Social desirability and controllability in computerized and paper-and-pencil personality questionnaires

Shaul Fox*, David Schwartz

Department of Psychology, Bar-Ilan University, Ramat-Gan 52900, Israel

Abstract

The main objective of this study was to examine Steel’s [Advances in Experimental Social Psychology 21 (1988) 261] self-affirmation theory, by checking the effects of different levels of manipulated control on participants’ performance on personality tests. Specifically, we assessed the impact of participants’ level of control (i.e. prior familiarity of questionnaire’s items, choosing of test content) and anonymous or identified data collection on two different forms of socially desirable responses: impression management and self-deception [Journal of Personality and Social Psychology 46 (1984) 598]. Second, we compared these effects on paper-and-pencil versus computerized testing conditions. In Study 1 (N=91) we showed that perceived control can be manipulated in computerized tests and that higher level of control is related to more positive attitudes towards the test and to lower levels of anxiety experienced during test taking. In Study 2 (N=200) we found a significant positive relationship between manipulated control and impression management. The hypothesis claiming a significant relationship between controllability and self-deception has not been corroborated. In addition, no differences were found between the paper-and-pencil and the computerized mode of administration on measures of perceived control, trust, candor, and social desirability. © 2002 Published by Elsevier Science Ltd.

Keywords: Perceived control; Social desirability; Computerized testing; Personality tests

1. Introduction

The prevalent use of self-report measures of personality in psychological research has raised doubts regarding the accuracy of these measures in assessing personality
attributes. Self-report measures may be so heavily influenced by response sets that they may not provide a valid picture of the respondents’ psychological condition. One of the most common of these response sets that has raised concern among personality researchers is social desirability. This tendency to create a positive impression is manifested especially when respondents are highly motivated to achieve some goal, such as when applying for a job. In these situations, people may be strongly motivated to make a positive impression, and may alter their responses to deny negative elements of their personality and to claim favorable qualities.

There is still controversy in the literature on whether socially desirable responses to self-report personality tests are only contaminants or whether they may disclose substantive information regarding the respondents’ personality. On the one hand, they could merely be reflecting respondents’ deceptive accounts of themselves that may adversely affect the validity of the tests (Levin & Montag, 1987; Zerbe & Paulhus, 1987). However, these responses could possibly represent certain traits related to other scores of personality measures as well as to external criteria, and thus may serve as predictors in their own right (McCrae & Costa, 1983; Nicholson & Hogan, 1990; Ones, Viswesvaran, & Reiss, 1996).

A significant contribution to this debate was introduced by Paulhus (1984). He distinguished between two categories of socially desirable responses: self-deception and impression management. Self-deception is an unconscious drive to perceive oneself in a favorable light, to reduce inconsistency, or to protect one’s ego. Respondents, in this case, do not try to fabricate their responses, rather they answer the questions honestly, as they believe. However, this belief is a concealment derived from their positively biased view of themselves (or from lack of self-insight). Self-deception is mainly manifested in the tendency of respondents to deny their negative qualities.

In contrast, the impression management component of social desirability is a conscious attempt to impress others, and to present a false facade in order to pursue a personally meaningful goal. The respondents purposely try to present an honorable front by claiming positive qualities they know they do not possess, or by denying negative qualities they know they have. Paulhus (1984) developed a social desirability scale (Balanced Inventory of Desirable Responding, BIDR) that measures these two aspects.

Respondents to self-report personality tests are required to expose their weaknesses to a stranger, who could possibly use the information to make critical decisions about their future. Socially desirable responses can be viewed, therefore, as a defensive means against this threat. Making a good impression on others is a common behavioral tactic in many interpersonal interactions, and it is assumed to be used more frequently in threatening or stressful social interaction situations, where the impression made may determine the likelihood of achieving personally significant goals.

Aside from the motivation to impress others, when responding to a self-report personality questionnaire one is directly exposed to his or her own faults. This is an unpleasant and threatening situation that may arouse ego-protective processes. Therefore, by denying one’s faults and at the same time stressing positive
self-aspects one is able to ward off ego-threatening information and maintain psychological equanimity.

One pertinent question regarding social desirability is whether the drive to respond in a socially desirable manner always exists? Are there conditions in which the person can take other steps against the threat instead of using socially desirable responses?

1.1. Self-affirmation theory

Self-affirmation theory (Steele, 1988) indicates that there are situations in which people do not protect the ego by dealing directly with the threatening cause. Rather, when some aspect of the self is threatened, the person may restore self-integrity by seeking other positive aspects of the self. Being aware of other aspects of the self than the one under threat, and choosing to behave in a manner that may strengthen positively perceived aspects of the self, may sustain self-esteem. This defensive strategy may enable the subject to cope adequately with the existing threat and to reaffirm his or her general sense of worth. A central aspect of the theory is that self preservation is conducted by referral to a broader self concept, and not necessarily focusing on those aspects geared to solve the specific self-concept threats. This, according to Steele (1988), illustrates the fluidity of the self-affirmation process, wherein one resource of the self may compensate for other negative components and thus maintain a sense of self-worth.

When applying self-affirmation theory to personality test-testing, we hypothesize that by inducing a general sense of self worth we may buffer participants' inclination to utilize defensive impression management strategies, since an awareness of their strengths and virtues may weaken their need to employ ego-protective maneuvers. It seems that one possible way of activating global feelings of self worth is by inducing a personal sense of control. It is hypothesized that inducing participants with a sense of control over the testing situation will create positive feelings that will enhance their self-integrity. Following this sense of controllability participants may perceive their general "totalitarian ego" (Greenwald, 1980) as positive and strong, reducing the need to affirm their worth and goodness by other means.

1.2. Perceived control

There is ample evidence in the psychological literature on the favorable psychological implications of having a sense of control (Friedman & Lackey, 1991), and the negative implications of loss of control (Abramson & Alloy, 1980; Averill, 1973; Greenberger, Strasser, Cummings, & Dunham, 1989; Langer & Rodin, 1976; Seligman, 1975). People are generally motivated to produce behavior-outcome contingencies and they seek to exert control over the environment (White, 1959). Control is a major component of the individual's emotional well-being (Greenberger & Strasser, 1991); it helps to adjust to major life events and to cope with stress and pain; it leads to better health outcomes; it helps make behavioral changes; and it improves performance (Thompson & Spacapan, 1991). Thus, we assume that
augmenting the personal sense of control may serve as a resource that may increase the overall sense of self-worth and consequently, may serve as a compensatory element in the face of threat to self-esteem. Our main hypothesis, therefore, is that as people affirm an image of themselves as competent and able to control important outcomes (i.e. they can control the outcomes of the personality test), they will cope with the threat of self-exposure without resorting to socially desirable responses.

