



# Fake Happy = Happy? Residual Effect of Feigning and Positive Response Bias

Irena Boskovic<sup>1</sup> · Claudia Tejada-Gallardo<sup>2</sup> · Nina Ludwig<sup>1</sup> · Ana Blasco-Belled<sup>2</sup>

Received: 12 May 2023 / Accepted: 7 June 2023 / Published online: 29 June 2023  
© The Author(s) 2023

## Abstract

*Residual effect of feigning* pertains to the findings that individuals, first instructed to feign symptoms, later endorse those symptoms despite being told to respond honestly. We explored whether the same effect would occur if participants were instructed to fabricate positive experiences, such as happiness, by employing a simulation design that included two groups and two testing sessions. The control group received instructions to respond honestly, and the “fake happy” had a task to identify with an exceptionally happy person depicted in a vignette. During Session 1, all participants received the Subjective Happiness Scale, the Satisfaction with Life Scale, and the Supernormality Scale-Revised. In Session 2, after eight days, all participants responded honestly to the Scale of Positive and Negative Experiences and the Flourishing Scale. Overall, fake happy participants, when compared to the control group, exhibited higher levels of supernormality (i.e., positive response bias) and well-being in Session 1. Hence, the instructions had an immediate effect on one’s self-presentation. Yet, the two groups did not differ in Session 2, indicating that the residual effect of feigning was not depicted, or was not strong enough to persist the 8-day period between the sessions. Looking within both groups, all participants significantly lowered their well-being scores overtime. Limitations and implications of this work are discussed.

**Keywords** Feigning · Faking good · Residual effect · Well-being · Supernormality

Research has shown that approximately 30% of the general public distorts their reports about their mental or physical state to reflect their health in a worse way than it actually is (i.e., feigns; Dandachi-Fitzgerald et al., 2020). Yet, such feigning behaviour in the general public does not necessarily signal intentional deception. But when we look at specific populations that have certain *external* incentive to *intentionally* portray their health worse than it is (i.e., exhibit negative response bias), such as forensic sample or compensation applicants, the terms malingering or feigning bad are more suitable (van Oorsouw et al., 2016). The exact prevalence

of malingering is unknown, as it is based on estimates, but the research has indicated that, for physical symptoms reports, prevalence ranges between 20% and 40%, whereas for psychological complaints, the prevalence goes up to 50% (Greve et al., 2009; Mittenberg et al., 2002; Freeman et al., 2008; see also Young, 2015). This discrepancy could be explained simply due to the subjective, therefore rarely verifiable, nature of psychological complaints. The key feature of malingered symptom claims is the largely inflated scores on symptom validity tests (SVTs), which usually contain bizarre or implausible complaints that genuine patients would not endorse (Bianchini et al., 2001). Hence, by endorsing such items, negative response bias is exhibited.

In a similar manner in which people elevate their complaints, they can also hinder them (i.e., exhibit positive response bias or fake good). People who engage in “faking good” intentionally deny or grossly minimize their physical or psychological issues (Rogers, 2018). This behaviour is estimated to be one of the most pervasive source of bias in evaluations, and an even more frequent issue than faking bad, with the estimated prevalence ranging up to 50% (see Griffith & Converse, 2012; Mazza et al., 2020), although

✉ Irena Boskovic  
boskovic@essb.eur.nl

Claudia Tejada-Gallardo  
claudia.tejada@udl.cat

Nina Ludwig  
nina.marlene.ludwig@online.de

Ana Blasco-Belled  
ana.blasco@udl.cat

<sup>1</sup> Erasmus University Rotterdam, Rotterdam, the Netherlands

<sup>2</sup> Universitat de Lleida, Lleida, Spain

the exact prevalence is yet unknown (Rogers, 2018). Recent research showed the engaging in underreporting of psychopathology seems to be more related to the assessment context (high vs. low stake) rather than to the evaluatees' personality (Novo et al., 2022). Therefore, it is not surprising that the typical contexts in which faking good behaviour is exhibited are: job applications (Armour, 2002; Donovan et al., 2003), child custody hearings (Baer & Miller, 2002), parole hearings (Kucharski et al., 2007), therapy (Blanchard & Farber, 2016), and personality assessments (Griffin & Wilson, 2012). In all of these contexts, admitting psychopathology, or any type of socially unacceptable behaviour (e.g., substance abuse) would be considered unfavourable. There are different types of faking good behaviour, such as defensiveness, socially desirable responding (see Rogers, 2018), and *supernormality* (Cima et al., 2008), among others. Supernormality presents systematic denial of even most common symptoms, regardless of the social context (Cima et al., 2003, 2008). Due to its low reliance on social norms, supernormality is a separate type of symptom underreporting than socially desirable responding. However, supernormality does imply social desirability, although this relationship does not go both ways (Cima et al., 2003).

