Supplementary Materials

Main Psychophysiological Experiment

Stimuli validation and modeling. We assessed the temporal dynamics of the FACS-coded video stimuli to be used in the main study. Our goal was to ensure that the parallel action units (AUs) associated with the two facial muscles to be measured – i.e., AU4 (corrugator supercilii) and AU12 (zygomaticus major) – had similarly standardized time-courses of activation and intensity in the stimuli.

***Method*.** To accomplish this, we used the Computer Expression Recognition Toolbox (CERT; Littlewort et al., 2011), an advanced automated FACS-coding software program. CERT outputs have been shown to significantly correlate with zygomaticus and corrugator fEMG measures, thus making it a valuable tool for time-course comparison. CERT provides continuous relative activations for 19 different AUs, as well as summary measures for six basic emotions. All video stimuli to be used in the main study (2500ms videos of four different models that were FACS-coded within the MMI Database for happiness and anger) were uploaded and processed for frame-by-frame coding. AUs 4 (corrugator supercilii) and 12 (zygomaticus major) were the primary focus, since those AUs corresponded to the muscles measured in the main study. Continuous support vector machine (SVM) outputs for these AUs were then aggregated and analyzed via repeated measures linear mixed-effects (LME) modeling – the same method used to analyze the participants' fEMG data.

***Results and discussion*.** CERT data were analyzed via a Muscle (2: AU4, AU12) × Valence (2: angry, happy) × Time (64: number of processed frames in 2500ms video). We observed main effects of Valence, $F(1,12)=12.0$, $p=.01$, and Time, $F(63,368)=1.69$, $p<.01$. The Valence main effect was due to greater SVM output values for the zygomaticus (which resulted in greater absolute values for AU12, in general), and this broad activation was greater in happy versus angry videos. The Time main effect demonstrated that the video stimuli had a dynamic change in general activation of both AUs over time. These main effects were qualified by three interactions:

(1) A Muscle × Valence interaction, $F(1,12)=23.5$, $p<.001$, which revealed that each facial muscle was more active with the congruent emotion – i.e., as would be expected, more AU4 (corrugator) activity during angry videos and more AU12 (zygomaticus) activity during happy videos.

(2) A Time × Valence interaction, $F(63,368)=1.92$, $p<.001$, which demonstrated a greater separation in activation for happy videos versus angry videos. This is not surprising, considering that AU12 (zygomaticus) outputted greater overall SVM absolute values in CERT, and this muscle was recruited more so for happy videos and less so for angry videos.

(3) A Muscle × Time × Valence interaction, $F(63,368)=2.98$, $p<.001$, which verified that each muscle activated differently over time in accordance with the congruent emotion. More specifically, AU4 (corrugator) had a clear arc of activation over the 2500ms time-course for angry videos but is relatively flat for happy videos, and vice versa for AU12 (zygomaticus).

In short, these effects and interactions are what would be expected from properly FACS-coded stimuli. The above-outlined analyses further standardize these videos for timing and activation on the muscles-of-interest for the main study.

Equipment and software. Stimuli were presented on a 17-inch Dell flat-screen monitor with an Intel® Core™ 2 Duo CPU with 4 GB of RAM and a 32-bit operating system (running

Windows 7 Professional, © Microsoft Co.). The experimental task was presented using E-Prime 2.0 (Psychology Software Tools, Pennsylvania, USA). EMG signals were recorded with a model MP150CE/EMG2-T BioNomadix Wireless data acquisition system (Biopac Systems Inc., California, USA), AcqKnowledge version 4.1.1 BIOPAC recording software, and BIOPAC EL504 disposable electrodes. SCR was also recorded with the MP150CE model on a SCR100 channel. All channel sampling rates were 2000 Hz. Physiological data were analyzed using MindWare EMG package version 2.52 (MindWare Technologies Ltd., Ohio, USA) and coding scripts in Matlab R2011a (MathWorks Inc, Massachusetts, USA) to calculate, clean, and standardize physio signals (data cleaning and standardizing are detailed in the "Data processing" section).

Participants. Note that our sample of participants for the main study included 82% females (similar to the follow-up experiment, which included 77% females; see "Methods, results, and discussion" section for the "Follow-up Rating Experiment" in the Supplementary Materials).While gender differences do seem to be present in spontaneous facial mimicry (as they are in most, if not all, social-cognitive and psychophysiological phenomena), past research has shown that both genders generally perceive and respond to facial expressions similarly, especially with rapid facial responses via fEMG, but that females usually show a stronger pattern of reactivity (e.g., Dimberg & Lundquist, 1990; Thunberg & Dimberg, 2000).

