Appears in the 10th IEEE International Conference on Automatic Face and Gesture Recognition, Shanghai, China 2013

Felt emotion and social context determine the intensity of smiles in a competitive video game

Jonathan Gratch, Lin Cheng, Stacy Marsella and Jill Boberg
University of Southern California, Institute for Creative Technologies
12015 Waterfront Drive Playa Vista, CA 90094

***Abstract*— The present study uses automatic facial expression recognition software to examine the relationship between social context and emotional feelings on the expression of emotion, to test claims that facial expressions reflect social motives rather than felt emotion. To vary emotional feelings, participants engaged in a competitive video game. Deception was used to systematically manipulate perceptions of winning or losing. To vary social context, participants played either with friends or strangers. The results support the hypothesis of Hess and colleagues that smiling is determined by both factors. The results further highlight the value of automatic expression recognition technology for psychological research and provide constraints on inferring emotion from facial displays.**

I. Introduction

Computers are becoming increasingly adept at recognizing and classifying displays of emotion. In sharp contrast, there has been less progress in using these displays to infer how a person actually feels. This, in part, reflects deep and unresolved theoretical disagreements concerning the meaning of emotional displays. Consider the case of smiles. On the one hand, smiles are claimed to be automatic and involuntary "readouts" of underlying joy [1]. On the other hand, smiles are seen as deliberate signals that communicate social motives and are not causally related to emotions [2]. These views are often cast as mutually exclusive, and both perspectives claim strong empirical support (see [3] for a discussion). Though theoretical in nature, this disagreement has profound practical implications for face and gesture research as, at its heart, it raises the question of whether it is even possible to infer emotional state from expressed behavior.

One challenge in resolving this dispute has been the difficulty in measuring affective displays in naturalistic settings. Many psychological studies utilize intrusive sensors such as facial electromyography (EMG) which involves abrading the skin and attaching wired electrodes to the side of the face [4]. Another common approach is to collect and annotate videos of facial movements using coding schemes such as the Facial Action Coding System (FACS). While allowing a more detailed analysis than EMG, FACS is extremely labor-intensive and requires considerable training before coders reach reliability. Recently, automatic facial expression recognition techniques have achieved some success at addressing other theoretical disputes in psychology [5]. One contribution of this article is we demonstrate the relevance of automatic smile detection to re-examine the relationship between felt emotion and social context on the display of emotion.

Another contribution of this work is that we address some limits in how prior studies have elicited emotions in their experimental settings. Emotions evolved to help organisms survive in a dynamic, semi-predictable, world and they typically reflect the extent to which an individual's needs are met or threatened. For example, positive emotions arise when an individual makes progress in achieving his or her goals. Nonetheless, many of the classic studies into the expression of emotion have not directly manipulated task-relevant factors such as perceptions of goal attainment nor examine how these perceptions (and resulting emotions) change over time. For example, the seminal studies by Fridlund [6], Hess and colleagues [4] and Jakobs and colleagues [3] all asked participants to watch emotional movie clips – a situation where the goal is at best implicit. Other work did examine situations where emotions are goal-relevant (e.g., Fernández-Dols and Ruiz-Belda looked at expressions of athletes during competitions [7]) but most such studies have been observational rather than experimental (meaning that they do not experimentally manipulate success and failure) and thus afford weaker conclusions. In this article, we engage participants in a competitive task and systematically manipulate perceptions of success or failure over time, studying the impact of these changes on observed expression.

Our results support claims that facial expressions do indeed reflect the underlying emotional state of participants, but that the social context is an important moderating factor. The results further highlight the value of automatic face recognition technology for psychological research and provide constraints on inferring emotion from facial displays.

II. Motivation and Hypotheses

As our study design builds on a series of prior findings, we first review this earlier work. In his classic 1991 study, Fridlund [6] created a potent counterweight to the then prevailing view of facial expressions as "readouts" of true emotion. Fridlund invited participants to come alone or with a friend and watch a funny movie. A true emotion perspective argues that smiles would reflect the intrinsic humor in the film. In contrast, Fridlund showed that different patterns of smiles occurred (measured by EMG) depending on if participants watched the film alone or together with a friend. He further presented data that the presence of smiles was not correlated with participants' felt emotion. These two findings (that smiling was altered by social context and uncorrelated with felt-joy) presented a serious and continuing challenge to the interpretation of facial expressions.

Subsequent research by Hess and colleagues identified several limitations in Fridlund's study which they sought to correct [4]. They argued that the lack of correlation between smiles and joy should be viewed with suspicion as Fridlund's design did not systematically manipulate the intensity of perceived humor. Thus, the lack of correlation could be more readily explained by the lack of variance in the original stimuli. They also sought to vary not only the presence of a co-participant, but also the nature of the relationship between participants (i.e., friends vs. strangers). Again like Fridlund, participants watched humorous films (but now varying in their rated funniness) and had participants watch alone or with strangers. Again, smiles were measured by EMG. Their findings illustrated that expressions were best predicted by considering both the intensity of the emotion elicitor and the social context. More recently, Jakobs et al [3] replicated the findings of Hess and colleagues using hand-annotated videos as a measure of facial activity.