It is important to note that, although equally relevant, faking good behaviour is often times neglected by researchers compared to the volume of research output referring to faking bad, although the study designs for these two research questions do not differ significantly. Response bias, whether negative or positive, is usually experimentally tested by employing a simulation design (Rogers, 2018). This type of study design consists of providing participants with a specific scenario depicting a person for whom it would be beneficial to act in a certain way (e.g., faking bad or faking good), and participants are asked to identify with the protagonist. If compliant, participants in the experimental conditions then exhibit significantly different scores on symptom inventories than healthy participants who were told to respond honestly (i.e., the control group). Interestingly, literature on faking behaviors showed that participants who were instructed to intentionally fabricate certain behaviors (e.g., overreporting neurocognitive and psychiatric symptoms), continue reporting such complaints even *after* being asked to stop faking and to act or respond honestly (Kunst et al., 2016; Merckelbach et al., 2011). Hence, that the reporting instructions employed in the simulation design can actually prompt prolonged changes in one's self-perception of their health (i.e., residual effect of feigning). Further research of this behaviour also showed that this *residual effect of feigning* seems to be resistant to corrective feedback (Merckelbach et al., 2015). The authors argued that the most probable explanation of this effect seem to fall under the cognitive dissonance theory (Festinger & Carlsmith, 1959). When one's beliefs and their actions are in conflict, the person

might adjust their beliefs in order to align them with exhibited behaviour, hence, decrease the discrepancy between the two. For instance, people usually wish to think of themselves as "good", yet, "good people" do not engage in faking. Hence, one potential way of solving that discrepancy would be to convince oneself that the faked behaviors/emotions are indeed genuinely experienced (i.e., internalization of symptoms; Merckelbach & Merten, 2012). However, the above-mentioned studies inspected only faking bad behaviour, therefore, it is currently unknown whether the residual effect of feigning works as well when people exhibit positive response bias (i.e., "fake good").

The majority of literature on faking good is mainly focused on personality traits (e.g., Mazza et al., 2020), hence, one important aspect is oftentimes neglected, such as well-being or, simply put, happiness. Happiness has been associated with positive social impressions and other beneficial outcomes (e.g., it facilitates the pursuit of goals or contributes to vital social bonds), and it is considered as an important societal value (Eid & Diener, 2001; Oishi et al., 2007). Hence, there are strong motives for people to still strive to appear as happy to others (Moore et al., 2017). Some of the previous research supports the notion that faking happiness would increase well-being and other, in contrast, provided evidence showing that faking happiness might have counterproductive effects (e.g., Abele, 2003; Ferguson and Sheldon, 2013; Hülshager & Schewe, 2011). Yet, none of these studies specifically tested the overtime impact the instructions prompting people to appear more happy could have on one's self-perception (i.e., residual effect of feigning).

## Current Study

Research has shown that attorneys very often coach their clients to fake good (see Baer & Miller, 2002). In the forensic context, for instance the custody hearings, presetting oneself as a happy and healthy person might aid being viewed as a better suited guardian, hence, this type of coaching is unsurprising. Therefore, in this study, we sought to achieve two goals: first, to test the efficacy of simulation instructions (i.e., coaching) in evoking positive response bias, and second, to check if such instructions could trigger the residual effects of feigning, which has not yet been tested in the domain of faking good behaviour. We operationalised happiness as subjective well-being (hedonic approach; Diener, 1984) and psychological well-being (eudaimonic approach; Ryff, 1989). Subjective well-being involves a combination of a cognitive (i.e., life satisfaction) and two affective components (i.e., positive and negative affect; Cummins & Cahill, 2000), whereas psychological well-being is explained by positive (and effective) functioning that comes from the development

and realization of the human potential (self-fulfillment; Waterman, 1993). Student participants were randomly allocated to either the control (genuine) group or to a fake happy group. All participants needed to fill out well-being measures during two sessions separated by 7-day period (period was extended to eight days, see Participants). Members of the control group received instructions to respond honestly, whereas the fake happy group first received instructions to fake good, and then to respond honestly during the second session. Based on prior research on the residual effect of feigning, we anticipated that the fake happy group would exhibit higher well-being scores than the control group on the measures provided in both sessions.

## Method

### Participants

Initially, 152 bachelor students joined our study. The inclusion criteria were that (1) participants completed all of the questions ( $n_{\text{excluded}} = 24$ ); (2) that participants provided a traceable code in order to combine their data from two sessions ( $n_{\text{excluded}} = 12$ ); (3) that participants completed session 2 seven days after session (the period was extended to eight days as nine participants finished the study just a few hours after the seventh day;  $n_{\text{excluded}} = 21$ ). This left a total of 95 participants.

The average age of the participants was 20 years old (range: 18–26;  $SD = 1.69$ ), and the majority was female (87.4%; male 10.5%, and two participants did not want to disclose their gender). The most frequent nationalities within our sample were Dutch (42%), German (11%), British (3%), Turkish (3%), and Bulgarian (3%). As the study was conducted in English, we checked participants' proficiency in English on a 5-point scale (1 indicating poor proficiency, 5 indicating C2 or native level). The results showed that participants had very good proficiency in English ( $M = 4.32$ ;  $SD = 0.64$ ; range: 3–5). Then, participants were randomly divided into the control group ( $n = 52$ ) and the manipulation (i.e., fake happy) group ( $n = 43$ ). The groups did not differ in terms of age ( $t(93) = 1.40$ ,  $p = .166$ ), nor in English proficiency ( $t(93) = 0.13$ ,  $p = .894$ ).

