ENGINEERING TEAM PERFORMANCE AND EMOTION:

AFFECTIVE INTERACTION DYNAMICS

AS INDICATORS OF DESIGN TEAM PERFORMANCE

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF MECHANICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Malte Friedrich Jung

August 2011

This dissertation is online at: http://purl.stanford.edu/th996ft5752

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Larry Leifer, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

James Gross

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Pamela Hinds

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Ralf Steinert

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

ABSTRACT

Emotions play a crucial role in the practice of designing in teams. Engineers engaged in the activity of designing complex technical systems constantly negotiate disagreements with their peers and sometimes these disagreements turn hostile. Excitement and interest are suddenly replaced by frustration and contempt, and what started as a promising project falls short of expectations. Even though there is now a growing body of research about the relationship between social dynamics and engineering team performance, we still lack a coherent body of knowledge about the role of emotions in engineering team interactions. The reasons for this lack of research might be partially found in an absence of appropriate research methods and theory. This lack of research on emotions in engineering design interactions is opposed to an extensive body of work on emotions in marital interactions, an area in which researchers were able to powerfully relate emotional interaction dynamics with short and long-term outcomes.

The present work adapts theory and methods that have been successful in predicting satisfaction and divorce in marriages to study relationships between affective interaction characteristics and performance relevant outcomes in engineering design teams. In three studies, I show that Group Hedonic Balance, the balance between positive and negative affect present in a team, is a strong correlate of engineering team performance. Drawing from this work, I then propose a theory about the relationship between intra-group conflict and performance based on affective distinctions of conflict types. This affect based intra group conflict theory departs from established theories, which are based on topical distinctions of conflict as either task, process or relationship-oriented.

ACKNOWLEDGEMENTS

To Larry Leifer, the best advisor I could wish for. Thank you for letting me join your dancing with ambiguity.

To the members of my dissertation committee: Pamela Hinds, James Gross, and Martin Steinert, whom I feel so privileged to call my advisors.

To Janine-Giese Davis, who inspired my interest in emotion.

To Ozgur Eris, without whom this work would not exist.

To Clifford Nass, for chairing my dissertation committee, and for his constant inspiration and motivation.

To Neeraj Sonalkar, who sharpened my thinking in so many ways.

To Ade Mabogunje, who was always there when I got stuck.

To Karen Piemme, who showed me how to train students in SPAFF.

To all the students, and especially to Julia Tang and Dan Lopez, who helped me with that tedious task of coding videos. Without you I would have never been able to finish all this.

To Syed Shariq and the members of the Kozmetsky Global Collaboratory for always making me feel welcome.

To Jonathan Edelman, Micah Lande, Becky Currano and the members of the Center for Design Research.

To my dissertation writing group: Nikki Velasco, Nicole Arshan, Rachel Kalmar, and Adam Rosenblatt.

To Sohyeong Kim and Eliver Ghosn, my amazing friends, who were always there when I needed them.

To my brothers Johannes, and Max.

To Lisa, my little big sister.

To my parents, who gave so much to make all this possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

INTRODUCTION

What makes some engineering teams deliver outstanding results while others fail to live up to expectations? Determinants of engineering team performance are often sought in technical knowledge, expertise, and interaction content. However, the practice of designing in teams is inherently social (Bucciarelli, 1988; Minneman, 1991). Team members interact with one another to generate and develop concepts and physical artifacts over time. The nature of their interaction has an impact on performance relevant outcomes such as the quality of final deliverables, the fulfillment of personal needs, or the willingness of a team to work together in the future (Hackman, 1987). Researchers interested in uncovering the relationship between social aspects of designing in teams and performance have looked at a large variety of phenomena such as question asking (Eris, 2004), gesturing activity (Tang, 1989), process changes (Frankenberger & Auer, 1997), and many other aspects (Finger & Dixon, 1989a, 1989b). Even specific laboratory environments were designed to study design interactions (Carrizosa, Eris, Milne, & Mabogunje, 2002). Despite these broad investigations of different interaction dynamics, not much has been done to look at the role of emotions in designing, and how the emotions designers express, or feel, shape subsequent performance relevant outcomes. In design, researchers have primarily looked at how products elicit certain emotions (Desmet, 2003, 2005; Desmet & Hekkert, 2007; Hekkert, 2006; Norman, 2004). Emotions, however, are recognized to play an important role regarding the effectiveness of teams in general (Barsade & Gibson, 2007; Janis, 1982; Kelly & Barsade, 2001; Tran, 2004). Especially the way a team handles disagreement and conflict has frequently been shown to relate to team outcomes (Garcia-Prieto, Bellard, & Schneider, 2003; Hinds & Bailey, 2003; Janis, 1982; Jehn & Mannix, 2001).

