Are facial feedback effects solely driven by demand characteristics? An experimental investigation

Nicholas A. Coles[1,2] (email) Brooke Frohlich[2], Jeff T. Larsen[2], Lowell Gaertner[2]

[1]Harvard Kennedy School, Harvard University, Cambridge, USA

[2]Department of Psychology, University of Tennessee, Knoxville, USA

Note: This manuscript is a pre-print that has not yet undergone peer review. Please cite responsibly. In addition to usual peer review, a team of financially-incentives research critics are being recruited to provide additional peer review. (See The 20% Statistician blog for a description of this "Red Team Challenge".) All materials, data, and code are available here. The pre-registration can be found here.

Abstract

The facial feedback hypothesis suggests that an individual's facial expressions can influence their emotional experience (e.g., that smiling can make one feel happier). However, a reoccurring concern is that demand characteristics drive these effects. To examine this, we had university students pose happy, angry, and neutral expressions and self-report their emotions following a demand characteristics manipulation, wherein we either (a) told participants we hypothesized their poses would influence their emotions, (b) told participants we hypothesized their poses would *not* influence their emotions, or (c) did not tell participants a hypothesis. Results indicated that demand characteristics moderated the effects of facial poses on self-reported emotion. However, facial poses still influenced self-reported emotion when participants were told we hypothesized their poses would *not* influence emotion. These results highlight the role of demand characteristics in facial feedback research but indicate that demand characteristics do not fully account for observed facial feedback effects.

Keywords: facial feedback hypothesis, demand characteristics, emotion, embodiment, facial expressions

Statement of Relevance: Consistent with theories that posit emotional experience is partially built off afferent signals from the peripheral nervous system, previous research suggests that posing facial expressions of emotion can influence emotional experience (e.g., that smiling can make one feel happier). However, a prevailing concern is that these observed facial feedback effects are merely an artifact of demand characteristics. We conducted an experiment that demonstrated that demand characteristics strongly shape the degree to which posing facial expressions influence emotion. However, even when people are told to not expect these effects, posing facial expressions can still influence emotion.

Most modern theories of emotion posit that emotional experience is partially built off afferent signals from the peripheral nervous system (for a review, see Moors, 2009). If true, it should be possible to modify people's emotional experience by altering their peripheral nervous system activity. Consistent with this possibility, a meta-analysis on the *facial feedback hypothesis* indicates that posing facial expressions of emotion can influence self-reported emotion (Coles, Larsen, & Lench, 2019). However, a prevailing concern is that these observed effects are merely an artifact of demand characteristics: cues that participants may use to form beliefs about, and subsequently adjust their behavior to fit, the researcher's hypothesis (Buck, 1980; Coles, Larsen, Kuribayashi, & Kuelz, 2019; Dimberg & Söderkvist, 2011; Schimmack & Chen, 2017). To address this concern, some researchers have used elaborate cover stories that mask the study's purpose. For example, in a study ostensibly about psychomotor coordination, Strack, Martin, and Stepper (1988) had participants view humorous cartoons while holding a pen in their mouth in a manner that either induced or inhibited smiling. Despite being apparently unaware of the study's true purpose, participants reported feeling more amused by the cartoons when smiling. However, more recently, a large multi-site study failed to replicate this pen-in-mouth effect (Wagenmakers et al., 2016). Although there are disagreements about its implications, this failure-to-replicate has revived discussions regarding the role of demand characteristics in facial feedback research (Noah & Mayo, 2018; Schimmack & Chen, 2017; Strack, 2016).

Whereas previous facial feedback researchers have strived to eliminate demand characteristics, the present study features a different tactic: the manipulation of demand characteristics. Prior to posing happy, angry, and neutral expressions and self-reporting their emotions, participants were randomly assigned to one of three demand conditions. In the *positive*

expectation condition, we told participants we expected their facial poses to influence their emotions. In the *null expectation* condition, we told participants we did not expect their facial poses to influence their emotions. In the *control* condition, we did not provide participants with information about a hypothesis. According to the facial feedback hypothesis, participants should report feeling (a) happier after posing happy vs. angry and neutral expressions, and (b) angrier after posing angry vs. happy and neutral expressions. According to the demand-characteristics account, this effect of pose on self-reported emotion will be largest in the positive expectation condition and non-existent in the null expectation condition. We pre-registered predictions consistent with this demand-characteristics account. Nevertheless, it is possible that both facial feedback and demand effects co-occur. Given that the null expectation condition tests a context where demand characteristics work *against* potential facial feedback effects, we pre-registered that we would consider a significant pose effect in this condition to be a critical test against the claim that observed facial feedback effects are solely driven by demand characteristics.