### Measures

**Self-Rating of Happiness Scale (SRHS; Abdel-Khalek, 2006).** The SRHS was developed to measure individuals' happiness using a single self-report item ("How satisfied are you with your life as a whole?") that is scored on a 11-point Likert scale (anchors: 0 = completely dissatisfied; 10 = completely satisfied). This measure was applied to check participants' baseline happiness.

**Subjective Happiness Scale (SHS; Lyubomirsky and Lepper, 1999).** The SHS is a 4-item measure of one's overall happiness that are scored on a 7-point Likert scale. Each item (e.g., "In general, I consider myself:") presents dichotomous sentence fragments (e.g., from "not a very happy person" to "a very happy person") to which the respondent indicates to what extent they identify with each item. A higher score represents greater levels of happiness. In the current study, the *Cronbach's alpha* of the scale was 0.84.

**Satisfaction with Life Scale (SWLS; Diener et al., 1984; Pavot & Diener, 2008).** The SWLS is a brief 5-item self-report measure of one's global life satisfaction. Responses for each item (e.g., "In most ways my life is close to ideal.") are recorded on a 7-point Likert scale (anchors: 1 = strongly disagree; 7 strongly agree) and possible scores range from 5 to 35 where a higher score indicates a higher satisfaction. In the current study, the *Cronbach's alpha* for the SWLS was 0.87.

**Supernormality Scale – Revised (SS-R; Cima et al., 2008).** The SS-R is a revised and extended version of the Supernormality Scale that was initially constructed by Cima et al. (2003) as a self-report measure of supernormality (Cima et al., 2008). The SS-R contains 50 items (e.g., "When I walk on the street late at night, I have the feeling that I am being followed") that are rated on a 4-point Likert scale (anchors: 1 = not at all applicable; 4 = extremely applicable) and subdivided into seven subscales. Namely, the subscales address (1) social desirability, (2) mood disorders, (3) obsessive compulsive symptoms, (4) psychotic symptoms, (5) dissociative symptoms, (6) aggression, and (7) anxiety. Additionally, 16 distractor items are included (e.g., "I like to watch television"). Subscale scores are computed as well as a total score. The *lower* the score, the more an individual appears to present supernormality and the cut off is set at 60. In the current study, the *Cronbach's alpha* for the total SS-R was 0.89.

**Scale of Positive and Negative Experiences (SPAN; Diener et al., 2009).** The SPAN is a 12-item self-report measure of positive and negative experiences that reflects various good or bad feelings such as boredom, interest, physical pleasure. The respondent is asked to think about what they have been doing in the past four weeks and then indicate on a 5-point Likert scale (anchors: 1 = very rarely or never; 5 = very often or always) how often they experienced each feeling (e.g., "positive", "unpleasant", or "joyful"). A positive (SPAN-P) and a negative scale score (SPAN-N) can be computed (range from 6 to 30), as well as an overall *balance score* (SPAN B) between those two by subtracting the negative score from the positive one (range from -24 to +24). The more extreme a balance score, the more an individual tends to identify those feelings in themselves (Diener et al., 2009). In the current study, the *Cronbach's*

*alphas* for the SPANE-P and the SPANE-N were 0.92 and 0.85, respectively.

**Flourishing Scale (FS; Diener et al., 2010).** The FS is a short self-report measure that taps into a broad range of components of social-psychological functioning essential to well-being. Its eight items (e.g., “I lead a purposeful and meaningful life”) are scored on a 7-point Likert scale (anchors: 1 = strongly disagree; 7 = strongly agree) and address the respondent’s own perception of their relationships, self-esteem, optimism, and purpose in life. The higher the total score (ranging from 8 to 56), the more positive the respondent seems to view themselves across those diverse domains. In the current study, the *Cronbach’s alpha* of 0.90 indicated excellent internal consistency.

The *SRHS* and the *SS-R* were included in the first measurement occasion (Session 1) to control for participants’ baseline happiness levels and check for instruction compliance, respectively, whereas the other measures were conceptually paired into measures of more affective well-being (Session 1: *SHS* and Session 2: *SPANE*) and more cognitive/eudaimonic well-being (Session 1: *SWLS* and Session 2: *FS*). The reasoning behind such pairings is justified by previous work showing significant correlations between these measures (*Pair 1.* *SHS* and *SPANE P* correlations ranging from 0.61; Proctor et al., 2015, to 0.83; Rice & Shorey-Fennell, 2020; with *SPANE N* – .65, Proctor et al., 2015; *Pair 2.* *SWLS* and *FS* correlations ranging from 0.71, Ramsay et al., 2023) to 0.76, Fong and Loi, 2016).