The question is whether this effect of heightened perceived control will affect both the impression management and the self-deception components of socially desirable behavior equally. Since self-deception is the mechanism people employ to protect their self-esteem, it is reasonable to assume that a heightened sense of control will serve as a self-affirming cognition regarding their self worth, that will diminish the need to protect the ego with a self-deception strategy.

However, since impression management is a conscious behavior that may be instrumental in achieving external goals, a greater sense of control may actually encourage participants to further pursue their goals, and may, therefore, increase their motivation to impress others. Consequently, we hypothesize that the use of impression management strategies will increase as feelings of control increase.

1.3. Anonymity

One of the methods suggested to curb the tendency to respond in a socially desirable manner is by soliciting participants' responses to personality tests under conditions of anonymity (Holden, Magruder, Stein, Sitarenios, & Sheldon, 1999). However, anonymity seems to affect only impression management. When respondents are convinced that their identity cannot be determined from the questionnaire, they will not expect their responses to help them in achieving their personal goals. On the other hand, since the self-deception component of social desirability is a mechanism used to protect the self, it should not be affected by anonymity, and the need to deny negative aspects of the self will remain. A few studies tested these assumptions, however, their results are equivocal. Although Paulhus (1984) found this anonymity effect, both Booth-Kewley, Edwards, and Rosenfeld (1992) and Lautenschlager and Flaherty (1990) found that anonymity decreased both impression management and self deception scores (however, in Lautenschlager and Flaherty's study, scoring of a subset of self-deception items that better represent anonymous and identified conditions, yielded no significant differences between anonymous and identified conditions, as initially expected regarding the effects of anonymity on self-deception).

We expected to find an interaction effect of assigned control and anonymity on the two social desirability scores. When participants are identified, those with a high level of control will have a higher score on impression management scale in comparison to those with low level of control. On the other hand, in the anonymous condition no difference is expected between high and low levels of control. In comparison to impression management scores, self-deception scores are expected to be lower when the respondents have a higher level of control, with no expected differences between anonymous and identified conditions.
1.4. Computerized testing

Computerized testing is the most advanced recent development in testing technology. The use of computers as a means of administering psychological tests has become increasingly prevalent due to its' advantages over paper and pencil testing. Computerized testing saves time, it eliminates coding and scoring errors, it enables more standardization, and it allows collecting information about many variables that are not accessible in the conventional non-computerized mode, such as time latency, order of responses, number of corrections and more. One of the greatest advantages of computer testing is that it enables the simple manipulation of variables and the examination of subjects' responses in carefully controlled experimental conditions. Thus, it is possible to provide instant feedback to performance, the duration of the exposure of each item can be controlled, and participants freedom to go backward and forward, or to change their answers can be constrained. These characteristics facilitate the control over selected variables and enable a precise examination of their effects on responses and scores. In the present study we took advantage of computerized testing to manipulate different levels of controllability when responding to personality questionnaires.

The question of whether participants tend to respond in a socially desirable manner in paper-and-pencil tests more often than in computerized tests is still not resolved. Although several studies showed that participants tend to be less defensive and respond more openly on computerized personality tests (Evan & Miller, 1969; Koson, Kitchen, Kochen, & Stodolsky, 1970; Lucas, Mallins, Luna, & McInrrey, 1977; Martin & Nagao, 1989), other studies did not find any real differences between computerized and paper and pencil tests in regard to social desirability responses (Greist, 1975; Skinner & Allen, 1983). In a recent meta-analysis conducted by Dwight and Feigelson (2000) a small significant effect was found for impression management, where scores were lower when responses were given by means of computer. However, this effect diminished over time and in recent studies virtually no differences were found. As expected no differences between computer and paper-and-pencil administration were found for self-deception factor. Similar results are reported in another meta-analysis conducted by Richman, Kiesler, Weisband, and Drasgow (1999), although Richman et al. found that when participants answered the questionnaires alone and when they could check and correct previous answers, computer mode of administration yielded lower social desirability scores. Since we assume that perceived control influences socially desirable responding, we will examine whether perceived control differentially affects social desirability scores on paper-and-pencil personality tests versus computerized modes of administration.

To summarize, the research's objectives are twofold: the research's main objective is to examine the influence of different levels of perceived control and identified/anonymous conditions on the impression management and self-deception components of social desirability; the second objective of the study is to compare computerized testing with paper-and-pencil test administration.

Two studies were conducted. The first study was a preliminary exploratory stage designed to ensure that sense of control on testing behavior could be manipulated.
In the second experiment we tested our hypotheses regarding the relationship between perceived control, anonymity, and socially desirable behavior and we also compared paper-and-pencil and computerized modes of administration on several criteria.

2. Study 1

It appears that previous studies have not examined the effect of level of control testing behaviour. Therefore, our first study was conducted to examine whether different levels of perceived control can be manipulated during test taking, and whether the effects of felt control will be reflected in cognitive, emotional and behavioral aspects related to test taking. We examined the extent of participants' trust in the test to tap the cognitive aspect, the level of their anxiety while test taking to tap the emotional aspect, and participants' readiness to provide candid answers to personality items to tap the behavioral aspect. We assumed that the positive implications of having a higher level of control would be reflected in expressing a higher degree of trust in the test, experiencing lower levels of anxiety during the test, and expressing a stronger readiness to cooperate and provide genuine answers to the tests' questions.

2.1. Method

2.1.1. Participants

The participants in the study were 91 men, 18–20 years of age who were candidates for a prestigious military course. Participation in the course was on a voluntary basis and dependent upon the successful completion of screening tests. After completing their standard tests they were asked to participate in the study. It was made clear that participation in the experiment was on a voluntary basis and that it would not affect the decision regarding their admittance to or rejection from the course. Most of the applicants agreed to partake in the experiment.