Method

Based on an a-priori power analysis, two hundred and fifty university students completed the experiment in exchange for course credit (68% female, 32% male, M_{age} = 18.84, SD_{age} = 2.37; see Supplemental Materials for details regarding power analysis). Upon arriving at the lab, we told participants the experiment examines how muscle movements and fluctuations in concentration/mood influence galvanic skin response and that they were assigned to move the muscles in their face. We provided this cover story to ensure that participants in all demand conditions were aware of our interest in facial movements. Next, two electrodes connected to a visibly active Biopac MP36 were attached to participants' left hands. (No physiological

recordings were actually taken.) Afterwards, we informed participants that they would complete 6 timed facial movements and answer questions about their concentration and mood.

We randomly assigned participants to one of three demand conditions. Participants in the control condition began the facial expression poses without further instruction. Participants in the positive expectation and null expectation conditions received the following message from the experimenter (null expectation instructions in brackets):

> "Researchers know that facial expressions influence skin responses. However, they are not sure why. Some researchers think that this happens because posing a facial expression of emotion causes you to feel that emotion. That is, smiling may make your happier and scowling may make you angrier. We believe this is [*not*] true. We are running this study to prove our hypothesis that posing facial expressions [*does not*] cause you to experience that emotion."

To ensure comprehension, experimenters were permitted to use a visual aid to explain the expected effect (see Figure S1 in Supplemental Materials). After leaving the participant alone in the lab to complete the study, the first page of instructions repeated the message regarding the expected effect.

All participants completed two blocks of poses, each of which included instructions to pose happy, angry, and neutral facial expressions in randomized order. For happy facial expression poses, we instructed participants to move the corner of their lips towards their ears, elevating their cheeks (Dimberg & Söderkvist, 2011). For angry facial expression poses, we instructed participants to move their eyebrows down towards their nose. For neutral facial expression poses, we instructed participants to maintain a blank expression. We asked

participants to hold each pose for 5 seconds in conjunction with an on-screen timer. After each pose, participants completed a modified Discrete Emotions Questionnaire (Harmon-Jones, Bastian, & Harmon-Jones, 2016), which included three items measuring happiness (overall α = .88), three items measuring anger (overall α = .85), and three filler items measuring fear (0 = "not at all" to 6 = "an extreme amount"). Cronbach alphas scores were calculated separately for each pose and block combination and then averaged. Participants also reported how difficult they found the posing task and how concentrated they felt. We report exploratory analyses related to self-reported difficulty in the supplementary materials (Table S3), but these exploratory analyses do not change the conclusions reported in the main text.

After the two blocks of poses, participants completed a funnel debriefing. In the funnel debriefing, the experimenter asked participants questions about what they believed the experiment hypothesis was and whether they believed their facial expressions would influence their emotions. Based on these funnel debriefings, experimenters provided ratings of the degree to which the participant (a) was aware we were testing facial feedback effects (0 = "not at all aware" to 4 = "completely aware"), and (b) believed facial feedback would influence their emotions (0 = "very convinced they would *not* influence their emotions" to 4 "very convinced they would influence their emotions").

Taken together, the experiment featured a 3 (Pose: happy, angry, neutral) x 2 (Block: first or second) x 3 (Demand: positive expectation, null expectation, control) mixed design, with demand characteristics manipulated between-subjects. The University of Tennessee-Knoxville Institutional Review Board approved the experimental protocol.

Results

Demand-Characteristics Manipulation Checks

To examine the efficacy of the demand characteristics manipulation, we first tested (a) whether participants were more aware that we were testing facial feedback effects in the positive and null expectation conditions (vs. the control condition) and (b) whether participants were more convinced that facial feedback influences their emotions in the positive expectation condition (and least convinced in the null expectation condition). Experimenters provided ratings of participants' awareness and beliefs based on funnel debriefings. Notably, these ratings should be interpreted with some caution since experimenters were aware of the participants' condition, which could have influenced their interview behavior and interpretation of the participants' responses. Furthermore, participants may have been unable or unwilling to express their beliefs during the funnel debriefing (Blackhart, Brown, Clark, Pierce, & Shell, 2012). With these caveats in mind, these ratings provide an opportunity to examine whether the demand characteristics manipulation was effective.