## Procedure

Students signed up online via the university-internal [removed for review] website and then received a Qualtrics link to the study. Data were collected in the period of April–May 2021. After receiving study information and providing the informed consent, participants were asked about their age, gender, level of education, nationality, as well as their English proficiency. To measure participants’ baseline happiness, they were administered the one-item *SRHS*. Thereafter, they were randomly assigned to one of two conditions (i.e., control and fake happy groups). The randomization was a priori set in Qualtrics, so it was done automatically. However, if a participant endorsed the highest score on the happiness baseline measure (i.e., 10), they were automatically assigned to the control condition because it would not be possible to evaluate whether those participants showed any potential residual effects of our manipulation ( $n = 3$ ). Participants in the control condition were instructed to answer all questionnaires truthfully. The fake happy group was asked to read a vignette about a person who is overall happy and has generally a very wholesome life, and were told to imagine being the protagonist. We based our vignette on two well-being dimensions; (1) indicators of momentary

positive affect/emotions (e.g., enjoys music in the morning) and satisfaction with life (e.g., is very fulfilled and loves s/ her life); and (2) indicators of positive relationships (e.g., loves meet people and is in a loving relationship), meaningful life (i.e., loves to learn and is very productive), and personal growth (e.g., is very understanding towards other people and her/himself). Subsequently, they were asked to respond to the questionnaires as if they were the happy person from the vignette. The first part entailed three questionnaires, namely, the *SHS*, the *SWLS*, and the *SS-R*. The latter scale was included to inspect whether participants were compliant with the instructions and if they were exhibiting positive response bias. Subsequently, participants answered three questions about their motivation to participate in this project, about the clarity of the questions and instructions, and about the difficulty of the survey. They were asked to create a personalized code that they should use in session 2, so that their responses could be merged.

Seven days later, the link to the second part was automatically sent to the participants via Qualtrics. The second part contained a question for the code and two well-being questionnaires, namely the *SPANE* and the *FS*. Rather than using the same measures at time 1 and time 2, multiple different measures of well-being were chosen to avoid the recursiveness of scales and carry-over effects. Among the scale items, several items that required a specific answer were added as attention and effort checks. All participants passed the checks. Finally, participants filled in the same evaluation questions as from part one. All participants received a debriefing form. Students who completed both parts of the study received one research credit as compensation. The study was approved by the Ethics Review Committee of the [removed for review].

## Data Analyses

The data were tested in statistical package SPSS, using the independent *t*-tests and the pair *t*-tests combining the measures in Session 1 and Session 2 (pair a: *SHS* – *SPANE* Balance; pair b: *SWLS* – *FS*). For the interpretation of effect sizes, we followed Cohen’s (1988) categorization of <0.20 as small, 0.50 medium, and 0.80 as large effect. The data and the outputs are available at Open Science Framework ([https://osf.io/rfh4w/?view\\_only=1e41f0e0e24440d382eca49a4a8490ac](https://osf.io/rfh4w/?view_only=1e41f0e0e24440d382eca49a4a8490ac)).

## Results

### Motivation, Clarity, Difficulty

Participants were asked to rate their motivation, the clarity, and the difficulty of the tasks for both study sessions on a 5-point scale. In Session 1, the overall ratings were high

**Table 1** Participants’ rating on motivation, clarity, and difficulty of the task across groups

		<b>Control group</b> <i>M (SD)</i>	<b>Fake Happy group</b> <i>M (SD)</i>	<b><i>t</i>-test</b>	<b><i>p</i></b>
	<i>N</i>	52	43	<i>df</i> =93	
<i>Session 1</i>	Motivation	3.81 (0.56)	3.72 (0.63)	0.71	0.480
	Clarity	4.75 (0.44)	4.47 (0.73)	2.34	<b>0.021</b>
	Difficulty	4.15 (0.78)	3.98 (0.88)	1.04	0.302
<i>Session 2</i>	Motivation	2.65 (0.68)	2.56 (0.85)	0.61	0.545
	Clarity	4.67 (0.83)	4.74 (0.49)	0.50	0.624
	Difficulty	4.42 (0.67)	4.33 (0.78)	0.66	0.510

( $M_{\text{motivation}} = 3.77, SD = 0.59; M_{\text{clarity}} = 4.62, SD = 0.60; M_{\text{difficulty}} = 4.07, SD = 0.83$ ). In Session 2, the ratings for motivation were slightly lower ( $M = 2.61, SD = 0.76$ ), whereas the ratings for clarity and difficulty remained high ( $M_{\text{clarity}} = 4.71, SD = 0.70; M_{\text{difficulty}} = 4.38, SD = 0.72$ ). The means and standard deviations for each group, as well as independent *t*-tests are presented in the Table 1. We checked whether the groups differed in their ratings using the independent *t*-tests. The only significant difference ( $p < .05$ ) was found in the clarity of the task in Session 1, although both groups provided high scores.

**Testing Happiness: Session 1 and Session 2**

We also inspected group differences on all of the used measures using the independent *t*-tests. The means and standard deviations on each measure across groups are provided in Table 2. Before introducing the instructions and other measures, we asked participants to report their overall life

satisfaction (0–10), and students in both groups reported similar, above-average levels on this measure. Subjects provided similar scores on most other measures as well. The significant group differences ( $p_s < .001$ ) were found only on SWLS (*Cohen’s d* = 1.11) and SS-R (*Cohen’s d* = 0.82), both given during Session 1, right after the instructions to either respond honestly or to fake happiness. The fake happy group exhibited higher levels of life satisfaction and higher rates of supernormality. The data were checked also using the non-parametric Mann-Whitney U test, but the results did not differ.