2.1.2. Questionnaires and tests

2.1.2.1. Tennessee Self Concept Scale (TSCS). This test was developed by Fitts (1965) and was adjusted to Israeli norms by Frankel and Arzi (1976). It consists of 100 items and answered on a five-point Likert scale. The test was transformed to a computerized version for the present study by a program developed by the TOP-C Company. The test assesses participants' feelings about themselves. It consists of a total self-concept score and three sub-scales measuring self-identity, self satisfaction, and behavior. In addition, the scale also includes a self-criticism scale. Each of the first three sub scales is measured by 30 items tapping five self-aspects: physical, moral-ethical, personal, familial and social. The self-criticism scale consists of 10 items taken from the MMPI Lie scale. In the present study we administered only the first three sub scales and removed four items that were judged unsuitable for our sample. Thus the final scale consisted of 86 items. We used this test as a stimulus to manipulate the extent of participants' sense of control.
2.1.2.2. Perceived control (PC). This questionnaire was developed by the authors for the present research. It examines respondents' feelings of personal control on their test behavior. The questionnaire was developed by constructing 14 items covering different aspects of perceived control in test settings. When selecting items, we attempted to cover most aspects of perceived control (see Langer, 1983) such as participants' ability to direct responses to items according to their objectives, their freedom to choose how to answer the items and in what order, having enough information about the test and how to respond to it, and the ability to manage the desired impression. This first version of the questionnaire was administered to 200 candidates to a prestigious military course (these subjects did not take part in the other studies) after completing the Tennessee Self Concept Test (Fitts, 1965). Outcomes of several standard psychometric tests led us to reframe seven items and to add one more item.

This second, revised version of the questionnaire, was administered to an additional sample of 150 military candidates (these subjects did not take part in other studies), after completing the Tennessee Self Concept Test (Fitts, 1965). In a factor analysis conducted on the revised scale (principal component analysis with Varimax rotation) three factors were extracted. The first factor was labeled: Perceived Control on Responses. This factor included eight items (e.g. “To what extent did you feel that you could present yourself as you wished?”; “To what extent did you feel that you had enough freedom while answering the test?”). The second factor was labeled: Ability to Influence the Outcome of the Test. This factor consisted of four items (e.g. “To what extent did you feel that you could influence the eventual impression of yourself on the basis of your responses to the questionnaire?”). The third factor was labeled: Choice. This factor included two items (e.g. “To what extent did you feel that you could choose what items to respond to and what items to skip?”). One item had low loadings on all factors, and was weakly correlated with the overall score of the questionnaire and was therefore removed. Participants answered all items on a five-point Likert-type scale from 1—a very low extent, to 5—a very high extent. Reliability measures for the three factors were quite low (Cronbach’s alphas = 0.72, 0.60, 0.52 for the three factor scores, respectively). The internal reliability for the entire questionnaire was suitable (Cronbach’s alpha = 0.73). Therefore, in the following analyses we used only the overall score for measuring perceived control.

2.1.2.3. Attitude toward the test. We developed this questionnaire in order to examine respondents’ attitude towards the test as a fair instrument to evaluate their personality, and their readiness to provide candid responses. First we constructed 17 items which covered these aspects. This first version of the questionnaire was administered to the same 200 participants who answered the first version of the perceived control questionnaire (see earlier). Half of the participants answered this questionnaire first (after completing the Tennessee Self-Concept Scale) and the other half answered it after completing the perceived control questionnaire. Following the results of this initial step, a few changes were made. We rephrased two items, removed five other items which had very low correlations with the overall score, and added two additional items.
This revised version of the questionnaire which consisted of 14 items was administered to 150 participants, the ones who took part in answering the second version of the perceived control questionnaire (counterbalancing the order of the two questionnaires). Factor analysis conducted on the responses of this sample yielded two factors. The first factor was labeled: Trust in the Test as a Just and Effective Instrument for Diagnosing Personality. This factor included eight items (e.g. "To what extent do you think the test is a fair test"? "To what extent do you think that responses to this test reliably tap respondents personality?"). The second factor was labeled: Candor. This factor included six items (e.g. "To what extent were your responses to the test candid?" "To what extent do you think that your friends will cooperate in providing candid answers to this test?"). Cronbach’s alpha for the first factor was 0.84 and for the second factor 0.85. For the overall questionnaire alpha = 0.87. Thus, the final questionnaire had 14 items, scored on a five-point scale from 1—a very low extent to 5—a very high extent. In the following analyses results will be provided separately for each factor.

2.1.2.4. State-Trait Anxiety Inventory (STAI, Spielberger, Gorsuch, & Lushene, 1983). We chose to test the level of anxiety experienced by the participants while answering the personality test with the state component of the STAI. This part of the inventory, which consists of 20 items (Hebrew version, translated by Teichman & Melnick, 1984) measures transitory anxiety response to a specific situation. In the present study we were interested in assessing the level of anxiety aroused by the personality questionnaire, and the differences in anxiety level as a function of the level of perceived control participants have over their testing behavior.

The three questionnaires (perceived control, attitude toward the test and the anxiety scale) were administered by computer. All the items were randomly mixed and were administered as one questionnaire. Each item was shown on a separate screen and the participants answered it by using the keyboard.

2.1.3. Experimental manipulations

The experimental manipulations used in the present study were improved following initial studies. The results of these studies led us to strengthen the control manipulations and to increase the differences between the experimental groups. Four experimental groups were manipulated as follows:

2.1.3.1. Fair control. This group received a computerized version of the TSCS generally equivalent to the original paper and pencil version. Each item was presented on a separate screen. The instructions were generally similar to the original paper and pencil version. However, additional instructions notified the respondents that they had to answer each item when presented on the screen, they had to answer all the items and would not be able to skip any items, or go back to view an item or change their responses.

2.1.3.2. Weakened control. In addition to the same instructions given in the fair control condition, the respondents of this group were given additional instructions
informing them that each item would be shown on the screen only for 3–9 s and they would have to respond to the item within this time limit. “You have to answer quickly, you will not be able to think too long, and you will not have enough time to change your answer.” They were further told that they would not be able to know in advance how long the item would be presented. For each item a red digital clock, which showed how many seconds passed from the time the item was shown, appeared on the screen. The location of the clock and its fonts were randomly changed along the test screens. If the allotted time expired without getting an answer, the item disappeared and on the screen appeared a warning in red letters: “You did not meet the time limit. Go back and answer the item within the time limit. Please notice that the number of times you do not meet the time limit is recorded”. After a few screens, an additional unexpected note was presented: “This questionnaire consists of items which cover several personality aspects. We calculate scores for each of these aspects”. We assumed that this unexpected and unclear information would create uncertainty and would further impair feelings of control. Finally, two times during the trial the respondents were told that the computer could not accept their answers, since they probably mistakenly chose a number larger than 5. The items reappeared on the screen and the participants were asked to be careful and choose only responses between 1 and 5 (this note was screened to all subjects, unrelated to their real responses).