Hypothesis 1: Following the work of Jakobs et al. and Hess et al., we hypothesize that smiles will be determined by both felt-joy and social context during a social task.

The previous studies are limited in that they focus on the behavioral consequences of emotion but do not explicitly consider the factors that produced the emotion in the first place, making it difficult to characterize how aspects of emotion-evoking situations might relate to facial expression. Indeed, these studies employed rather blunt instruments to elicit and measure emotion and its expression. Although emotion is often argued to arise in response to specific events, the above studies all analyzed both felt and expressed emotion in very broad temporal terms: they averaged expressions across the entire duration of a film, and did not consider how specific events within a film might relate to expression. Further, although emotions are frequently claimed to guide adaptive action in the world, participants in these studies were only asked to passively view pre-recorded actions of others.

To address these limitations, we must consider theories of how emotions arise from situations. There are many such theories. For example, theories of emotional contagion [8] argue that people can catch emotions from each other, but this sidesteps the question of where the emotion arose in the first place. Appraisal theories of emotion are one influential perspective on emotion elicitation [9, 10]. In appraisal theory, emotion is argued to arise from patterns of individual judgment concerning the relationship between events and an individual's beliefs, desires and intentions. These judgments, typically referred to as *appraisal variables* or *appraisal checks*, include such judgments as *goal-congruence* (i.e., is the triggering event congruent or incongruent with an individual's goals?), *expectedness* (i.e., was the event anticipated?), and *control* (i.e., can the individual causally shape outcomes following from the event?).

Patterns of appraisal are associated with specific feelings, physiological and behavioral reactions. For example, a controllable event with positive goal-congruence would likely elicit joy, whereas an uncontrollable event with negative goal-congruence would more likely evoke fear.

Appraisal theory suggests more systematic ways to explore the factors that might elicit emotion, emotional feelings, and their behavioral expression. In line with this, Kappas and Pecchinenda [11] highlighted the limitations of methodologies that treat complex sequences of events (such as a movie) as a single unit of analysis and, instead, argued for new experimental techniques for studying "event-related' facial activity. Inspired by appraisal theories of emotion, they created a Pacman-type video game and analyzed facial EMG responses to positive and negative events in the game such as reaching a power pill or being eaten by a monster. More recently, Wang and Marsella adopted a similar approach, using manual FACS coding to analyze responses [12]. Although both approaches showed initial promise, neither were fully explored. The current article can be seen as an attempt to continue and extend these lines of work.

In this article, we develop a video game, Mouse Wars, which allows us to systematically manipulate appraisal variables within the context of a social task. In this article we restrict our explorations to manipulations of goal-congruence. This, then leads to our second hypothesis:

Hypothesis 2: Following appraisal theories of emotion, we hypothesize that the intensity of both felt-joy and displayed smiles will be determined by appraisals of goal congruence (i.e., greater congruence will produce more smiles).

Finally, we aim to give further insight into how unfolding situations effect feelings and expression. Within a game like setting, goal congruence would seem to map straightforwardly to the subjective probability of winning or losing (which may differ from the objective probability) and this factor will change as a game unfolds. Several studies highlight immediate judgments and feelings may depend on what has come before. For example, perceptions of loss or gain are frequently made with respect to some reference point [13, 14]. Therefore, in the context of a game, feelings of joy might be influenced by whether the participant came from behind or not. This leads to the following research question:

Research Question 1: Do felt- and expressed-joy depend solely on the current probability of goal attainment, or are they reference-dependent (i.e., depend on if the person was previously ahead or behind)?

III. EMOTIONS IN TASKS

Mouse Wars is an online game we designed to systematically investigate how the structure of tasks impacts emotions and emotion displays.[1] The design is influenced by appraisal theories of emotion – which argue that emotions arise in response to judgments (possibly automatic) of how events in the world impact an individuals' beliefs, desires and intentions [9, 10]. For example, joy would be elicited in response to a certain event that facilitates an important goal, whereas fear would arise from an event that possibly threatens it. Following this theory, Mouse Wars aims to elicit

[1] The name stems from the fact that players must rapidly click their computer mouse to increase their chance of winning. The game board contains an icon of a rodent as a metaphorical homage to this.

Figure 1: Illustrates the red player's view of the game. The red player is experiencing the reversal-of-fortune-lose script and believes he is ahead of blue.

emotional feelings by manipulating the certainty, goal-congruence and control of task events during the course of an unfolding interaction. Note that, although Mouse Wars was developed to test the theoretical assumptions of appraisal theory (i.e., Hypothesis 2), for the purpose of testing Hypothesis 1, the only requirement is that it allows us to systematically manipulate participants' emotions.