**Testing Happiness over Time**

In order to test whether participants indicated stable levels of happiness within their groups overtime, we conducted *pair t*-tests. We compared (*standardized scores*) on SHS with SPANE B (Balance score = SPANE P - SPANE N), and (*standardized*) scores on SWLS and FS. The results are

**Table 2** Means, standard deviations, and *t*-tests for all used measures at Session 1 and Session 2

		<b>Total Sample</b> <i>M (SD)</i>	<b>Control Group</b> <i>M (SD)</i>	<b>Fake Happy Group</b> <i>M (SD)</i>	<b><i>t</i>-test</b>	<b><i>p</i></b>	<b><i>Cohen’s d</i></b>
		95	52	43	<i>df</i> =93		
<i>Baseline</i>	<b>SRHS</b> (0–10)	6.97 (1.34)	7.10 (1.39)	6.84 (1.27)	0.94	0.350	0.19
	<b>SHS</b> (1–7)	4.98 (1.17)	4.89 (1.01)	5.08 (1.34)	0.80	0.427	0.16
<i>Session 1</i>	<b>SWLS</b> (1–7)	5.18 (1.15)	4.67 (1.12)	5.79 (0.86)	<b>5.37</b>	<b>0.001</b>	1.11
	<b>SS-R*</b> (0-200)	77.67 (16.24)	83.31 (14.40)	70.86 (15.87)	<b>4.01</b>	<b>0.001</b>	0.82
	<b>SPANE P</b> (1–5)	3.51 (0.68)	3.55 (0.75)	3.47 (0.60)	0.55	0.582	0.11
<i>Session 2</i>	<b>SPANE N</b> (1–5)	2.80 (0.73)	2.75 (0.75)	2.86 (0.72)	0.66	0.512	0.13
	<b>FS</b> (1–7)	5.27 (0.95)	5.33 (1.00)	5.19 (0.91)	0.68	0.497	0.14

Self-Rating of Happiness Scale (SRHS), Subjective Happiness Scale (SHS), the Satisfaction with Life Scale (SWLS), Supernormality Scale – Revised (SS-R), Scale of Positive and Negative Experiences (SPANE P and SPANE N), Flourishing Scale (FS)

\*Lower values on SS-R indicate a stronger presence of positive response bias



**Table 3** Standardized Z values and pair t-tests for control and happy group

		Control Group	Pair t-test	<i>p</i>	Happy Group	Pair t-test	<i>p</i>
		<i>Z</i> values ( <i>SD</i> )	<i>df</i> =51		<i>Z</i> values	<i>df</i> =42	
Pair 1	SHS	0.074 (0.87)	1.05	0.296	0.09 (1.14)	0.87	0.389
	SPANE B	0.063 (1.07)			0.076 (0.92)		
Pair 2	SWLS	0.440 (0.97)	3.27	0.002	0.53 (0.75)	3.94	< 0.001
	FS	0.064 (1.04)			0.077 (0.95)		

Subjective Happiness Scale (SHS), Scale of Positive and Negative Experiences – Balanced Score (SPANE B), the Satisfaction with Life Scale (SWLS), Flourishing Scale (FS); Control group: Pair 1 Cohen's  $d_{rm}$  = 0.17, Pair 2 Cohen's  $d_{rm}$  = 0.47, Happy group: Pair 1: Cohen's  $d_{rm}$  = 0.12; Pair 2: Cohen's  $d_{rm}$  = 0.69

presented in Table 3. Our results indicated that both groups dropped their scores significantly ( $p < .001$ ) over time regarding life satisfaction. The data were also checked using Wilcoxon Signed Ranks test, and the results did not differ.

## Discussion

In this study, the goal was to investigate whether the *residual effect of feigning* (Merckelbach et al., 2011) occurs when used for faking happiness, specifically defined through the concepts of affective and cognitive indicators of well-being. Our results are as follows: First, looking at Session 1, our data from the supernormality scale (Cima et al., 2008) showed that feigners' scores were significantly more indicative of biased reporting than the scores of the control group. Hence, feigners significantly more denied even everyday, common symptoms (i.e., exhibited a tendency to show themselves as “supernormal”; Cima et al., 2003; Cima et al., 2008) than the control group. This confirms the immediate effect of instructions on the way one presents themselves also in the domain of symptom underreporting. It is important to note that such prompts could often be provided by attorneys or counselors in, for instance, parole or child custody hearings, when it is beneficial for a person to present themselves in a best light possible. What our results show is that such coaching can have an immediate effect.