Therefore, it is quite common in many mimicry studies to use most to all female subjects (similar to the samples for both of our experiments) since they tend to show stronger effects than males, that are usually in the same general direction (e.g., Hermans, Putman, & Van Honk, 2006; Hess & Bourgeois, 2010; Korb, Grandjean, & Scherer, 2010). For this reason, we are confident that our findings represent an accurate portrayal of the population, and would generalize beyond our sample of participants.

Materials. Our manipulation of perceiver-power required participants to recall and write about a past experience, based on random assignment to either the high-power (HP), low-power (LP), or control condition. After signing consent forms and answering some basic demographic questions, the participants received the instructions for the writing prime, which have been shown to reliably induce different levels of subjective power (Galinsky, Gruenfeld, & Magee, 2003). Instructions for the LP condition were as follows: *On the sheet of paper you were given, please recall a particular incident in which another individual (or individuals) had power over you. By power, we mean a situation in which another person (or persons) controlled your ability to get something you wanted, or were in a position to evaluate you. Write about and describe the situation – i.e., what happened, how it made you feel, etc.*

The instructions for the HP condition were as follows: *On the sheet of paper you were given, please recall a particular incident in which you had power over another individual (or individuals). By power, we mean a situation in which you controlled another person (or persons) ability to get something they wanted, or were in a position to evaluate them. Write about and describe the situation – i.e., what happened, how it made you feel, etc.*

The instructions for the control condition were as follows: *On the sheet of paper you were given, please recall the timeline of events that occurred for you yesterday. By events, we mean classes, meetings, appointments, etc. Write about and briefly describe each event – i.e., what happened, specific times, etc.*

We used the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) to assess self-reported mood after our perceiver-power manipulation. This was to ensure that systematic mood differences across conditions did not confound the fEMG effects. The PANAS has been found to be high in both validity and reliability when gauging positive and negative mood, as supported by cross-sectional and correlational studies in large, non-clinical samples (e.g., Crawford & Henry, 2004; Thompson, 2007).

Data processing. During each trial, the 3000ms fixation period was used as a baseline to which each of the subsequent 5000ms video trials was compared. In order to ensure that the data were properly cleaned and filtered, means and standard deviations were first calculated on each subject's raw dataset, and all trial points outside the ± 3SD range from the mean signal for that subject were removed (this resulted in the exclusion of approximately 0.5% of all trial points). Next, a new mean and standard deviation for each subject was calculated based on the remaining valid trials, and all signals were z-scored. Finally, the average of the z-scored baseline for each trial was subtracted from each 500ms z-scored trial point, and this process yielded a time-course of baseline-corrected, z-scored signals across 500ms intervals for all video trials (standardized by subject).

Linear mixed-effects (LME) modeling via restricted maximum likelihood (REML) was used for all repeated measures analyses to reduce information loss when evaluating our large, unbalanced datasets (Judd, Westfall, & Kenny, 2012) after signal standardization. All models were built using the *lme4* package in *R* (Bates, 2005) with a maximal random effects structure, after which a top-down approach and stepwise likelihood-ratio χ^2 significance tests were used to optimize model fit (West, Welch, & Galecki, 2007). All DFs were calculated using the Kenward-Roger method (*CERT*: AICc=2089, BIC=3356; *zygomaticus*: AICc=-685, BIC=141; *corrugator*: AICc=-1980, BIC=-1086).

Cross-muscle analyses. We were interested in comparing the development of perceivers' expressions over time, in order to gauge potential timing differences between the corrugator and zygomaticus (especially for situations where both muscles are utilized; e.g., HP perceivers' smiling *and* frowning toward angry HP targets). To investigate this, we used two different methods: (*i*) zero-lag cross-correlations between the muscles to assess simultaneous co-activation, and (*ii*) a new, expanded LME model that included Time (10: 500ms-5000ms, in 500ms time-windows) and Muscle (2: corrugator, zygomaticus) as factors and z-scored change in activation from baseline as the outcome variable.