Mouse Wars is loosely based on a board game called Battleship, which has been used to elicit emotion in past studies [15]. In Battleship, players take turns trying to sink their opponent's hidden ships by dropping bombs on a grid. Given that finding the opponents ships is largely a random process, this is approximately equivalent to a game of taking turns flipping a coin where the winner is the first person to achieve *N* heads, where *N* is the number of ships on the board. Mouse Wars retains this underlying "deep" structure but superficially appears to be a very different game.

Figure 1 illustrates the game interface. As in Battleship, players take turns. The goal is to get the mouse icon (seen in the upper right corner of the checkerboard) into the player's goal. To move, players spin a roulette wheel. If it lands on their color (henceforth referred to as a hit), the mouse advances one square towards their goal. If it reaches their goal, they have won. Otherwise, the next player takes a turn. Players can influence what percentage of the wheel contains their color (and thus the probability of goal attainment) by expending effort. In the current study, each turn is divided into three phases: a 10-second "effort-phase" during which players increase their chances by clicking the mouse button as rapidly as possible; a 2-second "spinning-phase" when the roulette wheel spins; and a 2-second "reaction-phase" where the game pauses to give players time to appreciate their success or failure. We expect to observe facial reactions to hits or misses during the 2-second reaction phase.

Mouse Wars is designed to alter several appraisals but in the current study we focus on participant perceptions of goal-congruence. Although participants assume the roulette wheel is fair and random, Mouse Wars allows us to script outcomes to create different turn-by-turn perceptions of winning the game. For example, a "reversal-of-fortune" trajectory might have one player build a commanding lead but ultimately lose the game. Participants fail to notice this deception as long as the turn-by-turn probability does not get too close to one or zero. Other manipulations are possible but not considered in this study. For example, we can manipulate perceptions of control how rapidly their effort increases the area of the wheel that contains their color.

In the present study, we use a two-phase script to manipulate both perceptions of winning and losing and reference-dependent effects. In the first phase, players experience a reversal of fortune: one player jumps out to an early lead (rounds 1-7), but then the other player pulls even (rounds 8-14). In the second phase, a player either wins or loses the game. The second phase is manipulated independently from the first resulting in four possible game experiences, listed in Table 1 (where "H refers to hit and "M" refers to miss): reversal-of-fortune-win (RW), reversal-of-fortune-lose (RL), close-call-win (CCW) and close-call-lose (CCL). These scripts were developed from pilot studies to maximize differences in perceived probability of winning/losing without creating suspicion of the manipulation. With these scripts, we are able to elicit and compare five qualitatively distinct classes of events that differ in terms of goal congruence – being even, being ahead, being behind, having won and having lost – and explore the impact of prior perceptions of winning or losing on current judgments (i.e. being ahead and wining vs. being ahead and losing). Note that RL and CCW are identical up to round 14. Similarly, RW and CCL are identical in the first phase.

Table1: Pattern of hits and misses by script and turn

Script	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
RL	H		H		M		H		M		M		M		M		
CCW	H		H		M		H		M		M		M		M		H
RW		M		M		M		H		M		H		H		H	
CCL		M		M		M		H		M		H		H		M	

To assist in manipulating the social context, the game allows the players to see live video of each other. The presence of the video can be manipulated by experimental design: in this article it is always visible. Whether it is displayed or not, video of each participant is captured, stored and synchronized with game events. We capture video data with the resolution of 640*480 at 12 frames per second.

IV. METHOD

A. Experimental Design

To investigate the relationship between felt-emotion, expressed-emotion and social context we employ a 2x4 between-participants repeated-measures design. We examine two levels of social context (games between friends and games between strangers) and participants experienced one of 4 possible scripted sequences of hits and misses in Mouse Wars (RL, RW, CCL or CCW). Self-reported emotions were repeatedly measured (within-subjects) at four points in each game: before the game started, at turn 7 (when one player is close to winning), at turn 14 (when the other player pulls even), and at the end of the game. Facial expressions were measured continuously, but to compare expressed emotion with self-reported emotion, we assess expressions repeatedly at these same points.

B. Measures of self-reported Joy

Participants were asked how much joy they experienced at the four points in the game discussed above, and also asked their subjective chance of winning during the game to assess if we successfully manipulated goal congruence. Following the approach in [15], these items were assessed by a single-item on a visual analog scale ranging from 0 to 100.

Participants were also given several other self-reported measures that were included to address hypotheses that are not considered in this current article. We will not discuss these items further but describe here for completeness. Before playing Mouse Wars, participants completed a 5-item *Game Experience Questionnaire*, the 20-item Positive And Negative Affect Scale [16] (a measure of participants' mood at the start of the study) and the 8-item Short Grit Scale [17] (which measures trait-level perseverance and passion for achieving goals). Joy was assessed as part of an 8-item *Emotion and Appraisal Questionnaire*. In addition to joy, this asked participants the extent to which they currently feel hope, disappointment and fear on a visual analog scale ranging from 0 to 100. The remaining four appraisal questions ask people the current importance they assign to winning, their perceived chance of winning, their sense of control over outcomes in the game and the amount of effort they are currently expending, again on a 0 to 100 visual analog scale. Control and chance were not asked post-game as these questions have no meaning at this point.