Second, and in accordance with the previous point, the results are indicating that, despite obtaining similar baseline scores, the two groups significantly differed also on the life satisfaction measure administered during Session 1. Specifically, feigners exhibited higher life satisfaction than the control group. These findings partially support our expectations that the fake happy group will exhibit higher scores than the control group. Yet, the difference was not found for the happiness measure. Looking at the affective well-being scores, both groups exhibited an above average level of it, hence, it is likely that the control group exhibited too high levels for feigners to overtake, especially if they did not want to “overdo” their task, in order to appear credible in their assignment. Also, we did have participants who reported

the maximum score, hence, we automatically allocated to the control condition, which could partially explain the lack of group difference.

Third, however, feigners' engagement was short-lived, as the effect of feigning was only present at Session 1, right after they received the instructions. During Session 2, our fake happy participants obtained scores similar to those of the control group. Therefore, looking at the lack of differences between groups at Session 2, we did not succeed to induce the residual effect of feigning. This lack of difference could be due to the time between the sessions, which might have been too long for participants to remember the instructions. Indeed, the studies by Merckelbach and colleagues (2011), as well as the replication by Kunst et al. (2016), included intervals not longer than two hours. The original authors suggested that future investigation should introduce longer intervals between testing sessions, however, it appears that the 8-day interval was too long for the residual effect of feigning, if existing, to remain. Therefore, further research should incorporate a few testing moments so as to determine the stability of this effect. For instance, employing an experience sampling method (Verhagen et al., 2016) might be the most appropriate as participants would report their state every day repeatedly (see also Boskovic et al., 2021). This investigation would be beneficial as it would potentially indicate a peak of the residual effect, as well as its lowest point, which could be informative for, for instance, validity of repeated assessment. Although these findings suggest that feigning happiness does not reflect on increased well-being scores overtime, it is possible that symptom underreporting, as a key feature of supernormality, would persist. However, as no other measure except SS-R was available to us, we did not administer such test at Session 2. Future studies should monitor symptom underreporting overtime as well.

Fourth, we also checked for any potential over-time changes *within* the groups. We compared the standardized scores on measures that tap into similar constructs in order to observe any potential changes in participants' happiness. It seems that both groups reported significantly lower levels of well-being during Session 2. This was not surprising for the fake happy group, considering

the instructions they previously received and the briefness of residual effect, however, the same direction of score change was found in the control group. Together, these findings might fall under the “*initial elevation phenomenon*”, which was found specifically in self-reported data (Anvari et al., 2022). Namely, research subjects given the self-report questionnaires tend to initially provide significantly higher scores than those provided at later times. Thus far, this phenomenon was only found in the context of negative subjective experiences (mood and physical health symptoms; Anvari et al., 2022; pain; Boskovic et al., 2021), hence, our data might be novel in a sense that it indicates that the initial elevation phenomenon could even apply to self-reported positive experiences.

There are a few limitations of our work that need to be addressed. First, our sample consisted of well-adjusted young (female) adults (i.e., students), which limits the generalizability of our findings. Future studies should try to employ the same intervention among clinical participants such as patients with depression or other psychological symptomatology. It is also important to mention that our sample might have been too small for the employed design, so we suggest that future investigations aim at including larger samples. Second, we did not check for the presence of any psychopathology among our sample, which could have impacted the stability of the initial significantly different scores between groups. Third, the given vignette might not accurately reflect the participants’ vision of a happy person, hindering the task of “embodying” of the simulated case to respond to the measures. Fourth, even though all participants attend their studies in English, there were not native English speakers, and our testing material was only provided in that language. Hence, although unlikely, it is possible that some of our participants did not comprehend all of the administered items. Fifth, as discussed above, the time period between the two sessions in our study was significantly longer than the intervals used in the previous work. Future work should possibly include either some mid-periods or employ the experience sampling method with multiple testing points. Sixth, based on the study evaluation questions, we could see that the fake happy group provided significantly lower scores at Session 1 regarding the clarity of the instructions. This is reasonable considering that their instructions were much more complex than those of the control, yet, their scores were above 4.4, indicating still very high levels of clarity. Finally, our participants came from diverse cultural backgrounds, which is important to consider, especially in the context of happiness investigation. While West countries thrive on enhancing happiness, expressing happiness in other cultures might be considered rude or even completely irrelevant (Joshani & Weijers, 2014). Therefore, future studies should build

on these limitations, aiming to include a more uniform sample in terms of cultural background and language.

## Conclusion

Overall, in this study, we tested the strength and stability of the residual effect of feigning in the domain of faking good. We showed that instructing people to fake good (i.e., to fake happiness) leads to significantly increased self-reported well-being, as well as to underreporting of everyday symptoms, when measured right after the instructions. However, this trend was short-lived, as no residual effect of fake happiness on the well-being scores was captured eight days after instructions, which was probably too long of an interval for the effect so far only investigated in short sequences (e.g., 1 h). Therefore, future investigations, preferably with multiple and shorter testing periods, are encouraged.

**Authors’ Contribution** All listed authors significantly contributed to the execution of the study and to this manuscript.