Cross-correlations. To assess whether or not the muscles simultaneously co-activated, we computed zero-lag cross-correlations between the corrugator and zygomaticus at every factor level, a method that determines the statistical association between waveforms over time (Orfanidis, 1988). All fEMG data for each participant was thus divided into each factor level for our different conditions – 3 (Perceiver-Power: HP, LP, control) × 2 (Target-Power: HP, LP) × 2 (Valence: angry, happy) – and cross-correlations were run using the Pearson *r* method. After individual subject cross-correlations were calculated, they were transformed using the Fisher-*Z* formula, and then averaged within each factor level. Next, all *Z* averages were backtransformed into Pearson *r*-values, which were then used to compute 95% confidence intervals and significance tests (see Table S1). This method has been shown to reduce bias when doing average correlation estimation, especially when the sample size is small (Silver & Dunlap, 1987).

Results are displayed in Table S1: Overall, none of the cross-correlations reached the level of significance, while a majority of the correlation coefficients were negative. In short, this seems to show that there is not a systematic pairing between corrugator and zygomaticus activation, across any of the factor levels, within our sample of subjects.

			Average Subject Cross-Correlations (Corrugator vs. Zygomaticus)
Power (Perceiver) Condition	Power (Target) Condition	Expression (Target) Valence	*Pearson's r (backtransformed Fisher Z values) [95% CI]*
Control (n = 18)	High-power (HP)	Anger	0.0721 [-0.4085, 0.5214]
		Happiness	-0.0523 [-0.5068, 0.4249]
	Low-power (LP)	Anger	-0.0567 [-0.5101, 0.4213]
		Happiness	-0.0555 [-0.5092, 0.4223]
High-power (HP; n = 19)	High-power (HP)	Anger	-0.1324 [-0.5533, 0.3424]
		Happiness	-0.0548 [-0.4967, 0.4096]
	Low-power (LP)	Anger	-0.0731 [-0.5104, 0.3942]
		Happiness	-0.0759 [-0.5125, 0.3918]
Low-power (LP; n = 18)	High-power (HP)	Anger	0.0710 [-0.4095, 0.5206]
		Happiness	0.0252 [-0.4470, 0.4863]
	Low-power (LP)	Anger	0.2397 [-0.2558, 0.6354]
		Happiness	-0.0917 [-0.5356, 0.3920]

Table S1. Results of cross-correlation analyses: Perceiver-power conditions are divided into each of the three main rows, and columns for target-power and expression valence subset the data to each factor level. Average subject cross-correlations (which represent *r*-values that were backtransformed from Fisher-*Z* averages) are depicted in the last column, along with the corresponding 95% confidence intervals.

Mixed-model for muscle timing. To evaluate whether there were systematic timing differences between the muscles, we built a separate 3 (Perceiver-Power: HP, LP, control) × 2 (Muscle: corrugator, zygomaticus) × 10 (Trial Time Point: 500ms-5000ms) nested LME model, with z-scored change in muscle activity as the outcome variable. Since all fEMG activations were z-scored within-subject, this allows for cross-muscle comparisons (Target-Power and Valence did not interact with time, so they were trimmed from the final model).

Aside from main effects of Time, $F(9,4254)=5.52$, $p<.001$, and Muscle, $F(1,104)=4.08$, $p=.05$, which demonstrated that participants elicited a response over the trial period that was relatively greater in the zygomaticus, we also observed a significant Muscle × Time interaction, $F(9,4254)=8.66$, $p<.001$. Interestingly, participants activated the corrugator much earlier during the trial period (i.e., before 1000ms), $b=.10$, $t=3.05$, $p=.003$, $d=.43$, while the zygomaticus activated later (i.e., 2000-4000ms), $b=.09$, $t=3.32$, $p=.001$, $d=.23$ (see Figure S1).

Linear mixed-effects (LME) modeling via restricted maximum likelihood (REML) was used for all repeated measures analyses to reduce information loss when evaluating our large, unbalanced datasets (Judd, Westfall, & Kenny, 2012) after signal standardization. All models were built using the *lme4* package in *R* (Bates, 2005) with a maximal random effects structure, after which a top-down approach and stepwise likelihood-ratio χ^2 significance tests were used to optimize model fit (West, Welch, & Galecki, 2007). All DFs were calculated using the Kenward-Roger method (*cross-muscle comparison:* AICc=-394, BIC=34).