C. Measures of Smiles

Video of each participant's face was collected throughout the game and analyzed after the experiment using the OKAO Vision system by Omron Inc. [18]. OKAO has been previously used with some success to measure facial expressivity [19, 20]. Briefly, it uses computer vision techniques to identify 16 facial landmarks. From this, it derives a variety of facial pose estimates including a smile intensity ranging from 0 (no smile) to 100 (full smile).

Figure 2: An illustration of the output of OKAO Vision software. It outputs an intensity value from 0 to 100 for each video frame. Here, it outputs values of 60 (left image) and 89 (right image) on one of our participants.

To assess how smiles changed in response to perceived changes in goal attainment, we only analyze video segments during the 2-second reaction-phase corresponding to the period immediately-following when the outcome of the roulette wheel is displayed. We examine the same four turns for which we query self-reported joy (turns 7, 14, and post-game). Expressed-joy is operationalized as the average intensity of smile detected by OKAO in each frame of the resulting 2.0 second video.

D. Procedure

Participants arrived in groups of 6-10. Upon arrival, they were asked to read and sign informed consent forms then seated at individual workstations in a classroom setting. The game and all questionnaires were administered from this computer. After completing a questionnaire about their background including demographic data and the pre-game questionnaires, participants watched a short video describing how to play Mouse Wars and complete the first of the four instances of the emotion and appraisal questionnaire. Participants were then paired with another individual in the room (either a stranger or their friend) and began playing the game. At the point at which the game begins, they see a video image of themselves and their partner on game board. As described above, the game was paused on turns 7 and 14 to allow participants to complete the emotion and appraisal questionnaire, and then immediately after the game.

Participants played two additional unscripted games with the same participant afterwards (these games have not yet been analyzed and won't be discussed further in this article). Afterwards, the participants were fully debriefed, paid their participation fee and thanked for their participation.

E. Participants

136 persons (66 friends; 70 strangers) participated in the study.[2] They were recruited via www.craigslist.com from

---

[2] We originally recruited 171 participants. Several were excluded due to prematurely terminated games due to software/OS problems, and a power outage in one session. Because of no-shows, we occasionally had unequal numbers of strangers. On these occasions (25 total), the unpaired-stranger played a confederate blind to the experimental goals and condition. All confederate data is excluded from analysis.

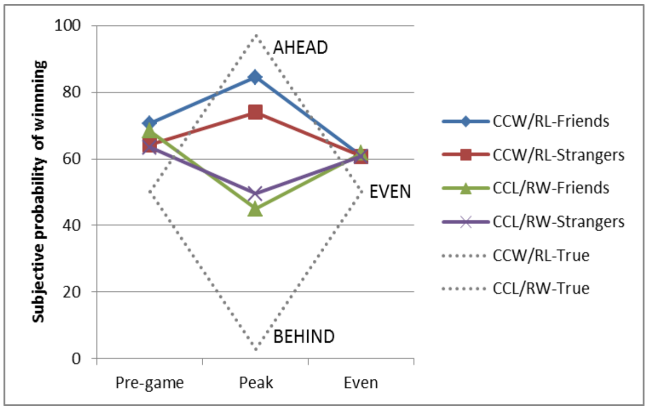

Figure 3: Subjective probability of winning as function of friends vs. strangers and collapsed script. Also illustrates objective probability.

the general Los Angeles area and were compensated $25 for one hour of their time. Participants were instructed to either bring a friend or come alone. We made no attempt to match gender. There were 75 males (55.1%) and 61 females. Mean age was 31.9 (*SD*=11.6), ranging from 18 to 61 years.

V. Results

A. Manipulation Check of Goal Congruence

We first verified that participants recognized that their chance of winning (i.e., goal congruence) changed across the measured events. We used a repeated-measure ANOVA during the two measured time periods in phase 1 (when the outcome of the game is still uncertain). We performed a 2(Ahead-Even vs. Behind-Even) x 2(time) ANOVA with Greenhouse-Greisser correction to examine the relationship of script and time on perceived probability of winning. There was a main effect of script ($F(1,132) = 17.53$, p=.000, η_ρ^2=.117) and a script-by-time interaction ($F(2,132)$=59.70, p=.000, η_ρ^2=.311). Post hoc t-tests revealed that this effect is solely due to a difference ($p < 0.5$) in probability of winning at the peak (Ahead > Behind). There was no significant difference in perceived chance of winning by script when the game was even and no significant effects of social context. This is illustrated in Figure 3. From this, we conclude the manipulation of goal-congruence was successful.