**Funding** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Declarations

**Ethical Approval** All procedures performed in this study were in accordance with the ethical standards of the Ethical committee of Erasmus University Rotterdam, the Netherlands, and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

**Conflict of Interest** The authors have no conflict of interest to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abdel-Khalek, A. M. (2006). Measuring happiness with a single-item scale. *Social Behavior and Personality: An International Journal*, 34(2), 139–150. <https://doi.org/10.2224/sbp.2006.34.2.139>.
- Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, 85, 768–776. <https://doi.org/10.1037/0022-3514.85.4.768>.

- Anvari, F., Efendić, E., Olsen, J., Arslan, R. C., Elson, M., & Schneider, I. K. (2022). Bias in self-reports: An initial elevation phenomenon. *Social Psychological and Personality Science*, 19485506221129160.
- Armour, S. (2002). *Security checks worry workers*. *USA Today*, 1b.
- Baer, R. A., & Miller, J. (2002). Underreporting of psychopathology on the MMPI-2: A meta-analytic review. *Psychological Assessment*, 14(1), 16–26. <https://doi.org/10.1037//1040-3590.14.1.16>.
- Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist*, 15, 19–45. <https://doi.org/10.1076/clin.15.1.19.1907>.
- Blanchard, M., & Farber, B. A. (2016). Lying in psychotherapy: Why and what clients don't tell their therapist about therapy and their relationship. *Counselling Psychology Quarterly*, 29(1), 90–112. <https://doi.org/10.1080/09515070.2015.1085365>.
- Boskovic, I., Zwaan, L., Baillie, V., & Merckelbach, H. (2021). Consistency does not aid detection of feigned symptoms, over-reporting does: Two explorative studies on symptom stability among truth tellers and feigners. *Applied Neuropsychology: Adult*, 1–9. <https://doi.org/10.1080/23279095.2021.188872>.
- Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. *The Clinical Neuropsychologist*, 17(2), 235–243.
- Cima, M., Van Bergen, S., & Kremer, K. (2008). Development of the supernormality scale-revised and its relationship with psychopathy. *Journal of Forensic Sciences*, 53(4), 975–981. <https://doi.org/10.1111/j.1556-4029.2008.00740.x>.
- Cohen, J. (1988). *Statistical Power Analysis for the Social Sciences (2nd Edition)*. Hillsdale, New Jersey, Lawrence Erlbaum Associates.
- Cummins, R., & Cahill, J. (2000). Avances en la comprensión de la calidad de vida subjetiva. *Psychosocial Intervention*, 9(2), 185–198.
- Dandachi-FitzGerald, B., Merckelbach, H., Bošković, I., & Jelicic, M. (2020). Do you know people who feign? Proxy respondents about feigned symptoms. *Psychological Injury and Law*, 13, 225–234. <https://doi.org/10.1007/s12207-020-09387-6>.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95, 542–575. <https://doi.org/10.1037/0033-2909.95.3.542>.
- Diener, E., Wirtz, D., Biswas-Diener, R., Tov, W., Kim-Prieto, C., Choi, D. W., & Oishi, S. (2009). New measures of well-being. *Assessing well-being: The collected works of Ed Diener*, 247–266.
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, 97, 143–156. <https://doi.org/10.1007/s11205-009-9493-y>.
- Donovan, J. J., Dwight, S. A., & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance*, 16(1), 81–106. [https://doi.org/10.1207/S15327043HUP1601\\_4](https://doi.org/10.1207/S15327043HUP1601_4).
- Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures: Inter- and intrainational differences. *Journal of Personality and Social Psychology*, 81(5), 869–885. <https://doi.org/10.1037/0022-3514.81.5.869>.
- Ferguson, Y. L., & Sheldon, K. M. (2013). Trying to be happier really can work: Two experimental studies. *The Journal of Positive Psychology*, 8, 23–33. <https://doi.org/10.1080/17439760.2012.747000>.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58, 203–210.
- Fong, M., & Loi, N. M. (2016). The mediating role of self-compassion in student psychological health. *Australian Psychologist*, 51(6), 431–441. <https://doi.org/10.1111/ap.12185>.
- Freeman, T., Powell, M., & Kimbrell, T. (2008). Measuring symptom exaggeration in veterans with chronic posttraumatic stress disorder. *Psychiatry Research*, 158, 374–380. <https://doi.org/10.1016/j.psychres.2007.04.002>.
- Greve, K. W., Ord, J. S., Bianchini, K. J., & Curtis, K. L. (2009). Prevalence of malingering in patients with chronic pain referred for psychologic evaluation in a medico-legal context. *Archives of Physical Medicine and Rehabilitation*, 90, 1117–1126. <https://doi.org/10.1016/j.apmr.2009.01.018>.
- Griffin, B., & Wilson, I. G. (2012). Faking good: Self-enhancement in medical school applicants. *Medical Education*, 46(5), 485–490. <https://doi.org/10.1111/j.1365-2923.2011.04208.x>.
- Griffith, R. L., & Converse, P. D. (2012). The rules of evidence and the prevalence of applicant faking. In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 34–52). Oxford University Press.
- Hülshager, U. R., & Schewe, A. F. (2011). On the costs and benefits of emotional labor: A meta-analysis of three decades of research. *Journal of Occupational Health Psychology*, 16, 361–389. <https://doi.org/10.1037/a0022876>.
- Joshanloo, M., & Weijers, D. (2014). Aversion to happiness across cultures: A review of where and why people are averse to happiness. *Journal of Happiness Studies*, 15(3), 717–735. <https://doi.org/10.1007/s10902-013-9489-9>.
- Kucharski, L. T., Tang, P., & Rosenfeld, B. (2007). Detecting denial of severe psychiatric disorder: An MMPI-2 investigation of criminal defendants. *International Journal of Forensic Mental Health*, 6(2), 145–152. <https://doi.org/10.1080/14999013.2007.10471259>.
- Kunst, M. J. J., Aarts, M., Frolijk, J., & Poelwijk, R. (2016). Exploring the residual effects of malingering in mock victims of violent crime: An experimental vignette study. *Psychology Crime & Law*, 22(3), 238–251. <https://doi.org/10.1080/1068316X.2015.1085981>.
- Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, 46, 137–155. <https://doi.org/10.1023/A:1006824100041>.
- Mazza, C., Monaro, M., Burla, F., Colasanti, M., Orrù, G., Ferracuti, S., & Roma, P. (2020). Use of mouse-tracking software to detect faking-good behavior on personality questionnaires: an exploratory study. *Scientific Reports*, 10(1), 1–13. <https://doi.org/10.1038/s41598-020-61636-5>.
- Merckelbach, H., & Merten, T. (2012). A note on cognitive dissonance and malingering. *The Clinical Neuropsychologist*, 26, 1217–1229. <https://doi.org/10.1080/13854046.2012.710252>.
- Merckelbach, H., Jelicic, M., & Pieters, M. (2011). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, 33, 131–139. <https://doi.org/10.1080/13803395.2010.495055>.
- Merckelbach, H., Dandachi-FitzGerald, B., van Mulken, P., Ponds, R., & Niesten, E. (2015). Exaggerating psychopathology produces residual effects that are resistant to corrective feedback: An experimental demonstration. *Applied Neuropsychology: Adult*, 22, 16–22. <https://doi.org/10.1080/23279095.2013.816850>.
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24, 1094–1102. <https://doi.org/10.1076/jcen.24.8.1094.8379>.
- Moore, C., Lee, S. Y., Kim, K., & Cable, D. M. (2017). The advantage of being oneself: The role of applicant self-verification in organizational hiring decisions. *Journal of Applied Psychology*, 102, 1493–1513. <https://doi.org/10.1037/apl0000223>.
- Novo, R., Gonzalez, B., & Roberto, M. (2022). Beyond personality: Underreporting in high-stakes assessment contexts. *Personality and Individual Differences*, 184(2022), 111190. <https://doi.org/10.1016/j.paid.2021.111190>