The reasons for a lack of research into the emotional dynamics of engineering design teams might be partially found in an absence of appropriate research methods. Currently, most research exploring the relationship between affect and performance

relies on self-report measures. Tran (2004), for example, explored how specific, self reported emotions predict performance in decision making teams. Many self-report measures of affect have been developed that are easy to obtain and that have been frequently validated (Russell, Weiss, & Mendelsohn, 1989; Watson, Clark, & Tellegen, 1988; Watson & Tellegen, 1985; Wiggins, Trapnell, & Phillips, 1988). However, studies that specifically explore performance relevant aspects using questionnaires are at danger of being biased by participants' implicit theories about performance (Staw, 1975). Dong's studies of affective appraisal verbalizations in design interactions seem to be so far the only investigation of the role of affect in design team interactions (Dong, 2006; Dong, Kleinsmann, & Valkenburg, 2009). These studies however only rely on the analysis of verbal protocols. They focus mostly on verbalized displays of affect towards ideas or artifacts (e.g."that's a good idea") and neglect other behavioral channels such as facial muscle movements or changes in voice tone as indicators of affect. Overall the current methods have not led to the development of a coherent understanding about the mechanisms underlying the relationship between the affective characteristics of engineering design team interactions and their performance relevant outcomes.

GROUP HEDONIC BALANCE

In contrast to researchers of engineering teams, researchers of married couples have developed powerful methods to gain insights into relationship between affective interaction dynamics and performance. In particular, the methods developed by Gottman and his colleagues led not only to new theory about the relationship between affective dynamics and subsequent outcomes (Gottman, 1994), but also to novel therapy practices (Gottman, 1999), and direct insights that couples could use on their own to improve the quality of their relationships (Gottman & Silver, 1994). Especially illustrative is a study in which Gottman and Levenson (2000) showed that it is possible to accurately (93%) predict the fate of a marriage based on the affective interaction quality determined from just a 15 minute video sample of a couple's in-

teraction. In another study divorce could be predicted based on the affective interaction quality during the first 3 minutes of a conflict episode with 80% accuracy (Carrere & Gottman, 1999). Using the same methods, it was possible to predict marital outcomes such as satisfaction and divorce across a wide range of studies (Gottman, 1994; Gottman, Coan, Carrere, & Swanson, 1998; Gottman, Katz, & Hooven, 1996; Jacobson, Gottman, Gortner, Berns, & Shortt, 1996; Levenson, Carstensen, & Gottman, 1994; Levenson & Gottman, 1983; Levenson & Gottman, 1985). Central to these studies was the idea of a balance theory of marriage (Gottman, 1993) which posits that couple's ability to regulate the hedonic balance of positive and negative affect is critical for the quality of their interaction and long-term outcomes. A classification of couples based on their hedonic balance into regulated (more positive than negative) and non-regulated (more negative than positive) couples in the studies cited above was repeatedly shown to be predictive of marital satisfaction and divorce. The notion of hedonic balance as a predictor of individual well-being has been established in several studies (Diener, 2000; Diener & Diener, 1996; Diener & Suh, 1997; Kahneman, 1999; Kahneman & Krueger, 2006), but besides the work of Losada (Fredrickson & Losada, 2005; Losada, 1999; Losada & Heaphy, 2004), not much has been done to explore the relationship between hedonic balance and outcomes in teams. Specifically the balance theory of marriage has not been applied to investigate the relationship between engineering team interaction and performance.

In this dissertation, I will therefore investigate the notion of Group Hedonic Balance, the balance between positive and negative affect expressed and experienced in a small group, as an indicator of performance in engineering teams.

RESEARCH AIMS AND HYPOTHESES

The broad long-term objective of my research is to further our understanding of the mechanisms underlying the relationship between the affective characteristics of small work group interactions and their immediate and long-term effects on performance relevant outcomes.

The first aim of the following three studies is therefore to adapt a set methods that have been successful in predicting satisfaction and divorce in marriages to solve the problem of understanding the relationships between affective interaction characteristics and performance relevant outcomes in design teams. Specifically I want to use these methods to predict performance from affective dynamics in short interaction samples across a variety of design scenarios.