B. Manipulation Check of Smile Intensity

OKAO purports to produce an accurate measure of smile intensity but we decided to independently assess this relationship by comparing our intended measure of smile (described in Section IV.C) with manual ratings of naïve coders recruited through Amazon's Mechanical Turk. Although our study focuses on specific points of the game (turns 7, 14, and post-game), we selected a subset of 2-second reaction-phase videos across the entire game to avoid any possible sampling bias that might result from focusing on these points. In total, the study produced approximately 2200 reaction videos (136 participants x 16-17 turns each) from which we selected 579 for manual coding. Each selected video was observed by 5 coders and rated using the Self-Assessment Manikin [21]. This scale assesses perceived valence, arousal and dominance on a 9-point likert scale. We used the average valence rating of the 5 coders as to represent the perceived joy in each video. We next correlated this manual measure with the OKAO-reported measure. We found a strong (r =.705) and significant (p=.000) correlation between these two measures. From this we conclude that OKAO is an acceptable automatic measure of expressed joy.

C. Emotion Results

We next performed a series of ANOVAs (using Greenhouse-Greisser correction) to assess which factors influenced felt and expressed-joy.[3] Although participants experience one of four possible sequences of hits and misses, we are able to exploit the shared structure of the scripts and significantly increase our statistical power by examining each time period separately. Thus, for both felt and expressed-joy, we performed three separate 2x2 ANOVAS looking at the relationship between goal congruence and social context at each time period (Peak, Even and Post-game). Specifically, we examined Ahead vs. Behind (i.e., RL/CCW vs. RW/CCL), falling-to-even vs. rising-to-even (i.e., RL/CCW vs. RW/CCL) and Won vs. Lost (i.e. RW/CCW vs. RL/CCL). This allowed us to examine Hypotheses 1 and 2, but does not completely assess the carry-over effects needed to test Research Question 1 (e.g., it does not test if winning after being ahead feels different than wining after being behind). For this we examined the main effect of script for the second ANOVA (falling-to-even vs. rising-to-even) but for post-game measures, we performed a 2(social context) x 2(won vs. lost) x 2(previously-ahead vs. previously-behind) ANOVA to examine if the previous game state had an influence.

Means of felt and expressed-joy as a function of goal congruence and social context are summarized in Figure 4. We discuss the significance of these differences for each variable separately.

1) Felt-joy

There were large main effects of script at Peak ($F(1,132)$=25.45, p = 0.000, η_ρ^2 = .162) and Post-game ($F(1,132)$ = 36.75, p = 0.000, η_ρ^2 = .218) but no difference at Even. Specifically, Ahead produced more joy than Behind ($t(134)$ = -4.72, p=.000) and Winning felt better than losing ($t(110)$ = 6.28, p=.000). There was little effect of social context: no interaction of context with script at any time and no main effect of social context at Even or post-game. The ANOVA did reveal a small main effect of social context at the Peak ($F(1,132)$ = 4.67, p = 0.033, η_ρ^2 = .034), but this did not reach significance in a post hoc t-test.

With regard to reference-dependent effects, there was no main effect of falling-to-even vs. rising-to-even at the second time period and no interaction with social context.

[3] When data was missing (e.g., if smile could not be detected because the face was out of the frame, we excluded these specific data points from analysis and reduced the DOF (i.e., we did not impute missing data).

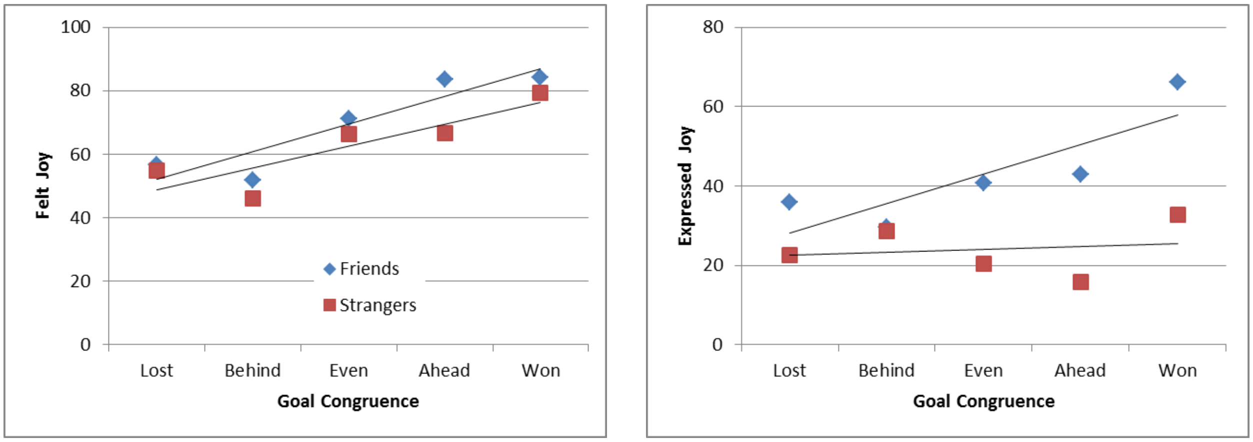

Figure 4: Illustrates participants' felt-joy (left) and expressed-joy (right) as a function of probability of goal attainment (goal congruence). Changes in felt-joy are associated with changes in goal congruence, regardless of social context. However smiles only show this association in games between friends.