- Oishi, S., Diener, E., & Lucas, R. E. (2007). The optimum level of well-being: Can people be too happy? *Perspectives on Psychological Science*, 2(4), 346–360. <https://doi.org/10.1111/j.1745-6916.2007.00048.x>.
- Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology*, 3(2), 137–152. <https://doi.org/10.1080/17439760701756946>.
- Proctor, C., Tweed, R., & Morris, D. (2015). The naturally emerging structure of well-being among young adults: “Big Two” or other framework? *Journal of Happiness Studies*, 16, 257–275. <https://doi.org/10.1007/s10902-014-9507-6>.
- Ramsay, J. E., Wang, D., Yeo, J. S., Khong, Z. Y., & Tan, C. S. (2023). Perceived authenticity, Machiavellianism, and psychological functioning: An inter-domain and cross-cultural investigation. *Personality and Individual Differences*, 204, 112049. <https://doi.org/10.1016/j.paid.2022.112049>.
- Rice, S. P., & Shorey-Fennell, B. R. (2020). Comparing the psychometric properties of common measures of positive and negative emotional experiences: Implications for the assessment of subjective wellbeing. *Journal of Well-Being Assessment*, 4, 37–56. <https://doi.org/10.1007/s41543-020-00025-1>.
- Rogers, R. (2018). An introduction to the response bias. In Rogers, R. & Bender, S. D. (Eds.). *Clinical Assessment of Malingering and Deception* (4th Edition). Guilford.
- Ryff, C. D. (1989). Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, 57, 1069–1081. <https://doi.org/10.1037/0022-3514.57.6.1069>.
- Van Oorsouw, K., Jelicic, M., & Cima, M. (2016). Deception: Faking good or bad. *The handbook of forensic psychopathology and treatment* (pp. 155–168). Routledge.
- Verhagen, S. J., Hasmi, L., Drukker, M., van Os, J., & Delespaul, P. A. (2016). Use of the experience sampling method in the context of clinical trials. *Evidence-Based Mental Health*, 19(3), 86–89. <https://doi.org/10.1136/ebmental-2016-102418>.
- Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, 64, 678–691. <https://doi.org/10.1037/0022-3514.64.4.678>.
- Young, G. (2015). Malingering in forensic disability-related assessments: Prevalence 15 ± 15%. *Psychological Injury and Law*, 8, 188–199. <https://doi.org/10.1007/s12207-015-9232-4>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.