2.1.3.3. The choice group. After reading the general instructions participants in this group were given information about the test. They were told that the test measures three aspects: self-identity, present behavior, and behavior in childhood (the two first measures were real, while the third was false. The original behavior factor of the TSCS was included but no information was provided regarding it). For each aspect a few sample questions were presented. For the third aspect—behavior in childhood—which was not part of the original scale, sample items were “When I was very young I wept a lot at night”, “When I was a young child I had difficulties to stop suckling”. The objective of this measure was to manipulate participants’ belief that they had control over the tests’ content. Accordingly, participants were told they could choose which aspects they would like to respond to. The measure of behavior in childhood was added, in fact, since it was assumed that most participants would choose not to answer it, and thus all participants would undergo the same test. Indeed, most of the participants chose not to respond to this scale, and the four participants who chose to answer it were not included in the sample. In order to assure participants that their choice was real, an announcement appeared on the screen for five seconds that said: “please wait until the computer prepares the test according to your choice.” In addition, participants were allowed to go back and change their responses, and there was no time limit for answering the test. To remind participants that the test was structured according to their preference, at the end of the test another note was screened that read: “The test included only questions which examined two aspects—self-identity and present behavior—as you chose. We did not include items regarding behavior in childhood which you preferred not to answer”. This procedure was assumed to lead participants to believe
that they chose the topics they would be tested on while in fact all participants completed the exact same scales.

2.1.3.4. Prior familiarity of stimuli and responses group. According to Langer (1983) familiarity of the responses and of the stimuli increases the person's illusion of control. We applied this idea in the current study by providing the respondents with a detailed explanation on the self-aspects the test examines (physical self, personal self, etc.), and gave them examples of sample items. They were urged to practice answering these items, knowing that this part is a practice stage in which their answers would not be taken into consideration. In addition, they were provided with information on answers' alternatives, and on the way their score is computed. They were assured that they would not need to memorize this information. The information about the test's aspects reappeared when the respondents started answering the test: For each item information about the aspect that the item examines also appeared. The instructions informed respondents that they could go back and change their responses, and that the test has no time limit.

2.1.4. General procedure

After completing the standard military screening tests for which the participants were summoned to, all those who volunteered to take part in the experiment were seated before a computer. They were given detailed explanations on the procedure and completed practice exercises to familiarize them with the computer (most of the subjects had prior knowledge on computers and they did not need too long practice). Before administering the tests they were told that they do not need to provide any information about themselves, since their responses are anonymous. Following this initial procedure the TSCS and the three other questionnaires—perceived control, attitudes, and STAI in their computerized form were given. As mentioned before, the items of these three questionnaires were combined and randomly mixed forming one questionnaire. Each item was presented on a separate screen. The test and the questionnaires were administered to a group of three respondents each time, and each group was randomly assigned to one of the four experimental groups (i.e. fair control, weakened control, etc.).

2.2. Results

Table 1 presents the means and standard deviations for each of the four reactions toward the personality test, according to the four experimental control manipulations. An omnibus MANOVA was significant, $F(12, 222) = 2.32, P < 0.008$ (Wilks Lambda). Follow-up one-way ANOVAs revealed significant differences between the four experimental groups for all the dependent variables: perceived control, $F(3, 87) = 6.28, P < 0.01$, Trust, $F(3, 87) = 2.82, P < 0.04$, Candor, $F(3, 87) = 3.42, P < 0.02$, and Anxiety, $F(3, 87) = 5.83, P < 0.001$.

Contrast analysis across all the dependant variables, yielded significant differences ($P < 0.05$) between the weakened control group and the choice and familiarity groups. Differences that approached significance were found between the weakened
Table 1
Means and standard deviations of perceived control, trust, candor and state anxiety for the four experimental groups

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Perceived control</th>
<th>Trust</th>
<th>Candor</th>
<th>State anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weakened control</td>
<td>25</td>
<td>44.24a (6.36)</td>
<td>24.12a (5.07)</td>
<td>23.84a (3.49)</td>
<td>32.56a (9.22)</td>
</tr>
<tr>
<td>Fair control</td>
<td>24</td>
<td>45.62a (7.79)</td>
<td>26.58ab (4.52)</td>
<td>25.25ab (3.39)</td>
<td>26.29b (7.89)</td>
</tr>
<tr>
<td>Choice</td>
<td>21</td>
<td>50.90b (6.60)</td>
<td>27.24b (4.15)</td>
<td>25.95b (2.13)</td>
<td>25.33b (7.59)</td>
</tr>
<tr>
<td>Familiarity</td>
<td>21</td>
<td>50.76b (5.09)</td>
<td>27.90b (5.23)</td>
<td>26.52b (2.69)</td>
<td>24.14b (4.33)</td>
</tr>
</tbody>
</table>

* Within columns, means with the same letter are not significantly different at $P<0.05$.

control group and the fair control group ($P<0.06$), and between the fair control group and the choice group ($P<0.08$).

Contrast analyses conducted for the perceived control scores yielded significant differences ($P<0.01$) between the weakened control group and the choice and familiarity groups—where, as expected, the weakened group also had significantly ($P<0.05$) lower perceived control scores in comparison to the choice and familiarity groups. No other significant differences were found (weakened control vs. fair control groups, and choice vs. familiarity groups).

For both the trust and candor factors, contrast analyses revealed significant differences between the weakened control manipulation and both the choice and familiarity groups ($P<0.03$). Participants with a stronger manipulated control had more favorable feelings towards the test.

The contrast analyses of the state anxiety questionnaire yielded a significantly ($P<0.01$) higher level of anxiety among participants who belonged to the weakened control group in comparison to the other three experimental groups. No other significant differences were found.

Correlations between perceived control and the other three variables across the groups yielded the following outcomes (all $r$'s are significant at the $P<0.01$ level): $r=0.51$, $r=0.42$, and $r=-0.55$ for trust, candor and anxiety, respectively. Participants who felt they had more control expressed more positive attitudes toward the questionnaire, they were more ready to respond candidly and they were much less anxious.