The second aim of the following studies is to extend Gottman's balance theory of affect to small groups and explore the role of Group Hedonic Balance as an indicator of team performance empirically across different design relevant contexts. Specifically I hypothesize that Group Hedonic Balance correlates with performance relevant outcomes: Teams with a higher group hedonic balance will outperform those with a lower group hedonic balance.

DESIGNING THE INSTRUMENT

My studies draw from three core techniques that I bring together for the study of design interactions: (1) Thin slicing of team behavior, (2) systematic observation of behavior and (3) Lab-sampling of team behavior. These techniques were adapted in several rounds of prototyping with different types and sizes of engineering teams. I introduce these techniques here to give an overview about the approach used in my studies. My central goal was to design procedures that would allow me to capture a small sample of a group's interaction that would be representative of its affective interaction style. Systematic observation of behavior and thin slicing are used in all three studies, while lab-sampling of team behavior is only used in the third study.

THIN SLICING OF TEAM BEHAVIOR

Thin-slicing refers to the process of making accurate classifications based on small samples, or "thin slices" of expressive behaviors (Ambady & Rosenthal, 1992). The thin-slicing research showed powerfully that certain behavioral characteristics are stable over time and that only a small interaction sample is necessary to make meaningful judgments about behavior occurring over longer durations such as hours, or even months. For example, in a frequently cited study, Ambady and Rosenthal (1993) showed that end-of-semester teacher evaluations could be accurately classified based on judgments of 30 second silent video clips of the respective teachers. In addition to its more prominent applications in predicting teacher evaluations (Ambady & Rosenthal, 1993) and marital outcomes (Carrere & Gottman, 1999; Gottman & Levenson, 2000), the method of thin slicing has been applied successfully across a wide range of other contexts such as doctor-patient interactions, family interactions, interviews, or work related interactions (Ambady & Rosenthal, 1992). In a meta-analysis across 38 different studies, Ambady and Rosenthal (1992) were able to show that short behavioral samples ranging between 30 seconds and 5 minutes, are highly

indicative of long-term characteristics, irrespective of the specific context they were taken in. The method of thin slicing is particularly relevant when gathering data about the interactions of real teams that exist not only for a few minutes or hours in the lab but over weeks, months or even years.

SYSTEMATIC OBSERVATION OF BEHAVIOR

Systematic observation refers to a particular approach of quantifying behavior according to which trained observers record occurrences of specific behaviors in correspondence with a preciously defined coding scheme (Bakeman & Gottman, 1986). The goal of this approach is "for properly trained observers to produce identical protocols, given that they observed the same stream of behavior" – (Bakeman & Gottman, 1986). Often audio or video recordings are used as a basis for coding.

One of two central components of the method of Systematic Observation is the catalog of behavior codes or "Coding Scheme". Two approaches can be distinguished in developing a coding scheme: To (1) derive it from existing theory "top down", or (2) to develop it using a more grounded approach "bottom up", deriving it from the close observation of behavior itself (Weingart, Olekalns, & Smith, 2004). The line between these two approaches is necessarily blurred in practice and most approaches are somewhat hybrid, as their development often iterates between the formation of new codes from close observation and their comparison to existing theory. A final coding scheme has to both define how a behavioral is unitized, and then how the isolated units are classified. Some coding approaches start by unitizing the data and then categorize the units afterwards. This can even be done be separate observers. Other approaches use more simultaneous procedures for unitizing and categorizing behavioral streams. Finally coding schemes can be placed on a continuum between physically based and socially based coding schemes. According to Bakeman and Gottman (1986), physically based schemes are schemes with clear and well-understood roots in the organism's physiology, and socially based schemes are

schemes that deal with behavior whose very classification depends far more on ideas in the mind of the investigator (and others) than on mechanisms in the body.

The other central component of Systematic Observation is the demonstration of reliability of coders. Coder reliability refers to the level of agreement between two independent observers of the same stream of behavior. A common measure for the assessment of coder reliability is Cohen's Kappa (Cohen, 1960), which determines the level of agreement corrected for chance agreements, in comparison to the often used percent-agreement. Weingart and colleagues have distinguished between two types of reliability: Unitizing reliability and interpretive reliability (Weingart, et al., 2004). Unitizing reliability refers to the degree of agreement regarding identifications of units to be categorized and interpretive reliability refers to the degree of agreement in assigning the labels to the units.