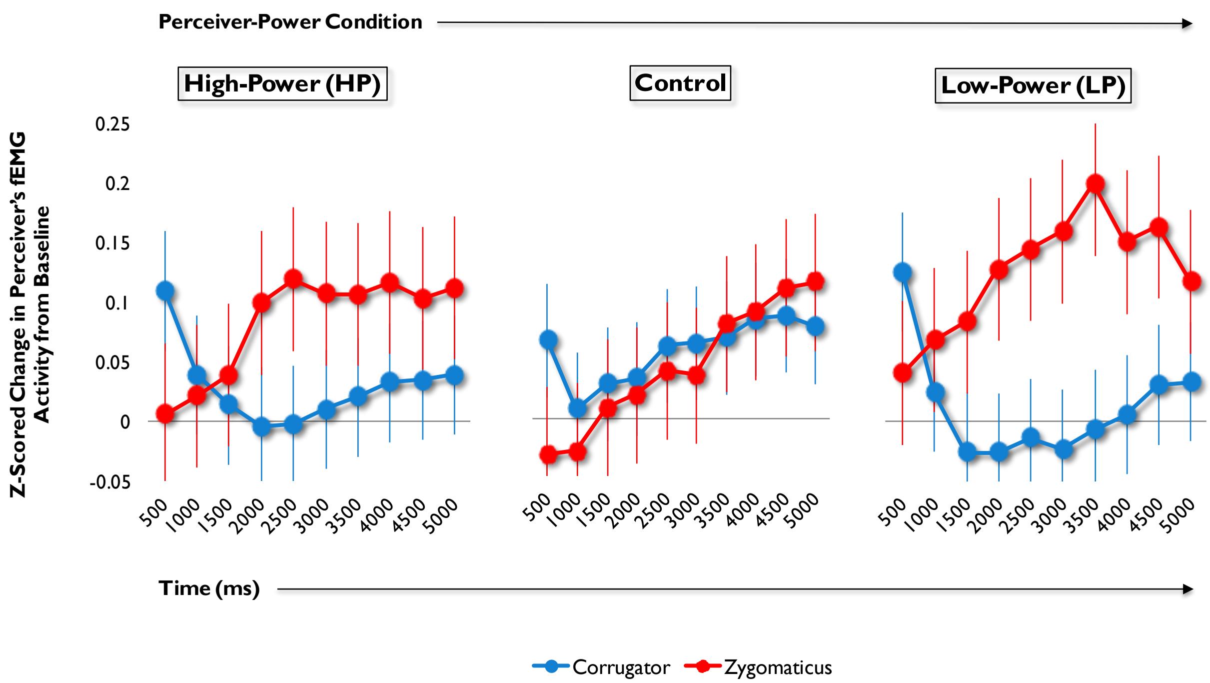

Figure S1. Corrugator and zygomaticus activation over time: For all perceiver-power conditions, participants activated the corrugator (blue line) much earlier during the trial period (i.e., before 1000ms), while the zygomaticus (red line) activated later (i.e., 2000-4000ms). Trial time (in milliseconds) is plotted along the horizontal axes, and z-scored change in muscle activity from baseline is plotted on the vertical axes. Error bars=±1 SEM.

Follow-up Rating Experiment

The follow-up experiment to the main psychophysiological study evaluated whether perceptual differences in the interpretation of facial expressions would arise after our power manipulation. We used a rating experiment that required a new set of subjects to make judgments about different smiles that were exhibited by the models in our video stimuli. While the main study demonstrated that power changes the direct-matching of emotional expressions, this was especially true for smiles. Those results alone, however, are also consistent with the possibility that HP and LP perceivers merely interpret smiles differently. For example, given the multiple meanings of smiles across different social contexts, LP perceivers may view smiles as affiliation signals, whereas HP perceivers may view smiles as dominance displays (Niedenthal et al., 2010).

Affiliative, dominant, and enjoyment smiles are the three primary smile types put forth in the Simulation of Smiles (SIMS) model (Niedenthal et al., 2010). This framework posits that meaning representations through embodied simulations within the perceiver's own motor, affective, somatosensory, and reward systems drives these expressions. Thus, while a "smile" technically involves only the upward raising of the corners of the mouth, such an action can have vastly different motivations. An "enjoyment smile" can be a traditional display of pleasure and happiness, an "affiliative smile" can be one that merely reinforces or expresses positive social motives, and a "dominant smile" can reflect attempts to obtain status or authority. Since we were mainly interested in any potential differences in the interpretation of smiles, the SIMS perspective was a suitable framework for the follow-up, and we used these three smile types as the classification categories for the experiment.