If the demand characteristics manipulation was effective, participants should have been more aware we were testing facial feedback effects in the positive and null expectation conditions (vs. the control condition). Results from a one-way ANOVA confirmed that awareness ratings varied by expectation condition, $F(2, 247) = 31.07$, $p < .001$. Follow-up pairwise comparisons indicated that, compared to participants in the control condition ($M = 1.39$, $SD = 1.22$), participants were more aware we were testing facial feedback effects in the positive ($M = 2.93$, $SD = 1.21$), $t(247) = 7.66$, $p < .001$, M_{diff} 95% CI [1.14, 1.94], and null expectation conditions ($M = 2.52$, $SD = 1.38$), $t(247) = 5.62$, $p < .001$, M_{diff} 95% CI [0.73, 1.52]. Results also

indicated that participants were slightly more aware we were testing facial feedback effects in the positive vs. null expectation conditions, $t(247) = 2.12$, $p = .04$, M_{diff} 95% CI [0.03, 0.80].

We also examined participants' beliefs about facial feedback effects. Conceptually, demand characteristics are separable from beliefs about an effect. It is generally assumed that demand characteristics influence behavior because participants want to please the experimenter —not because they believe the experimenter's hypothesis is valid (Orne, 1962). Nevertheless, we speculated that many participants may come to believe the experimenter's stated beliefs. Results from a one-way ANOVA confirmed that participants' belief in the facial feedback hypothesis varied by expectation condition, $F(2, 231) = 27.35$, $p < .001$[1]. Follow-up pairwise comparisons indicated that participants were more convinced that facial feedback influences emotion in the positive expectation ($M = 2.84$, $SD = 1.22$) vs. control ($M = 2.31$, $SD = 1.38$), $t(231) = 2.37$, $p = .02$, M_{diff} 95% CI [0.09, 0.96], and null expectation conditions ($M = 1.34$, $SD = 1.39$), $t(231) = 7.31$, $p < .001$, M_{diff} 95% CI [1.09, 1.90]. Participants were also more convinced that facial feedback influences emotion in the control vs. null expectation condition, $t(231) = 4.40$, $p < .001$, M_{diff} 95% CI [0.54, 1.41]. Taken together, these results suggest that the demand characteristics manipulation was effective.

Self-Reported Emotional Experience

Due to the nested structure of the data (emotion reports nested within participants), we examined emotion reports using linear mixed effects modeling with Pose, Block, and Demand entered as factors, all higher-order interactions, and random-intercepts for each participant. The primary analysis of interest was the pose by expectation interaction but results from the full

[1] The facial feedback belief *F*-test has smaller denominator degrees of freedom than the facial feedback awareness *F*-test because experimenters were originally unsure if they had to complete the question in the control condition.

model are reported in the Supplemental Materials (Tables S1 and S2). *F*-values are from ANOVAs with Satterthwaite method degrees of freedom.

To quantify the size of the pose effect in the positive expectation, control, and null expectations conditions, standardized mean difference scores were computed (Cohen's d_{rm}; Cohen, 1988). For ease of comprehension, self-reported emotional experience was averaged across block and effect sizes were calculated as the mean difference between (a) the emotion-congruent pose and (b) the average of the emotion-incongruent poses. For example, to calculate the size of the effect of pose on happiness, we compared happiness scores between (a) the average of the two happy pose trials, and (b) the average of the two neutral pose and two angry pose trials.

Self-Reported Happiness

Consistent with a demand characteristic account, there was a significant Pose by Demand interaction, $F(4, 1235) = 24.60$, $p < .001$. As depicted in the top panel of Figure 1, the Pose effect was largest in the positive expectation condition, $F(2, 1235) = 191.47$, $p < .001$, $d_{rm} = 1.19$, second largest in the control condition, $F(2, 1235) = 85.53$, $p < .001$, $d_{rm} = 0.92$, and smallest in the null expectation condition $F(2, 1235) = 16.79$, $p < .001$, $d_{rm} = 0.51$. In all three Demand conditions, participants reported more happiness when posing happy vs. neutral and angry expressions (Table 1). These results confirm that the effect of posed facial expressions on happiness is moderated by demand characteristics. However, the finding that the effect of posed facial expressions on happiness is significant in the null expectation condition suggests that there was a facial feedback effect that was not entirely driven by demand characteristics.

Self-Reported Anger

Consistent with a demand characteristic account, there was a significant Pose by Demand interaction, $F(4, 1235) = 17.55$, $p < .001$. As depicted in the bottom panel of Figure 1, the Pose effect was largest in the positive expectation condition, $F(2, 1235) = 116.86$, $p < .001$, $d_{rm} = 0.75$, second largest in the control condition, $F(2, 1235) = 54.42$, $p < .001$, $d_{rm} = 0.69$, and smallest in the null expectation condition $F(2, 1235) = 8.89$, $p < .001$, $d_{rm} = 0.34$. In all three Demand conditions, participants reported more anger when posing angry vs. neutral and happy expressions (Table 1). These results confirm that the effect of posed facial expressions on anger is moderated by demand characteristics. However, finding that the effect of posed facial expressions on anger is significant in the null expectation condition indicates that there was a facial feedback effect that was not entirely driven by demand characteristics.