2.3. Discussion

The findings of Study 1 demonstrated that perceived control in testing behavior can be manipulated. The results showed that perceived control is positively associated with respondents trust in the test, to their readiness to respond candidly to the test, and negatively associated with their level of anxiety while test-taking.

Participants who could choose the topics they would be tested on, and those who were given the opportunity to familiarize themselves with the test's items and mode of response, felt that they had more control, more trust in the test, were more ready to react genuinely, and their level of anxiety was lower compared to the weakened
control group. A similar yet non-significant trend was found when comparing the choice and familiarity groups with the fair control group. No significant differences were found between the choice and the familiarity groups.

Unexpectedly, no significant differences were found between the weakened control group and the fair control group, although the means for the four dependent variables were in the expected direction (the weakened control group had lower scores than the fair control group). This may be due to the constraints imposed upon the fair control group to respond to all items and their inability to go back and correct their answers.

Following our findings from Study 1 we conducted an additional study in which a paper-and-pencil administration group was added to the four experimental groups of Study 1. The objectives of this study were to examine how different levels of control affect the impression management and self-deception aspects of social desirability, and to examine the degree of perceived control and the social desirability scores while answering paper-and-pencil scales in comparison to their computerized version.

3. Study 2

3.1. Method

3.1.1. Participants

In the present study an independent sample of 200 males, 18–20 years of age, who voluntered to a highly military prestigious course participated in the study. As in Study 1, after completing screening tests for the course, participants were asked to volunteer for the present study. It was made clear to them that the results of this study will only be used for research purposes and will not affect their chances of admittance to the military course.

3.1.2. Measures

The Perceived Control questionnaire, the attitudes toward the test, with its two factors (Trust in the test and Candor), and the State-Trait Anxiety Inventory (STAI, Spielberger et al., 1983) were administered in the same way as in Study 1. The Tennessee Self Concept Scale (Fitts, 1965), was used as in Study 1. However, we added the self-criticism scale as an additional measure of social desirability.

Three more instruments were used in the present study:

3.1.2.1. Eysenck and Eysenck's (1975) Personality Questionnaire (EPQ). This is a 90-item forced-choice measure which assesses four constructs: Psychoticism (P), Neuroticism (N), Extraversion (E), and Lie (L). We used a Hebrew version of the EPQ translated by Montag (1985). The English and the Hebrew versions of the test were found to be similar in their psychometric properties and their relationships with other personality inventories (Levin & Montag, 1987; Montag & Comrey, 1982). The test was administered in its original pencil- and-paper form. The test was
used in the present study to verify that the different experimental groups were similar in their social desirability (Lie scale) and personality characteristics.

3.1.2.2. Balanced Inventory of Desirable Responding (BIDR, Paulhus, 1984). This questionnaire measures the two aspects of social desirability: impression management and self-deception. We used version 6 of the questionnaire which consists of 40 items, 20 items for each factor. Paulhus (1994) found that the two scales are quite independent (the correlations between the scales ranged between 0.20 and 0.30). Scores on the impression management scale are affected by situational demands (public versus anonymous administration; Paulhus, 1984; Paulhus & John, 1998). Self-deception, on the other hand, is more stable and represents an unconscious motivation for self enhancement. The questionnaire was translated to Hebrew by the authors. Three senior psychologists checked the translations and following their comments the translation was improved. Two items (one of each scale) were omitted because of their unsuitability to our sample. The revised Hebrew version was administered to 275 subjects (an independent sample of candidates to the prestigious military course that did not take part in any of our present studies). Factor analysis (principal component analysis, with Varimax rotation) generally confirmed the factor structure reported by Paulhus. Cronbach's Alphas yielded alpha = 0.84 for impression management, and alpha = 0.71 for self-deception. The correlation between impression management and self-deception was $r = 0.59$, much higher than the results Paulhus (1994) reported, yet similar to what Ones et al. (1996) found.

The questionnaire was computerized and its items were randomly incorporated within the items of the TSCS. The order of the items of the two original questionnaires was maintained.

3.1.2.3. The Marlowe–Crowne Social Desirability Scale (MC, Crowne & Marlowe, 1964). This questionnaire consists of 33 items. It is considered the most common instrument for measuring social desirability (Furnham, 1986). The questionnaire is widely used for examining the influence of various variables on social desirability (Furnham, 1986; Martin & Nagao, 1989). Paulhus (1984) found that it correlated both with self-deception and impression management. The questionnaire was administered in a computerized version, where each item was shown separately on a different screen.

3.1.3. Procedure

As in the first study, after the subjects completed the standard tests used for screening to the military course, they were asked to volunteer to take part in this research. Almost all participants volunteered. First, they completed the EPQ. Then, they were randomly divided into two groups: anonymous responding, and identified responding. The participants of the first group were instructed to answer the questionnaires without typing their names on the computer. Participants from the other, identified group, had to type their names and were aware that their responses were not anonymous. Then, the subjects were given the TSCS and the BIDR in five different control manipulations: Fair control, weakened control, choice, prior familiarity of stimuli and responses, and paper-and-pencil administration. The
assignment to the different groups was random. The manipulations used on the first four groups were the same as in Study 1. These groups completed the computerized version of the questionnaire. The fifth group completed the same questionnaires in a paper-and-pencil version. After completing the TSCS and the BIDR all participants responded to the questionnaires that examined perceived control, attitude towards the test, and anxiety. The items of these questionnaires were mixed and presented as one questionnaire. Finally they answered the MC Scale. Again, the fifth group answered the original paper-and-pencil version while the other groups answered the test in its' computerized version. Similar control manipulations to those of TSCS and BIDR were employed while answering the MC Scale. However, four manipulations were employed instead of five: In the fair control group and in the weakened control group the manipulations were the same as in the TSCS and the BIDR. The choice group and prior familiarity groups were given the same instructions as the familiarity group was given earlier. Namely, in addition to the standard instructions, a separate screen informed the subjects that the test examined social attitudes, beliefs and behaviors. They were given explanation regarding the scale and they practiced answering sample items. In addition, they were allowed to go back and change their responses and there were no time limits (note that the same instructions were given to the choice group and the prior familiarity group only for the MC Scale).