Methods, results, and discussion. 69 UCSD undergraduates participated for course credit (77% female, M_{age}=21.3 years, SD_{age}=2.53 years). After random assignment to the HP, LP, or control condition using the same procedure from the main study (n=23 in each condition), participants then watched the video stimuli from the main study (target-power counterbalanced). All participants were given extensive information regarding the different smile types (Niedenthal et al., 2010), classified each video as one of three smile types (i.e., affiliative, dominant, or enjoyment smile), and rated expression intensity using a 1 (not at all) to 5 (very much) scale. Lastly, all participants completed the PANAS to assess mood, before being debriefed (Watson et al., 1988).

Controls and manipulation check. With the same essay scoring procedure as the main study (α=.87), the power manipulation was successful in inducing the desired states, $F(2,66)=132.3$, $p<.001$, where LP expressed the least power (M=2.50, SD=.77), HP expressed the most power (M=5.87, SD=.80), and controls fell in-between (M=4.15, SD=.49), *Tukey HSD*, *p*s<.001. Also, the PANAS revealed no significant differences by perceiver-power for positive affect (M=2.45, SD=.79), $F(2,66)=1.50$, *ns*, or negative affect (M=1.52, SD=.57), $F(2,66)=.75$, *ns*, demonstrating that self-reported mood did not vary across conditions.

Smile judgments. We analyzed both the counts of the different smile classifications (using McNemar's χ^2 tests with continuity correction, for repeated measures) and smile intensities (using a mixed-model ANOVA) by Perceiver-Power and Target-Power. Participants showed no significant differences based on their own (perceiver) power or target-power for smile classifications, χ^2s<2.13, *ns*, nor smile intensities, *F*s<2.80, *ns* (see Table S2). The follow-up rating experiment demonstrates that the power manipulation does not significantly affect smile perception, and therefore, the mimicry effects observed in the main study do not appear to be driven by differing interpretations of the expression (above the level of consciousness).

		Affiliative		Dominant		Enjoyment	
Smile Type →							
Power (Target) Condition →		*High (HP)*	*Low (LP)*	*High (HP)*	*Low (LP)*	*High (HP)*	*Low (LP)*
Power (Perceiver) Condition	*Control*	23: 3.39 (0.78)	24: 3.38 (1.10)	6: 3.83 (0.41)	6: 3.00 (1.67)	17: 3.71 (1.31)	16: 3.69 (0.95)
	High (HP)	22: 3.77 (0.97)	29: 3.62 (1.08)	9: 4.00 (1.00)	6: 3.00 (0.89)	15: 3.53 (1.25)	11: 3.55 (0.93)
	Low (LP)	18: 3.78 (1.17)	26: 3.73 (1.12)	12: 2.58 (1.08)	6: 3.00 (0.89)	16: 3.88 (0.62)	14: 3.86 (1.23)

Table S2. Follow-up study results: No significant differences in smile type classifications (i.e., affiliative, dominant, or enjoyment) or smile intensities were found according to perceiver-power or target-power. Tabulated counts, along with mean and standard deviation intensity ratings (1-5 scale), are included within each cell (i.e., count: mean (SD)).

References

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*, 245–265.

Dimberg, U., & Lundquist, L. O. (1990). Gender differences in facial reactions to facial expressions. *Biological Psychology, 30*, 151–159.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology, 85,* 453–466.

Hermans, E. J., Putman, P., & Van Honk, J. (2006). Testosterone administration reduces empathetic behavior: A facial mimicry study. *Psychoneuroendocrinology, 31,* 859–866.

Hess, U., & Bourgeois, P. (2010). You smile–I smile: Emotion expression in social interaction. *Biological Psychology, 84,* 514–520.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103,* 54–69.

Korb, S., Grandjean, D., & Scherer, K. R. (2010). Timing and voluntary suppression of facial mimicry to smiling faces in a Go/NoGo task—An EMG study. *Biological Psychology, 85,* 347–349.

Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). The computer expression recognition toolbox (CERT). *Face and Gesture, 2011,* 298–305.

Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The simulation of smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *Behavioral and Brain Sciences, 33,* 417–433.

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's *z* transformation be used? *Journal of Applied Psychology*, *72*, 146–148.

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). *Journal of Cross-Cultural Psychology, 38*, 227–242.

Thunberg, M., & Dimberg, U. (2000). Gender differences in facial reactions to fear-relevant stimuli. *Journal of Nonverbal Behavior*, *24*, 45–51.

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: Practical guide using statistical software*. New York, NY: Chapman & Hall/CRC.