Post-game, there was no main effect of previously-ahead vs. previously-behind and no interactions with other factors.

To summarize, friend and stranger pairs felt nearly identical levels of joy in of a game. Both friend and stranger pairs exhibit a clear association between felt-joy and goal congruence. We found no reference-dependent effects.

D. *Expressed-joy*

There were significant main effects of social context at all time periods: Peak ($F(1,127)$ = 6.10, p = 0.015, η_ρ^2 = .046), Even ($F(1,124)$ = 13.98, p = 0.000, η_ρ^2 = .101) and Post-game ($F(1,127)$ = 9.48, p = 0.003, η_ρ^2 = .075). There was a main effect of script at two time points -- Even ($F(1,124)$ = 4.79, p = 0.031, η_ρ^2 = .037) and Post-game ($F(1,112)$ = 12.05, p = 0.001, η_ρ^2 = .097) – and a script-by-context interaction at the Peak ($F(1,127)$ = 6.10, p = 0.015, η_ρ^2 = .046). Post hoc *t*-tests examining the impact of social context at the five levels of goal-congruence are summarized in Table 2.

Table 2: Expressed-joy – Friends vs. Strangers

	Friends	Strangers	All	Difference	
	M (SD)	M (SD)	M (SD)	Sig.	Effect Size η_ρ^2
Won	66.07 (35.85)	32.65 (34.43)	49.36 (38.69)	.001*	.190 (large)
Ahead	42.94 (35.44)	15.87 (23.38)	27.58 (31.98)	.000*	.178 (large)
Even	40.65 (32.44)	20.31 (30.70)	30.16 (33.05)	.007*	.095 (medium)
Behind	29.62 (34.16)	28.64 (36.59)	29.19 (34.92)	.918	.000 (none)
Lost	35.76 (41.09)	22.59 (32.85)	28.74 (37.19)	.173	.032 (small)

*Significant at $p <= .01$

With respect to reference-dependent effects on expressed-joy, the ANOVA revealed a main effect of falling-to-even vs. rising-to-even at the second time period ($F(1,124)$ = 4.79, p = 0.031, η_ρ^2 = .037), although the subsequent *t*-test failed to confirm this difference. Post-game, there was no main effect of previously-ahead vs. previously-behind on felt-joy and no interaction with social context.

Overall, friends showed more expressed-joy in response to positive events (Ahead and Won) whereas strangers' expressed-joy remained essentially constant, regardless of their perceived chance of winning. The expressions of friends and strangers reflected the current state of the game, and did not show reference-dependent effects.

VI. Discussion

Our main findings are illustrated in Figure 4. Participants' felt-joy is clearly influenced by our manipulation of goal congruence (i.e., perceptions of winning or losing) and friends and strangers largely feel the same emotions in response to this manipulation. In other words, the social context has no effect on emotions experienced by planers in this game. Expressions of joy, however, tell a more complex story. Smiles are impacted by our manipulation of goal congruence, but only in games between friends. When examining games between strangers, smiles are not correlated with perceptions and feelings of winning or losing.

The results support our Hypothesis 1: smiles are determined by both felt-joy and social context during a social task. The data fail to support Fridlund's contention that smiles are unrelated to felt emotion. Thus, these findings replicate and extend findings by Jackobs et al. [3] and Hess et al. [4] that social situation and emotion combine to determine the intensity of facial expression. The current study extends these previous findings by focusing on an active social task (Mouse Wars) rather than passive watching of videos, and by focusing on event-related facial activity (which is arguably more appropriate for studying emotion), than averaging expressions over longer windows of activity

Figure 5: The image shows one frame from a video of stranger-participant reacting to a miss in the game. She is exhibiting apparent masking behavior, briefly produces a frown display that quickly followed by a smile

(and thereby breaking the association between expressions and the events and cognitions that potentially elicit them).

The study also supports our Hypothesis 2: task appraisals help determine the intensity of both felt and expressed emotion, and thereby provide support for appraisal theories of emotion. Smiles positively correlate with both objective and subjective differences in goal congruence. It should be noted, however, that these differences failed to materialize in the expressions of strangers despite them feeling similar differences in felt-joy. Appraisal theories try to account for the effect of social context in at least two ways. On the one hand, aspects of the social situation can be appraised and result in different felt and expressed emotions [22]. For example, when playing with a friend, one might feel a social emotion of pride which might also impact facial expressions. A limitation of the current study is we did not explicitly ask participants to rate any social emotions. On the other hand, appraisal theories and basic emotion theories allow that expressions can be deliberately masked depending on the social context. For example, strangers might mask their smiles through the use of “smile controls” as a way to manage the level of intimacy in the social interaction [23].

We also examined if joy (both expressed and felt) was dependent on reference points – i.e., did people feel and express differently if they pulled-even when previously behind or fell-to-even when previously ahead. We found no evidence that prior game events influenced current joy. Rather, joy was best predicted by the current probability of goal attainment.