Three prominent examples of coding systems for affect are the Facial Action Coding System (FACS) (Ekman & Friesen, 1978), the Specific Affect Coding System (SPAFF) (Coan & Gottman, 2007) and the Rapid Couples Interaction Scoring System (RCISS) (Krokoff, Gottman, & Hass, 1989). Of these three FACS could be characterized as a physically based scheme as it classifies behavior on the level of movement. FACS assigns numbered Action Units (AUs) to specific muscle movements in the face and a specific subset of AUs have been associated with emotional expression. SPAFF and RCISS, on the other hand, classify behavior on the level of the affective meaning that behavior has in a particular cultural context. They both make affect distinctions not only based on facial muscle movements, but also on changes in vocal tone, content, and body movement. Additionally, coding in RCISS and SPAFF relies on a cultural informant approach (Coan & Cottman, 2007), by which coders make categorizations on a gestalt level. These unique characteristics make RCISS and SPAFF very powerful coding schemes, but also very difficult to train and use.

LAB-SAMPLING OF TEAM BEHAVIOR

Most researchers studying interactions in design teams make a trade-off between studying "real" teams in the field, or "artificial" teams in the lab. The study of real teams offers advantages regarding external validity of results, but due to varying contexts the teams are observed in, it is often difficult to separate the phenomena of interest from the context they were observed in and to generalize the results obtained. The study of laboratory teams, on the other hand makes it easier to isolate phenomena of interest and produce generalizable results, but it is questionable how externally valid insights are that were gained with teams that only existed for the duration of a study. The approach used by Gottman and his colleagues in their studies with couples, combines advantages of both approaches. As it has been done with couples (Gottman & Levenson, 1992), I can bring "real" teams into a laboratory environment and observe their interaction behavior under controlled conditions. Even though teams are expected to behave differently in a lab context, I expect them to have different affective styles of interaction and that those styles will be observable in the lab. As with couples, I assume those styles to be indicative of a team's behavior in the field. To generate a sample of behavior that is reflective of a team's interaction style I modified a emotion elicitation task that was developed to elicit high arousal emotions in dyads (Roberts, Tsai, & Coan, 2007) for the use with teams. The specific tasks sets up a conflict discussion that "amplifies" the emotional characteristics of a team interaction so that differences between teams can be observed more clearly.

STUDY 1 – SOFTWARE ENGINEERING

INTRODUCTION: PAIR PROGRAMMING

The goal of my first study was to investigate the relationship between affective dynamics and performance in pair programming - a specific practice according to which two software engineers work together shoulder-to-shoulder at a single computer to solve a programming task. I was particularly interested in exploring the idea of affective balance as a critical indicator of team performance.

Studies of pair programming in university programming classes have shown that pair programming yields better design, more compact code, and fewer defects for roughly equivalent person-hours (Cockburn & Williams, 2000; Hanks, 2005; Nicolescu & Plummer, 2003; Williams & Kessler, 2002; Williams, Kessler, Cunningham, & Jeffries, 2000). Studies have also noted that pair programmers exhibit greater confidence in their code and more enjoyment of the programming process (Beck & Andres, 2004; Hanks, McDowell, Draper, & Krnjajic, 2004; Hanks, 2005; McDowell, Werner, Bullock, & Fernald, 2003; Nosek, 1998). Positive results with pair programming have lead to speculation that a collateral benefits of the practice may include improved morale and project knowledge shared efficiently across the development team in a manner that can be expected to improve productivity in subsequent development cycles (Beck & Andres, 2004). In spite of these compelling results, the adoption of pair programming has faced resistance and skepticism from both managers and programmers. While this may simply be a result of either the novelty of the practice or skepticism of the larger methodological context (Extreme Programming/Agile methods) in which pair programming is often introduced, there is some evidence that pair programming may not necessarily be appropriate for everyone (Dick & Zarnett, 2002). Furthermore there is evidence that demonstrates that pair programming can be less efficient than solo programming (Nawrocki & Wo-

jciechowski, 2001). Given these contradictions it is still an open question as to when and why pair programming works.

While past research focusing on pair programming interactions has looked at personality type influences (Hannay, Arisholm, Engvik, & Sjoberg, 2010; Salleh, Mendes, Grundy, & Burch, 2009; Sfetsos, Stamelos, Angelis, & Deligiannis, 2009; Walle & Hannay, 2009), and collaborative behavior such a keyboard switching (Chong & Hurlbutt, 2007), driving and navigating behavior (Chong & Hurlbutt, 2007; Chong, et al., 2005), and communication (Dick & Zarnett, 2002; Stapel, Knauss, Schneider, & Becker, 2010), no research up to date has focused on the quality of pair programming interactions by investigating affective interaction dynamics nor have methods been proposed to measure and study these dynamics.