Table 1

Mean differences in self-reported happiness and anger after posing happy vs. neutral vs. angry facial expressions in the positive expectation, control, and null expectation conditions.

		Positive expectation		Control		Null expectation	
Emotion	**Pose contrast**	***t***	***M*$_{diff}$ 95% CI**	***t***	***M*$_{diff}$ 95% CI**	***t***	***M*$_{diff}$ 95% CI**
Happiness	Happy – Neutral	16.33***	[1.12, 1.43]	10.82***	[0.73, 1.05]	4.25***	[0.18, 0.48]
	Happy – Angry	17.50***	[1.21, 1.52]	11.77***	[0.81, 1.13]	5.54***	[0.28, 0.58]
	Anger – Neutral	-1.17	[-0.24, 0.06]	-0.95	[-0.24, 0.08]	-1.29	[-0.25, 0.05]
Anger	Happy – Neutral	-2.58**	[-0.26, -0.04]	0.25	[-0.10, 0.13]	1.15	[-0.05, 0.18]
	Happy – Angry	-14.34***	[-0.93, -0.71]	-8.91***	[-0.66, -0.42]	-2.94**	[-0.28, -0.06]
	Anger – Neutral	11.76***	[0.56, 0.78]	9.16***	[0.43, 0.67]	4.09***	[0.12, 0.34]

Note: *All pairwise comparison tests have 1,235 degrees of freedom. Boxes with grey fill represent contrasts that the demand characteristics account predicts will be null. Boxes with dotted outline represent contrasts that the facial feedback hypothesis predicts will be null. p-values are uncorrected for multiple comparisons.*

**$p < .01$; ***$p < .001$

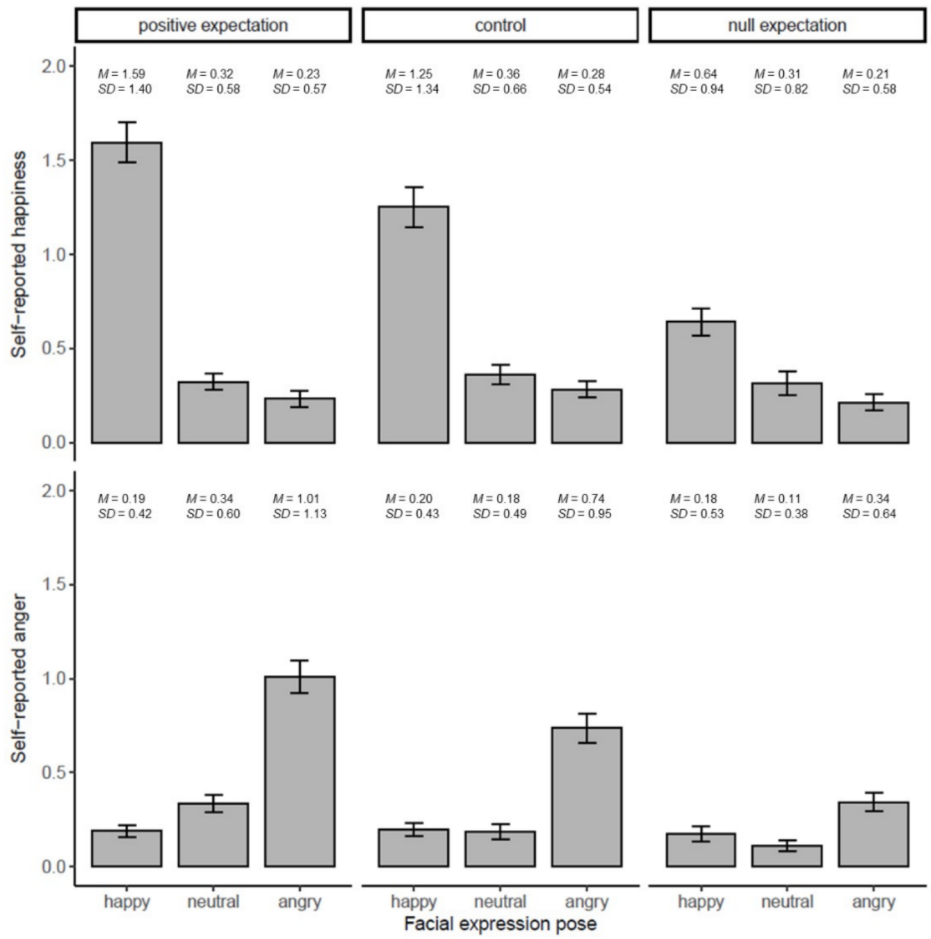

Fig. 1.