To summarize, the design of the experiment was a 5 (experimental manipulations—fair control, weakened control, choice, prior familiarity, and paper and pencil)×2 (anonymous and identified) factorial design. The dependent variables were the impression management and self-deception scores (from the BIDR), scores on the self criticism scale of the TSCS, and scores on Marlowe–Crowne Social Desirability Scale.

3.2. Results

3.2.1. Similarity of the experimental groups

In order to insure the random assignment of the participants into the experimental groups, we first conducted an ANOVA to examine the difference between the experimental groups in regard to the Lie scale scores of the EPQ. No significant differences were found between the anonymous responding group and the identified group as well as between the five experimental groups. We also found no interaction effect between anonymity and control.

Table 2 presents the means and standard deviations of the Lie scale scores across all groups. Examinations of the scores of other EPQ factors (E, N, and P) also yielded no significant differences between the groups and no significant interaction effects. It can be concluded therefore, that the groups are very similar in their basic personality characteristics, and that no differences exist in their tendency to respond in a socially desirable way.

3.2.2. Manipulation checks

To insure that our manipulation of control was effective in this study we conducted a manipulation check to reexamine it. We were interested in exploring the
Table 2  
Means and standard deviations of the experimental groups in Lie scale scores of the EPQ (N=20 in each group)

<table>
<thead>
<tr>
<th>Anonymity/controllability</th>
<th>Anonymous responses</th>
<th>Identified responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weakened control</td>
<td>11.00 (4.37)</td>
<td>11.60 (4.62)</td>
</tr>
<tr>
<td>Fair control</td>
<td>11.85 (4.69)</td>
<td>11.25 (3.55)</td>
</tr>
<tr>
<td>Control by choice</td>
<td>11.00 (2.55)</td>
<td>10.57 (3.76)</td>
</tr>
<tr>
<td>Prior familiarity</td>
<td>12.20 (4.21)</td>
<td>10.95 (3.83)</td>
</tr>
<tr>
<td>Paper and pencil</td>
<td>11.90 (2.93)</td>
<td>12.35 (3.68)</td>
</tr>
</tbody>
</table>

The effect of the anonymity variable on perceived control, trust, candor and state anxiety as well, and to examine the scores of the participants assigned to the paper-and-pencil group on these variables in comparison to other experimental groups. Table 3 reports the means and standard deviations of these dependent variables across all groups. The data were analyzed by a two-way MANOVA which yielded a control main effect, $F(16, 572) = 3.91$ (Wilks), $P<0.001$, and an interaction effect between control and anonymity, $F(16, 572) = 2.28$ (Wilks), $P<0.01$. The anonymity main effect was not significant, $F(4, 187) = 0.16$ (Wilks), $P>0.10$. A series of $5 \times 2$ ANOVAs conducted separately for each of the dependent variables generally replicated the outcomes obtained in Study 1. The experimental groups were found to be significantly different on all four variables: Perceived control, $F(4, 190) = 9.51$, $P<0.001$, Trust, $F(4, 190) = 6.16$, $P<0.001$, Candor, $F(4, 190) = 3.91$, $P<0.001$, and Anxiety, $F(4, 190) = 6.58$, $P<0.001$. A significant interaction effect between control and anonymity variables was found only for the trust variable, $F(4, 190) = 2.57$, $P<0.05$. Comparison analyses revealed that familiarity groups expressed stronger trust when responding was anonymous in comparison to the identified instructions. In the fair control group trust was stronger in identified versus anonymous responding. In the other groups no differences were found between anonymity and nomination groups. The highest level of trust in the test, in comparison to all other conditions, appeared among members belonging to the familiarity group.

For the perceived control variable, follow up comparisons yielded significant differences ($P<0.01$) between weakened control and all other groups (perceived control for the weakened control group was lower). In addition, the familiarity group was found to be significantly ($P<0.01$) higher on perceived control than the fair control group, the choice group and the paper-and-pencil group. Generally, perceived control was lowest among participants belonging to the weakened control group. The fair control and paper-and-pencil groups felt a moderate sense of control, and the strongest level of control was perceived among members of the familiarity group. In contrast to expectation, the manipulation of choice did not significantly affect the extent of perceived control, although the perceived control of those who answered the test anonymously, was quite high.

Similar findings were revealed in regard to the trust variable. The weakened control group had the lowest level of trust in comparison to all other groups, and the
familiarity group expressed higher trust in comparison to the fair control group, the choice group, and the paper-and-pencil group. It is interesting to note that no difference was found between paper-and-pencil administration and fair control computerized administration for the trust variable.

Weaker differences were found in the comparison analyses of the candor variable. Although the weakened control group obtained the lowest score, it was significantly different only in comparison to the familiarity and the paper-and-pencil groups. In addition, familiarity group participants expressed their readiness to provide candid responses more strongly than members belonging to the fair control group and the choice group. Again, no differences were found between the paper-and-pencil group and the computerized fair control group for the candor variable.

Comparison analyses conducted on the anxiety variable revealed that the weakened control group and the paper-and-pencil group were more anxious than all other groups, yet, the weakened control group received significantly higher anxiety scores than the paper-and-pencil group.

Finally, correlations between perceived control scores and the other three variables for the entire sample yielded the following results: Perceived control with Trust, \( r = 0.54 \) \((P<0.01)\), with Candor, \( r = 0.24 \) \((P<0.01)\) and with Anxiety, \( r = -0.37 \) \((P < 0.01)\).

Generally, these findings replicated the general findings of Study 1. It seems that our control manipulation was successful in creating differences among the groups in their perceived control and in other variables as well.