The dyadic nature of pair programming lends itself to an investigation from a marital interaction perspective. This study was a first approach to the question of whether I can apply methods from the study of affect in marital interaction meaningfully to understand affective dynamics and performance in engineering interactions. Based on Gottman's Balance Theory of Marriage (Gottman, 1993) I hypothesized that:

> *H: The Group Hedonic Balance assessed from a thin slice of a pair programming interaction will be positively correlated with subjective and objective programming outcomes.*

METHOD

For the purpose of my investigation into the emotional dynamics of pair programming interactions I re-investigated a data-set that was collected in a previous laboratory study (Chong, et al., 2005; Leifer, Plummer, Chong, & Toye, 2008). The original study was designed to explore differences in programming performance between pair programming and solo programming and therefore compared the performance of 19 pair programmers with that of 9 solo programmers. For this investigation, I was only interested in the pair programming groups and their performance. The previous study had shown that the performance of the various pairs was spread out considerably but without any pattern that could be explained by the previously collected metrics. For example, performance did not improve with similarity of the task to normal work tasks, proficiency in Java, length of acquaintance with partner (possibly zero), or prior experience in pair programming with a specific partner. This made it particularly interesting for a detailed analysis of the differences in the affective quality of their interactions.

Materials and Measures

Video Data: Starting point for my analysis were 16 of the originally 19 hi-8 videotapes that were recorded during the original study. Two video records were missing at the time of my analysis and one additional pair had to be dropped because they misinterpreted the task. A video record always showed one pair programming team engaging in a short (ranging between 30 and 90 minutes) warm-up task followed by the longer calendaring task to develop a java program to find the best meeting times for a group of required and optional participants. Each pair was working on a single computer with two large flat-panel monitors, a single keyboard, and a single mouse.

The videos were recorded in a quad-view format, showing the pair from four different angles, resulting in a very low video resolution per stream. All subjects in the study were professional programmers, who were paid near-industry-standard rates for their time. Some of the programmers had known each other previously and some had not. The average time for a team to finish the calendaring task was 5 hours

and 2 minutes ranging from 2 hours and 23 minutes for the fastest pair to 7 hours and 55 minutes for the slowest pair.

TABLE 3.1
Participant Statistics for Pair Programming Study.

	Age	Professional Experience in years	Years to know Partner
M	27.6	5.05	3.38
SD	3.5	3.17	3.77
Minimum	23.5	1.75	0.0
Maximum	37.0	13.5	10.0

N = 16

Performance Measures: I used two subjective and two objective performance measures that had been generated as part of the original study. The two subjective measures were (1) satisfaction with the programming experience, and (2) satisfaction with the developed code. Both measures were operationalized using one Likert scale survey item ranging from 1 (low) to 7 (high). The two objective measures were (3) duration to solve the calendaring task, and (4) code performance. Code performance was operationalized by scoring the developed code. The scoring procedure consisted of running the subjects' code on a comprehensive suite of 19 test cases that exercised all of the features they were meant to implement. Scores were determined by awarding one point for each test passed, resulting in a scale of 0 to 19 points. The average score for all pairs was 11.5, ranging from 0 points for the lowest performing pairs to 19 points for the highest performing pairs.

Table 3.2
Performance measures: Pearson correlations and descriptive statistics.

	(1)	(2)	(3)	(4)
Outcome Variables				
(1) Code Performance [0-19]	–	-.59*	.71**	.68**
(2) Duration [min]		–	-.07	-.60*
(3) Feelings about Code [1-7]			–	.79**
(4) Feelings about Experience [1-7]				–
Descriptives				
M	12.27	289.33	4.3	5.43
SD	7.52	98.96	1.53	0.92
Minimum	0	143	1.5	4
Maximum	19	447	7	7

Note: All correlations are Pearson correlations.
N = 16
*p < .05, **p < .01 (two-tailed)

Procedures

As an overall procedure for coding the occurrence of affective behavior in the videos, I combined a thin-slices approach with that of systematic observation of behavior.

Sampling of 5-Minute Slices: I extracted five-minute "slices" from each of the 16 video clips inspired by studies on dyadic negotiation interaction (Curhan & Pentland, 2007) and marital interactions (Carrere & Gottman, 1999). I always selected the first five minutes of each video showing the pairs working on the calendaring task for a detailed affect-interaction analysis (See Figure 3.1).