Self-reported happiness and anger after participants posed happy, neutral, and angry facial expressions in the positive expectation, control, and null expectation conditions. Bar heights represent means and error bars represent one standard error. Descriptive statistics are reported above the bars.

Discussion

Consistent with theories that posit emotional experience is partially built off afferent signals from the peripheral nervous system (for a review, see Moors, 2009), previous research suggests that posing facial expressions of emotion can influence emotional experience (Coles, Larsen, & Lench, 2019; Coles, March, et al., 2019). However, a reoccurring concern is that these effects are driven by demand characteristics (Buck, 1980; Dimberg & Söderkvist, 2011; Schimmack & Chen, 2017). Our results provide the first experimental evidence that *both* demand characteristics and facial feedback effects can occur in parallel. Demand characteristics moderated the effects of posed facial expressions on feelings of happiness and anger. However, contrary to our own predictions, smiles and frowns influenced emotional experience even when an experimenter told participants they expected it would not. This is notable given that this is a condition where demand characteristics work *against* potential facial feedback effects. These results underscore the potentially large role of demand characteristics in facial feedback research but suggest that demand characteristics do not fully account for the observed effects.

Observing that posed facial expressions influence emotional experience is consistent with theories that posit emotional experience is partially built off afferent signals from the peripheral nervous system (for a review, see Moors, 2009). Nevertheless, it is possible that the effects of facial poses on emotion reports are more indirect. For example, some researchers have suggested that facial feedback can impact other emotion-related cognitive processes, such as memory or the processing of emotional information (Niedenthal, 2007). Others have suggested that facial feedback may initiate emotion-specific changes in the central nervous system (Coan, Allen, & Harmon-Jones, 2001) and/or other components of the peripheral nervous system (Levenson, Ekman, & Friesen, 1990). The next important step for affective scientists is to model the

cognitive, neural, and social processes that underlie facial feedback effects, all while keeping at bay the potentially large influence of demand characteristics.

Author Contributions

N. A. Coles developed the study concept. N. A. Coles, B. Frohlich, and L. Gaertner developed the methodology. N. A. Coles and B. Frohlich programmed the study and supervised data collection. N. A. Coles and L. Gaertner analyzed and visualized the data. N. A. Coles drafted the manuscript. J. T. Larsen, L. Gaertner, and B. Frohlich provided critical feedback on the manuscript. All authors contributed to revisions and approved the final version of the manuscript for submission.

Acknowledgments

Thank you to Leo Tiokhin, Daniël Lakens, and Ruben Arslan for comments on an earlier version of this manuscript.

References

Blackhart, G. C., Brown, K. E., Clark, T., Pierce, D. L., & Shell, K. (2012). Assessing the adequacy of postexperimental inquiries in deception research and the factors that promote participant honesty. *Behavior Research Methods*, *44*(1), 24–40.

Buck, R. (1980). Nonverbal behavior and the theory of emotion: the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *38*(5), 811–824.

Coan, J. A., Allen, J. J. B., & Harmon-Jones, E. (2001). Voluntary facial expression and hemispheric asymmetry over the frontal cortex. *Psychophysiology*, *38*(6), 912–925.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Vol. 2nd). New York: Lawrence Erlbaum Associates.

Coles, N. A., Larsen, J. T., Kuribayashi, J., & Kuelz, A. (2019). Does blocking facial feedback via botulinum toxin injections decrease depression? A critical review and meta-analysis. *Emotion Review*, *11*(4), 294–309.

Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin*, *145*(6), 610–651.

Coles, N. A., March, D. S., Marmolejo-Ramos, F., Banaruee, H., Butcher, N., Cavallet, M., … Gorbunova, E. (2019). A multi-lab test of the facial feedback hypothesis by The Many Smiles Collaboration.

Dimberg, U., & Söderkvist, S. (2011). The voluntary facial action technique: A method to test the facial feedback hypothesis. *Journal of Nonverbal Behavior*, *35*(1), 17–33.

Harmon-Jones, C., Bastian, B., & Harmon-Jones, E. (2016). The discrete emotions questionnaire: A new tool for measuring state self-reported emotions. *PloS One*, *11*(8), e0159915.

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, *35*(1), 7–31.

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, *27*(4), 363–384.

Moors, A. (2009). Theories of emotion causation: A review. *Cognition and Emotion*, *23*(4), 625–662.

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*(5827), 1002–1005.

Noah, T., & Mayo, R. (2018). When both the original study and its failed replication are correct: Feeling observed eliminates the facial-feedback effect. *Journal of Personality & Social Psychology*, *114*(5), 1–26.

Orne, M. T. (1962). On the social psychology of the psychological experiment. *American Psychologist*, *17*(11), 776–783.