<table>
<thead>
<tr>
<th>Check variables/ experimental groups</th>
<th>Anonymity</th>
<th>Perceived control</th>
<th>Trust</th>
<th>Candor</th>
<th>State anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( M )</td>
<td>SD</td>
<td>( M )</td>
<td>SD</td>
</tr>
<tr>
<td>Weakened control</td>
<td>Anonymous</td>
<td>44.15</td>
<td>7.82</td>
<td>24.10</td>
<td>5.72</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>46.15</td>
<td>4.95</td>
<td>23.75</td>
<td>6.05</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>45.15a</td>
<td>6.54</td>
<td>23.92a</td>
<td>5.81</td>
</tr>
<tr>
<td>Fair control</td>
<td>Anonymous</td>
<td>47.50</td>
<td>5.74</td>
<td>25.95</td>
<td>4.62</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>48.05</td>
<td>7.04</td>
<td>28.60</td>
<td>5.21</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>47.77c</td>
<td>6.35</td>
<td>27.27b</td>
<td>5.04</td>
</tr>
<tr>
<td>Choice group</td>
<td>Anonymous</td>
<td>51.25</td>
<td>5.22</td>
<td>26.55</td>
<td>3.63</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>47.75</td>
<td>4.36</td>
<td>26.65</td>
<td>4.17</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>49.50c</td>
<td>5.06</td>
<td>26.60b</td>
<td>3.86</td>
</tr>
<tr>
<td>Prior familiarity</td>
<td>Anonymous</td>
<td>53.70</td>
<td>6.90</td>
<td>31.15</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>52.40</td>
<td>5.88</td>
<td>27.10</td>
<td>4.51</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>53.05b</td>
<td>6.36</td>
<td>29.12c</td>
<td>4.97</td>
</tr>
<tr>
<td>Paper and pencil</td>
<td>Anonymous</td>
<td>48.30</td>
<td>5.65</td>
<td>26.55</td>
<td>3.74</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>48.55</td>
<td>4.33</td>
<td>26.80</td>
<td>4.58</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>48.42c</td>
<td>4.97</td>
<td>26.67b</td>
<td>4.13</td>
</tr>
</tbody>
</table>

\[ a \] Within columns, only for all participants, means with the same letter are not significantly different at \( P < 0.05 \).
3.2.3. Control and social desirability

Table 4 presents the means and standard deviations of social desirability measures across the experimental groups. Although no differences were found between the experimental groups in terms of all personality factors measured by the EPQ, we decided to take extra caution and examine the effect of the control manipulation on the social desirability measures after controlling for the Lie scale scores and the Neuroticism scores. A MANCOVA analysis (controlling for these two variables) yielded only a main effect for the perceived control variable, $F(16, 562) = 2.87$ (Wilks), $P < 0.001$. No significant effects were found for the anonymity variable, $F(4, 184) = 1.43$, (Wilks), $P > 0.10$, and for the interaction between control and anonymity, $F(16, 562) = 1.37$ (Wilks), $P > 0.10$. Thus our hypothesis about the expected interaction effect of control and anonymity was not corroborated. The control variable alone affected social desirability scores.

In a following series of ANCOVAs, a main effect of the control variable was found only for the impression management component, $F(4, 187) = 5.32$, $P < 0.001$. Comparison analyses (Scheffe) yielded significantly higher impression management scores for the familiarity group and paper-and-pencil group in comparison to the weakened control group. No other comparisons were found to be significant. The paper-and-pencil group had the highest impression management scores, while the weakened control group received the lowest scores. ANCOVAs for the self-deception scores and the Marlowe Crowne scores yielded only marginally significant effects ($F(4, 187) = 1.94$, $P < 0.10$ and $F(4, 187) = 2.24$, $P < 0.06$, respectively). In these analyses, the familiarity and paper-and-pencil groups had higher scores than

Table 4
Means and standard deviations of social desirability measures across all experimental groups$^a$

<table>
<thead>
<tr>
<th>Experimental groups/ Social desirability measures</th>
<th>Anonymity</th>
<th>Impression management</th>
<th>Self-deception</th>
<th>Self-criticism</th>
<th>Marlowe-Crowne</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Weakened control</td>
<td>Anonymous</td>
<td>61.80</td>
<td>7.77</td>
<td>65.90</td>
<td>5.53</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>62.70</td>
<td>8.22</td>
<td>69.50</td>
<td>5.45</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>62.25$^{b}$</td>
<td>7.91</td>
<td>67.70$^{b}$</td>
<td>5.72</td>
</tr>
<tr>
<td>Fair control</td>
<td>Anonymous</td>
<td>66.45</td>
<td>11.06</td>
<td>68.00</td>
<td>6.71</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>67.45</td>
<td>9.76</td>
<td>69.55</td>
<td>6.87</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>66.95$^{ab}$</td>
<td>10.31</td>
<td>68.77$^{b}$</td>
<td>6.75</td>
</tr>
<tr>
<td>Choice</td>
<td>Anonymous</td>
<td>68.10</td>
<td>5.94</td>
<td>68.60</td>
<td>5.87</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>64.52</td>
<td>8.12</td>
<td>68.36</td>
<td>6.86</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>66.35$^{ab}$</td>
<td>7.22</td>
<td>68.48$^{b}$</td>
<td>6.29</td>
</tr>
<tr>
<td>Familiarity</td>
<td>Anonymous</td>
<td>69.50</td>
<td>11.30</td>
<td>71.00</td>
<td>6.34</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>68.25</td>
<td>7.83</td>
<td>70.40</td>
<td>5.63</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>68.87$^{a}$</td>
<td>9.61</td>
<td>70.70$^{a}$</td>
<td>5.92</td>
</tr>
<tr>
<td>Paper and Pencil</td>
<td>Anonymous</td>
<td>66.35</td>
<td>8.79</td>
<td>68.75</td>
<td>6.04</td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>72.20</td>
<td>8.86</td>
<td>70.95</td>
<td>5.43</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>69.37$^{a}$</td>
<td>9.17</td>
<td>69.85$^{a}$</td>
<td>5.78</td>
</tr>
</tbody>
</table>

$^a$ Within columns, only for all participants, means with the same letter are not significantly different at $P < 0.05$. 


all other groups. No significant differences were found in the self-criticism scores of the TSCS. Thus, higher perceived control did not lead to lower self-deception as we expected. It seems that there is even a tendency towards an opposite effect—stronger control leads to higher socially desirable responding. This effect was most pronounced in the impression management component of the socially desirable behavior. Calculating the correlation between impression management and self-deception for the whole sample yielded quite high relationship, $r = 0.49$ ($P < 0.05$).

3.3. Discussion

3.3.1. Control and social desirability

Our hypotheses on an interaction effect between perceived control and anonymity on the impression management and self-deception components of social desirability were not corroborated. We did not find significant interaction effects between perceived control and anonymity nor did we find a main effect of anonymity. In accordance with our expectation, participants' level of perceived control affected impression management. Participants in the prior familiarity group, who had the strongest sense of control, had the highest impression management scores, whereas the group who had the least control—the weakened control group—had the lowest scores. However, unexpectedly, control manipulations did not affect self-deception scores. No significant differences were found in self-deception scores between groups with heightened level of control and groups which were assigned to lower control conditions. Participants tended to give more socially desirable responses when they had more control.

