29	3.70		-3.70
Moment X Face	8.12	.017	Moments 1 & 2			Moments 1 & 2		
			Moment 3			Moment 3		
			Face A	Face B	Face C	Face A	Face B	Face C
			.23	-.18	-.05	2.14	-1.61	-0.47
			-.23	.18	.05	-2.14	1.61	0.47

Partial association tests, parameter estimates ( $\beta$ ) and z statistics corresponding to the significant effects in the Judgment X Face X Moment hierarchical loglinear analysis for S1 *Deceptive* in Experiment 2.

	Partial association chi-square	p		$\beta$	z																				
<i>First-order effect</i>																									
Judgment (of truthfulness / of deceptiveness)	53.52	.000		<table border="1"> <tr> <td>J. of Truthfulness</td> <td>J. of Deceptiveness</td> </tr> <tr> <td>-.55</td> <td>.55</td> </tr> </table>	J. of Truthfulness	J. of Deceptiveness	-.55	.55	<table border="1"> <tr> <td>J. of Truthfulness</td> <td>J. of Deceptiveness</td> </tr> <tr> <td>-6.46</td> <td>6.46</td> </tr> </table>	J. of Truthfulness	J. of Deceptiveness	-6.46	6.46												
J. of Truthfulness	J. of Deceptiveness																								
-.55	.55																								
J. of Truthfulness	J. of Deceptiveness																								
-6.46	6.46																								
<i>Second-order effects</i>																									
Judgment X Moment	11.55	.001	J.Truthfulness J.Deceptivnss	<table border="1"> <tr> <td>Moments 1 and 2</td> <td>Moment 3</td> </tr> <tr> <td>-.28</td> <td>.28</td> </tr> <tr> <td>.28</td> <td>-.28</td> </tr> </table>	Moments 1 and 2	Moment 3	-.28	.28	.28	-.28	<table border="1"> <tr> <td>J.Truthfulness</td> <td>J.Deceptivnss</td> </tr> <tr> <td>-3.27</td> <td>3.27</td> </tr> <tr> <td>3.27</td> <td>-3.27</td> </tr> </table>	J.Truthfulness	J.Deceptivnss	-3.27	3.27	3.27	-3.27								
Moments 1 and 2	Moment 3																								
-.28	.28																								
.28	-.28																								
J.Truthfulness	J.Deceptivnss																								
-3.27	3.27																								
3.27	-3.27																								
Moment X Face	18.70	.000	Moments 1 & 2 Moment 3	<table border="1"> <tr> <td>Face A</td> <td>Face B</td> <td>Face C</td> </tr> <tr> <td>.44</td> <td>-.29</td> <td>-.15</td> </tr> <tr> <td>-.44</td> <td>.29</td> <td>.15</td> </tr> </table>	Face A	Face B	Face C	.44	-.29	-.15	-.44	.29	.15	<table border="1"> <tr> <td>Moments 1 &amp; 2</td> <td>Moment 3</td> </tr> <tr> <td>Face A</td> <td>Face B</td> <td>Face C</td> </tr> <tr> <td>3.64</td> <td>-2.53</td> <td>-1.16</td> </tr> <tr> <td>-3.64</td> <td>2.63</td> <td>1.16</td> </tr> </table>	Moments 1 & 2	Moment 3	Face A	Face B	Face C	3.64	-2.53	-1.16	-3.64	2.63	1.16
Face A	Face B	Face C																							
.44	-.29	-.15																							
-.44	.29	.15																							
Moments 1 & 2	Moment 3																								
Face A	Face B	Face C																							
3.64	-2.53	-1.16																							
-3.64	2.63	1.16																							

APPENDIX 3: Partial association tests, parameter estimates ( $\beta$ ) and z statistics corresponding to the significant effects in the Moment X Value of Truth X Hit/Miss hierarchical loglinear analysis for the *first* presentation in Experiment 2.

	Partial association chi-square	<i>p</i>		$\beta$		<i>z</i>				
<i>Second-order effect</i>										
Value of truth X Hit/miss	51.29	.000	Truthful Statemnt.	Miss	Hit	Truthful Statemnt.				
			Deceptive Sttmnt	.55	-.55	Deceptive Sttmnt				
						Miss				
						Hit				
						6 .70				
						-6.70				
						-6.70				
						6 .70				
<i>Third-order effect</i>										
Moment X Value of truth X Hit/miss	17.54	.000	Momnts 1 and 2	Truthful	Miss	Hit	Momnts 1 and 2	Truthful	Miss	Hit
				Deceptive	.32	-.32		Deceptive	3.95	-3.95
			Moment 3	Truthful	-.32	.32	Moment 3	Truthful	-3.95	3.95
				Deceptive	-.32	.32		Deceptive	-3.95	3.95
				.32	-.32			3.95	-3.95	

Partial association tests, parameter estimates ( $\beta$ ) and z statistics corresponding to the significant effects in the Moment X Value Of Truth X Hit/Miss hierarchical loglinear analysis for the *second* presentation in Experiment 2.

	Partial association chi-square	<i>p</i>		$\beta$		<i>z</i>				
<i>Second-order effect</i>										
Value of truth X Hit/miss	36.57	.000	Truthful Statemnt. Deceptive Sttmnt	Miss	Hit	Truthful Statemnt. Deceptive Sttmnt	Miss	Hit		
				.42	-.42		5 .60	-5.60		
				-.42	.42		-5.60	5 .60		
<i>Third-order effect</i>										
Moment X Value of truth X Hit/miss	11.83	.001	Momnts 1 and 2	Truthful	Miss	Hit	Momnts 1 and 2	Truthful	Miss	Hit
				Deceptive	.25	-.25		3.35	-3.35	
			Moment 3	Truthful	-.25	.25	Moment 3	Truthful	-3.35	3.35
				Deceptive	.25	-.25		Deceptive	3.35	-3.35

**Article submitted for publication: 9 September 2001**

**Revision submitted: 11 November 2002**

**Article accepted for publication: 12 April 2003**

**Article published 29 July 200**