Schimmack, U., & Chen, Y. (2017). The power of the pen paradigm: A replicability analysis. Replicability-Index.

Strack, F. (2016). Reflection on the smiling registered replication report. *Perspectives on Psychological Science*, *11*(6), 929–930.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality*

and Social Psychology, *54*(5), 768–777.

Wagenmakers, E. J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., … Zwaan, R. A. (2016). Registered replication report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, *11*(6), 917–928.

Supplemental Materials

Pre-Registration and Power Analysis

We hypothesized that participants would report higher levels of happiness (a) when posing happy vs. neutral and angry expressions and (b) when in the positive expectation vs. null expectation and control conditions. More importantly, we hypothesized a Pose by Demand interaction, wherein the effects of pose would be largest in the positive expectation condition, second largest in the control condition, and smallest (and non-significant) in the null expectation condition. We pre-registered the same expectations about the effects of posing angry vs. neutral and happy expressions on self-reported anger. We hypothesized that participants would be more aware we were testing facial feedback effects in the positive and null expectation conditions (vs. the control condition). We also hypothesized that participants would believe in the facial feedback hypothesis the most in the positive expectation condition and the least in the null expectation condition.

To simplify our power simulation, we ignored the nested nature of our design and used the overall difference in self-reported happiness in the happy vs. neutral trials as our dependent variable. This means we simulated a non-nested single-factor design, with demand characteristics as the independent variable. This simplified approach allowed us to avoid making several difficult assumptions needed for multi-level power simulations (Lane & Hennes, 2018), but likely led to highly conservative estimates of power. For our power simulation, we generated 1,000 iterations where we drew 100 random samples from a normal distribution for each of the Demand conditions: positive expectation $N(0.75, 1.00)$, control $N(0.25, 1.00)$, and null expectation $N(0.00, 1.00)$. We then calculated the proportion of iterations that detected (a) an

effect of pose, (b) a pose by demand interaction, and (c) significant pose effects within each Demand condition. Results indicated that 300 participants would provide approximately 99% power to detect our expected main effect for Pose and our expected Pose by Demand interaction (using $\alpha = .05$). Furthermore, results indicated that this sample size would provide good power to detect small Pose effects within each level of Demand (power $\approx .72$ for $d = .25$). We pre-registered that we would end data collection when we either (a) collected 300 participants or (b) reached the end of the Spring 2020 data collection calendar. We used the latter stopping rule.

Visual Aids

Early in the study, experimenters noted that a visual aid may help them more effectively deliver the demand characteristics manipulation. Consequently, experimenters were permitted to use the visual aids in Supplemental Figure 1 to explain the expected effect to participants.

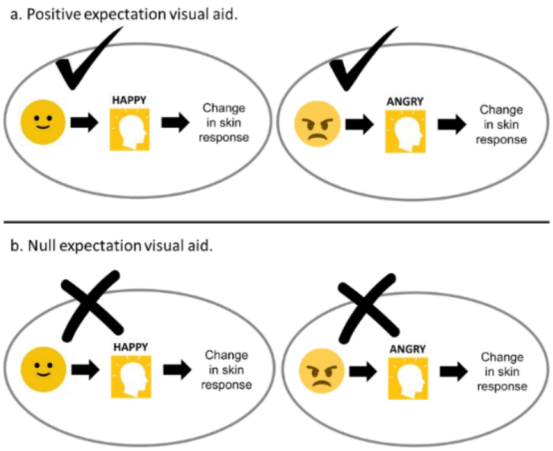

Fig. S1.

Experimenters were permitted to use the following visual aids to explain the expected effects in the positive expectation (a) and null expectation (b) conditions to participants.

Supplemental Analyses

Self-reported emotional experience.

Due to the nested structure of the data (emotion reports nested within participants), emotion reports were examined using linear mixed effects modeling with Pose, Block, and Demand entered as factors, all higher-order interactions, and random-intercepts for each participant. In the main text, we only reported the critical Pose by Demand interactions. However, effects from the full model examining happiness (Supplemental Table 1) and anger (Supplemental Table 2) are below

Table S1

Results from a linear mixed effects model with Pose, Block, and Demand entered as factors, all higher-order interactions, random-intercepts for each participant, and reported happiness as the dependent variable.