Participants with higher perceived control took advantage of the control they had to better manage their impression. They portrayed themselves more favorably than others with less sense of control. Since making a good impression may be instrumental to achieving external goals, participants who felt they could more easily manipulate this impression felt less restraints to take the opportunity. It should be noted that the study took place in a setting where all participants were undergoing selection to an elite military unit. Although participants were told that their answers on the questionnaires are part of a separate research that would not affect the decision regarding their selection to the unit, we could not be sure that this did not affect their impression management motivations. Many of them could still be concerned that the study was just a guise for obtaining more information about their personality that could be used by the military. That may also explain the lack of differences between the anonymous and identified context. It is reasonable to assume that in this specific context of test taking, many of them did not believe that the condition was really anonymous. Thus, it seems that when the motivation to achieve some goal is strong, and the possibility to make a good impression exists (i.e. there is more control over the test), participants will take the advantage and will overreport desirable behaviors and underreport their faults in order to maximize the likelihood of achieving their goal.

Yet, the present study offers a way by which social desirability may be reduced. We found that when participants had the lowest levels of perceived control (i.e.
weakened control group) their impression management scores were the lowest, in comparison to all other groups, including the paper-and-pencil group. Thus, it seems that restricting the control over participants' test behavior may reduce their tendency to fabricate their answers.

Contrary to our hypothesis, the self-deception component of social desirability did not decrease in the heightened control conditions. Since self-deception is assumed to be an unconscious process by which the person maintains his or her self-worth, we hypothesized, in accordance with self-affirmation theory (Steele, 1988), that participants who had higher levels of control would not need to deny their faults. The control that they felt would provide them with sufficient ego protection.

Self-affirmation theory was not supported in the present study. Having high control over the situation did not serve as an ego protector. Participants with a stronger sense of control tended to act similarly to those without control in terms of socially desirable responding. There is even some indication that they tended to cope with the threat of exposure by providing responses that are even more acceptable than those with less control.

One possible reason for not obtaining reduced self-deception scores under heightened control situations is that the self-deception measure does not successfully tap the specific elements which people are likely to deny, that are clearly different from the elements which are used for impression management. Tetlock and Manstead (1985) claimed that there is no difference between these two concepts, that the distinction between impression management and intrapsychic explanations such as self-deception for behavior are arbitrary and empirically indiscriminable. Indeed, there are indications that the impression management and self-deception scores of the BIDR (Paulhus, 1984) are not independent. Ones et al. (1996) report observed correlations in the magnitude of 0.40–0.70 between these two components, as well as a similar pattern of correlations between the two measures and the Big Five dimensions. In the present research we used an advanced version of the test. Whereas former studies found lower inter-correlations (0.20–0.32, see Paulhus, 1994), in our sample we found a correlation of 0.49 ($P < 0.05$) and in our pre-test group we found a correlation of 0.59 ($P < 0.01$) between the two scales. Paulhus (1994) points out that the correlations between the scales tend to be particularly high under conditions that exert high demands for self-presentation. It seems that the two scales are not entirely independent in selection settings. However, there is still a great deal of variance for each scale, which raises expectations for different results on impression management and self-deception as a function of relevant manipulations.

A more substantive explanation for the present findings is that the manipulation of perceived control in this study does not serve as an ego protective element. It seems to be a specific form of control that does not affect the person's whole self-image. Participants may not be able to generalize from the sense of control they have in this testing context to their total personality. Therefore, when confronting the threatening items in the questionnaire, the respondents may have not been able to rely on this temporary form of control to overcome the threat. Instead, their sense of control may have served as a means to deny the negative characteristics they possess. Further studies should be conducted to examine what specific
virtues affect the broader self-concept which may serve as a protective compensation minimizing the need for socially desirable behaviors.

3.3.2. Anonymity

The results show that no differences were found between anonymous and identified modes of responses across all groups. The only significant interaction effect was found for the trust variable. It seems that this interaction is due to the highest level of trust expressed by the familiarity—anonymous group, in comparison to all other groups. These results may indicate that strong feelings of control under anonymous testing conditions lead to the highest level of trust in the test.

The general nonsignificant differences between the anonymity and identity situations are in accordance with the results of other studies (Britton, Richardson, Smith, & Hamilton, 1983; Holden et al., 1999; Szekely, Raffeld, & Snodgrass, 1989; Welch, 1990). Yet, there are some studies that demonstrated the effect of anonymity on social desirability scores (Becker, 1976; Booth-Kewley et al., 1992; Lautenschlager & Flaherty, 1990; Paulhus, 1984). Thus, this issue is still not resolved (see also the meta-analysis findings of Richman et al., 1999). As mentioned, it may be assumed that in the present context of testing, the participants were highly motivated to pass the screening tests for the military unit. This may have led them to disregard the fact that they were not required to enter their names as well as our assurance that their responses will have no affect on their admittance to the course. Moreover, responding on the computer may have raised respondents' suspicion that their identity may be identified even without providing their names ("the big brother syndrome", see Rosenfeld, Booth-Kewley, Edwards, & Thomas, 1996).

3.3.3. Comparisons between the paper-and-pencil modes and computerized testing

Participants who answered the tests on the paper-and-pencil mode were no different from their counterparts who answered the computerized version (fair control) in the degree of control that they experienced, in their trust in the test, and in their declared readiness to give genuine responses. Yet, their level of anxiety during test taking was higher. Furthermore, no significant differences were found in the various social desirability scores between the paper-and-pencil and the computerized standard mode.

The findings regarding the effect of mode of administration on social desirability are in accordance with both Dwight and Feigelson's (2000) and Richman et al.'s (1999) meta-analyses data. However, our study highlights the facts that paper-and-pencil and computerized testing are equivalent in other important aspects such as perceived control, trust in the test, and readiness to react candidly.

The present study demonstrates the specific advantages and possibilities of the computerized mode of personality testing. We showed how perceived control, attitude toward the test, and anxiety level can be manipulated by special instructions, constraints, and notes appearing on the screen during the test. These manipulations which are difficult or impossible to conduct on conventional paper-and-pencil studies, open wide and new research possibilities on testing in general and on personality tests in particular.
References