A. *Mechanism?*

The present study doesn’t directly resolve the question as to why friends act differently than strangers. For example, it may be the case that friends are deliberately choosing to express felt emotion through smiles or it could be that strangers are suppressing their true feeling. However, our existing data is inconsistent with some other proposed mechanisms and further analysis of our existing data may reveal further insights.

One promising direction is to look for evidence of masking by analyzing the co-occurring facial action units or the temporal dynamics of smiles under the different experimental conditions. For example, using automated techniques, Mohammed et al. [24] found evidence that felt and false smiles had different temporal characteristics and Duckworth et al. [17] found that depressed individuals masked felt smiles with so-called ‘smile controls.”

Anecdotally, we find several hints that strangers are, on the one hand, using smile controls to mask positive expression of emotion, and on the other hand, using smiles to mask negative displays. For example, consider the expression produced by a stranger-participant, illustrated in Figure 5. The participant has just experienced a “miss” event. A manual FACS annotation of the video indicates she initially produces AU14 (dimpler), AU23 (lip tightener) and AU24 (lip pressor) which is rapidly followed and temporally overlaps with AU12 (smile). While this is only a single example, it highlights both the complex nature of facial displays and the potential value of detailed automatic analysis. We plan to investigate this more systematically in our future analyses.

Although our findings do not clearly distinguish between “readout” vs. social motive explanations of expressions, they imply constraints on other proposed mechanisms for expressed emotions in social settings. Recent years have seen an increase of interest in bottom-up mimicry explanations for social emotions. In contrast, the results are inconsistent with a simple contagion or mimicry explanation for emotional expression (e.g., [25]). Simple contagion would predict that smiles would be correlated across participants. Rather, we observe counter mimicry: events that evoked the most smiles in one participant (Ahead and Won) evoked the least smiles in their partner (Behind and Lost). This is consistent with older findings by Lanzetta and Engles [26] that competitive situations elicit patterns of counter-mimicry and counter-empathy. We prefer the explanation that the observed counter-mimicry is an emergent consequence of the fact that players are appraising game events from opposing perspectives, but a bottom-up “counter-mimicry” mechanism would produce a similar pattern of behavior, and thus cannot be ruled out.

B. *Concluding thoughts*

Pragmatically, the distinction between emotion as “readouts” of true emotion or deliberate choices to show their true feelings may be irrelevant for many applications. A great number of social interactions involve settings where people are familiar with each other and there is no overt motive to mislead. For example, in the current study, expressions between friends were diagnostic of their mental state, regardless as to whether this expression was automatically or deliberately produced.

The results demonstrate the promise of automatic facial expression analysis to address psychological questions. Previous studies on the relationship between emotion and facial expressions have relied on intrusive (EMG) or tedious

(FACS) methods for measuring facial activity. The present study gives confidence that video analysis has reached parity with these approaches, and moving forward, can allow a much richer analysis of human behavior. For example, automatic techniques should allow much larger databases of behavior to be analyzed more efficiently and naturally than physiological approaches like EMG.

Despite this promise, these results offer several cautions to automatic face and gesture researchers. Clearly, attempts to infer the meaning of an expression of emotion must explicitly consider the social context. It should also be noted that Mouse Wars does not give players an explicit motive to dissemble, and more strategic games like poker may produce a different pattern of results. However, this also creates opportunities. For example, the results suggest that if appraisals like goal congruence are known to the expression-understanding system, then the presence or absence of smiles should allow inferences about the social context (i.e., the system could infer if players are strangers based on the absence of their observed reactions).

To summarize, the present study used automatic facial expression software to examine the relationship between social context and emotional feelings on the expression of emotion, with the aim of testing claims that facial expressions reflect social motives rather than felt emotion. The results support the hypothesis of Hess and colleagues that smiling is determined by both factors. Further, the results give evidence that appraisal variable of goal congruence predicted the intensity of both felt and expressed smiles. The results further highlight the value of automatic face and gesture recognition technology for psychological research and provide some cautions on directly inferring emotion from facial displays.

ACKNOWLEDGMENT

We are grateful to Arvid Kappas, Rainer Reisenzein and Jeff Cohn for advice on experimental design and data analysis. This work is supported by the Air Force Office of Scientific Research (FA9550-09-1-0507) and the U.S. Army RDECOM. The content does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

[1] P. Ekman, "Facial expression and emotion," *American Psychologist,* vol. 48, pp. 384-392, 1993.

[2] A. J. Fridlund, "Evolution and facial action in reflex, social motive, and paralanguage," *Biological Psychology,* vol. 32, pp. 3-100, 1991.

[3] E. Jakobs, A. S. R. Manstead, and A. H. Fischer, "Social Motives and Emotional Feelings as Determinants of Facial Displays: The Case of Smiling," *Personality and Social Psychology Bulletin,* vol. 25, pp. 424-435, April 1, 1999 1999.