Omnibus test	Decomposition	Result
Pose	-	$F(2, 1235) = 246.03$, $p < .001$
	Happy – Neutral	$t(1235) = 18.16$, $p < .001$, M_{diff} 95% CI [0.74, 0.92]
	Happy – Angry	$t(1235) = 20.11$, $p < .001$, M_{diff} 95% CI [0.83, 1.01]
	Anger – Neutral	$t(1235) = -1.96$, $p = .05$, M_{diff} 95% CI [-0.18, 0.18]
Demand	-	$F(2, 247) = 7.34$, $p < .001$
	Positive vs. Control	$t(247) = .93$, $p = .35$, M_{diff} 95% CI [-0.09, 0.26]
	Positive vs. Null	$t(247) = 3.70$, $p < .001$, M_{diff} CI [0.15, 0.50]
	Control vs. Null	$t(247) = 2.66$, $p = .008$, M_{diff} 95% CI [0.06, 0.42]
Block	-	$F(1, 1235) = 19.38$, $p < .001$
	Block 1- Block 2	$t(1235) = 4.40$, $p < .001$, M_{diff} 95% CI [0.09, 0.24]
Pose by Demand	-	$F(4, 1235) = 24.60$, $p < .001$
	See main text	*See main text*
Pose by Block	-	$F(2, 1235) = 4.24$, $p = .01$
	Block 1 Pose Effect	$F(2, 1235) = 157.04$, $p < .001$
	Block 2 Pose Effect	$F(2, 1235) = 93.23$, $p < .001$
Demand by Block	-	$F(2, 1235) = 0.40$, $p = .67$
Pose by Demand by Block	-	$F(4, 1235) = 0.15$, $p = .96$

Note: *Omnibus test F-values are from ANOVAs with Satterthwaite method degrees of freedom. For main effects, decompositions are least squared pairwise difference tests. For interactions, decompositions are ANOVAs split by the second term in the interaction. Grey boxes are tests that we did not have a-priori hypothesis for. p-values are uncorrected for multiple comparisons.*

Table S2

Results from a linear mixed effects model with Pose, Block, and Demand entered as factors, all higher-order interactions, random-intercepts for each participant, and reported anger as the dependent variable.

Omnibus test	Decomposition	Result
Pose	-	$F(2, 1235) = 146.33$, $p < .001$
	Happy – Neutral	$t(1235) = 0.67$, $p = .51$, M_{diff} 95% CI [-0.09, 0.04]
	Happy – Angry	$t(1235) = -15.14$, $p < .001$, M_{diff} 95% CI [-0.57, -0.44]
	Anger – Neutral	$t(1235) = 14.47$, $p < .001$, M_{diff} 95% CI [0.42, 0.55]
Demand	-	$F(2, 247) = 9.50$, $p < .001$
	Positive vs. Control	$t(247) = 1.94$, $p = .05$, M_{diff} 95% CI [0.002, 0.48]
	Positive vs. Null	$t(247) = 4.35$, $p < .001$, M_{diff} 95% CI [0.17, 0.44]
	Control vs. Null	$t(247) = 2.29$, $p = .02$, M_{diff} 95% CI [0.02, 0.30]
Block	-	$F(1, 1235) = 8.30$, $p = .004$
	Block 1- Block 2	$t(1235) = -2.88$, $p < .001$, M_{diff} 95% CI [-0.13, -0.23]
Pose by Demand	-	$F(4, 1235) = 17.55$, $p < .001$
	See main text	*See main text*
Pose by Block	-	$F(2, 1235) = 1.13$, $p = .32$
Demand by Block	-	$F(2, 1235) = 0.48$, $p = .62$
Pose by Demand by Block	-	$F(4, 1235) = 1.28$, $p = .28$

Note: *Omnibus test F-values are from ANOVAs with Satterthwaite method degrees of freedom. For main effects, decompositions are least squared pairwise difference tests. For interactions, decompositions are ANOVAs split by the second term in the interaction. Grey boxes are tests that we did not have a-priori hypothesis for. p-values are uncorrected for multiple comparisons.*

Self-reported difficulty.

One possibility is that the effects of Pose on emotion reports is caused by differences in the difficulties of the pose (rated from 1 = "Not at all" to 7 = "An extreme amount"). For example, perhaps participants report feeling angrier after posing angry vs. happy and neutral facial expressions because it is most difficult to pose angry expressions.

To examine whether self-reported difficulty varied across conditions, we examined difficulty reports using linear mixed effects modeling with Pose, Block, and Expectation entered as factors, all higher-order interactions, and random-intercepts for each participant. Results indicated that there was a main effect of Pose on self-reported difficulty, $F(2, 1235) = 34.36$, $p < .001$. Follow-up least-squares pairwise comparisons indicated that participants rated the angry posing task ($M = 1.50$, $SD = 0.94$) as more difficult than the happy ($M = 1.28$, $SD = 0.66$), $t(1235) = 5.43$, $p < .001$, M_{diff} CI [0.14, 0.30], and neutral posing task ($M = 1.17$, $SD = 0.62$), $t(1235) = 8.14$, $p < .001$, M_{diff} CI [0.25, 0.41]. Participants also rated the happy facial expression posing task as more difficult than the neutral posing task, $t(1235) = 2.71$, $p = .007$, M_{diff} CI [0.03, 0.19].

Results also indicated that there was a main effect of Demand. Descriptively, mean difficulty ratings were the largest in the control condition ($M = 1.42$, $SD = 0.84$), second largest in the positive expectation condition ($M = 1.34$, $SD = 0.82$), and smallest in the null expectation condition ($M = 1.20$, $SD = 0.61$). Pairwise comparisons indicated that participants found the posing tasks to be more difficult in the control vs. negative expectation condition, $t(247) = 3.09$, $p = .002$, M_{diff} CI [0.08, 0.37]. There was marginally significant evidence that participants found the posing task more difficult in the positive vs. null expectation condition, $t(247) = 1.94$, $p = .05$, M_{diff} CI [0.00, 0.27]. No other higher-order interactions were statistically significant.

We subsequently re-examined the main effects of Pose, main effects of Demand, and Pose by Demand interaction effects on happiness and anger controlling for difficulty. Once again, the hypothesized Pose by Demand interactions were significant, $F_{happiness}$ (4, 1239.28) = 24.84, F_{anger} (4, 1239.28) = 17.10, $ps < .001$. In all three Demand conditions, participants reported (a) more happiness when posing happy vs. neutral and angry expressions and (b) more anger when posing angry vs. neutral and happy expressions (Supplemental Table 3).

Table S3

Mean differences in self-reported happiness and anger after posing happy vs. neutral vs. angry facial expressions in the positive expectation, control, and null expectation conditions in a model controlling for the reported difficulty of the pose.

		Positive expectation		Control		Null expectation	
Emotion	**Pose contrast**	***t***	***M*$_{diff}$ 95% CI**	***t***	***M*$_{diff}$ 95% CI**	***t***	***M*$_{diff}$ 95% CI**
Happiness	Happy – Neutral	16.56***	[1.13, 1.44]	11.03***	[0.74, 1.07]	4.36***	[0.18, 0.49]
	Happy – Angry	17.25***	[1.19, 1.49]	11.46***	[0.78, 1.10]	5.37***	[0.26, 0.57]
	Anger – Neutral	-0.73	[-0.21, 0.10]	-0.46	[-0.20, 0.12]	-1.01	[-0.23, 0.07]
Anger	Happy – Neutral	-3.04**	[-0.28, -0.06]	-0.16	[-0.13, 0.11]	0.95	[-0.06, 0.16]
	Happy – Angry	-13.82***	[-0.89, -0.67]	-8.19***	[-0.60, -0.37]	-2.49*	[-0.25, -0.03]
	Anger – Neutral	10.73***	[0.50, 0.72]	7.99***	[0.36, 0.59]	3.43***	[0.08, 0.30]

Note: *Boxes with grey fill represent contrasts that the demand characteristics account predicts will be null. Boxes with dotted outline represent contrasts that the facial feedback hypothesis predicts will be null. p-values are uncorrected for multiple comparisons.*

*p < .05; **p < .01; ***p < .001, uncorrected for multiple comparisons.

Belief in the facial feedback hypothesis.

Conceptually, demand characteristics are separable from beliefs about an effect. It is generally assumed that demand characteristics influence behavior because participants want to please the experimenter—not because they believe the experimenter's hypothesis is valid (Orne, 1962). As reported in the main text, results from a one-way ANOVA confirmed that participants' belief in the facial feedback hypothesis varied by expectation condition, $F(2, 231) = 27.35$, $p < .001$. As a follow-up exploratory analysis, we examined whether belief in the facial feedback hypothesis moderates facial feedback effects. To do so, we modeled happiness and anger reports using linear mixed effects models with Pose and Block entered as factors, Belief entered as a continuous predictor, a Pose by Belief interaction, and random intercepts for each participant. For both happiness and anger reports, the effects of Pose were moderated by Belief in the facial feedback hypothesis, $F_{happiness}(2, 1165) = 74.51$, $p < .001$, $F_{anger}(2, 1165) = 49.25$, $p < .001$.

To decompose the Pose by Belief interaction, we re-ran the models with Belief as a mean-split factor and re-examined the effects of Pose within each level of Belief. For both happiness and anger reports, the effects of Pose was larger in participants who had high, $F_{happiness}(2, 1165) = 269.00$, $p < .001$, $F_{anger}(2, 1165) = 155.56$, $p < .001$, vs. low, $F_{happiness}(2, 1165) = 25.20$, $p < .001$, $F_{anger}(2, 1165) = 10.53$, $p < .001$, belief in the facial feedback hypothesis.