[4] U. Hess, R. Banse, and A. Kappas, "The intensity of facial expression is determined by underlying affective state and social situation," *Journal of Personality and Social Psychology,* vol. 69, pp. 280-288, 1995.

[5] J. F. Cohn, "Advances in Behavioral Science Using Automated Facial Image Analysis and Synthesis [Social Sciences]," *Signal Processing Magazine, IEEE,* vol. 27, pp. 128-133, 2010.

[6] A. J. Fridlund, "Sociality of social smiling: Potentiation by an implicit audience," *Journal of Personality and Social Psychology,* vol. 60, pp. 229-240, 1991.

[7] J. M. Fernández-Dols and M.-A. Ruiz-Belda, "Are Smiles a Sign of Happiness? Gold Medal Winners at the Olympic Games," *Journal of Personality and Social Psychology,* vol. 69, pp. 1113-1119, 1995.

[8] E. Hatfield, J. T. Cacioppo, and R. L. Rapson, Eds., *Emotional Contagion* (Studies in Emotion and Social Interaction. Cambridge: Cambridge University Press, 1994, p.^pp. Pages.

[9] R. Lazarus, *Emotion and Adaptation*. NY: Oxford University Press, 1991.

[10] K. R. Scherer, "Appraisal Considered as a Process of Multilevel Sequential Checking," in *Appraisal Processes in Emotion: Theory, Methods, Research*, K. R. Scherer, A. Schorr, and T. Johnstone, Eds., ed: Oxford University Press, 2001, pp. 92-120.

[11] A. Kappas and A. Pecchinenda, "Don't wait for the monsters to get you: A video game task to manipulate appraisals in real time," *Cognition and Emotion,* vol. 13, pp. 119-124, 1999.

[12] N. Wang, S. Marsella, and T. Hawkins, "Individual differences in expressive response: a challenge for ECA design," presented at the Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 3, Estoril, Portugal, 2008.

[13] A. Tversky and D. Kahneman, "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quarterly Journal of Economics,* vol. 106, pp. 1039-1061, November 1, 1991 1991.

[14] B. A. Mellers, A. Schwartz, K. Hu, and I. Ritov, "Decision affect theory: Emotional reactions to the outcomes of risky options," *Psychological Science,* vol. 8, pp. 423-429, 1997.

[15] J. Gratch, S. Marsella, N. Wang, and B. Stankovic, "Assessing the validity of appraisal-based models of emotion," presented at the International Conference on Affective Computing and Intelligent Interaction, Amsterdam, 2009.

[16] D. Watson, A. Tellegen, and L. A. Clark, "Development and validation of brief measures of positive and negative affect: The PANAS scale," *Journal of Personality and Social Psychology,* vol. 54, pp. 1063-1070, 1988.

[17] A. L. Duckworth and P. D. Quinn, "Development and Validation of the Short Grit Scale (Grit–S)," *Journal of Personality Assessment,* vol. 91, pp. 166-174, 2009/02/17 2009.

[18] S. Lao and M. Kawade, "Vision-based face understanding technologies and their applications," *Advances in Biometric Person Authentication,* pp. 339-348, 2005.

[19] N. Sawai, N. Hato, N. Hakuba, H. Takahashi, M. Okada, and K. Gyo, "Objective assessment of the severity of unilateral facial palsy using OKAO Vision® facial image analysis software," *Acta Oto-Laryngologica,* pp. 1-5, 2012.

[20] Y. Horiuchi, T. Komatsu, and F. Nakaya, "Should Candidates Smile to Win Elections? An Application of Automated Face Recognition Technology," *Political Psychology,* 2012.

[21] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry,* vol. 25, pp. 49-59, 1994.

[22] A. S. R. Manstead and A. H. Fischer, "Social appraisal: The social world as object of and influence on appraisal processes," in *Appraisal processes in emotion: Theory, methods, research*, K. R. Scherer, A. Schorr, and T. Johnstone, Eds., ed New York: Oxford University Press, 2001.

[23] L. I. Reed, M. A. Sayette, and J. F. Cohn, "Impact of depression on response to comedy: A dynamic facial coding analysis," *Journal of Abnormal Psychology,* vol. 116, pp. 804-809, 2007.

[24] E. H. Mohammed, D. J. McDuff, and W. P. Rosalind, "Exploring Temporal Patterns in Classifying Frustrated and Delighted Smiles," *IEEE Transactions on Affective Computing,* vol. 99, 2012.

[25] P. M. Niedenthal, M. Mermillod, M. Maringer, and U. Hess, "The simulation of smiles (SIMS) mode: embodied simulation and the meaning of facial expression," *Behavioral and brain sciences,* vol. 33, pp. 417-480, 2010.

[26] J. T. Lanzetta and B. G. Englis, "Expectations of cooperation and competition and their effects on observers' vicarious emotional responses," *Journal of Personality and Social Psychology,* vol. 45, pp. 543-554, 1989.

Appears in the 10th IEEE International Conference on Automatic Face and Gesture Recognition, Shanghai, China 2013