Five minutes segments were chosen because of what was thought to be a reasonable compromise between accuracy of the analysis and effort to code the videos. This choice was supported by Ambady and Rosenthal's (1992) meta-analysis on thin slicing. They found that that assessments made based on just 30 seconds did not vary significantly from those made from samples that were 5 minutes long. It is especially encouraging that they could show strong correlations across 38 different interaction contexts. Finally, their study about teacher evaluations showed that a 30-second sample is enough to even assess an interaction that lasts several months thus giving me confidence that a five minute sample is more than enough to assess a pair's interaction style (Ambady & Rosenthal, 1993). The first five minutes were chosen for several reasons. (1) When looking at interactions qualitatively from start to finish, I observed, that the general interaction style did not change over time. This notion is also supported by the thin slicing literature that shows that accurate predictions are possible independent of the time the sample is taken at. (2) In research on marital interactions, it has been shown that particularly the startup phase of an interaction can give valuable insights and higher predictive accuracy. Therefore, even though any five-minute sample would have probably given me sufficient insight of the interaction style of a pair, this research suggests that the first 5 minutes are particularly useful in making accurate assessments.

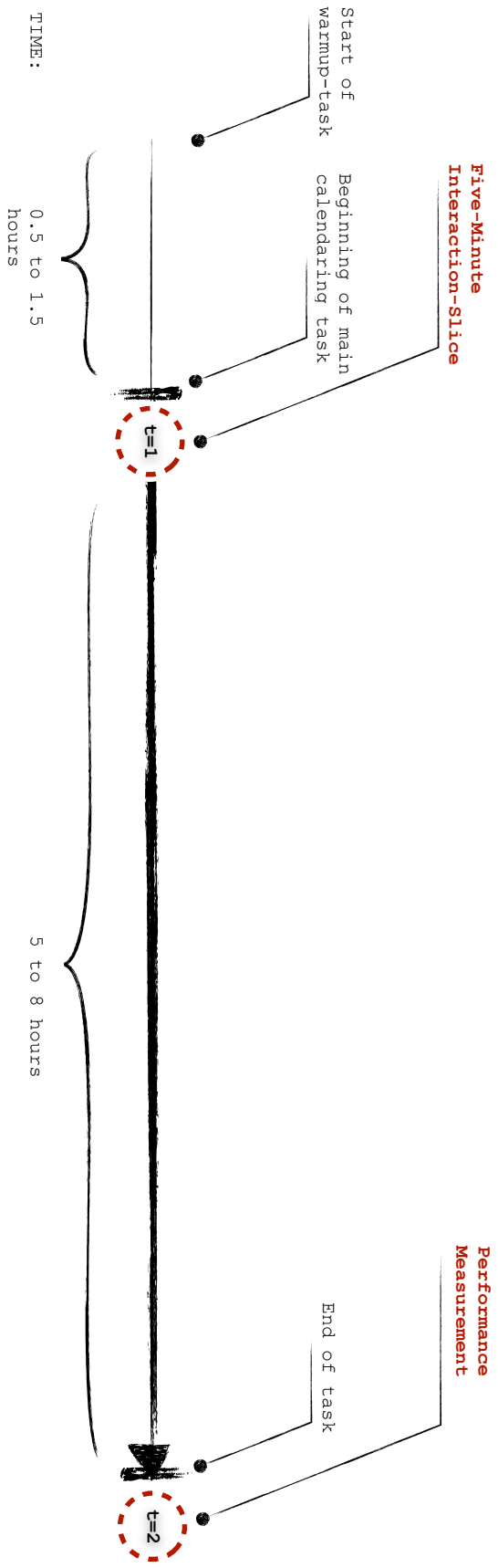


FIGURE 3.1
Pair programming study: Procedure overview.

Coding of the Selected Samples: For systematic observation, I used a slightly adapted version of the Rapid Couples Interaction Scoring System (RCISS) (Krokoff, Gottman, & Hass, 1989). I chose RCISS for two reasons: First, it is faster than the Specific Affect Coding System (SPAFF), and second, it is more granular in capturing minute differences in listener behavior. The modifications to the original coding scheme were necessary to adapt the coding system so that it would capture the affective differences between the various pairs. Couples interactions such as the ones that were used to develop RCISS, show huge differences in expressive behavior because they were generated in the context of a highly engaging conflict discussion. The pair programming teams, on the other hand, were not set out to be in conflict, and they did not share a long history together and thus affective differences in between them were much more subtle. For example expressions of contemptuous or frustrated behavior are common in marital conflict interactions but could hardly be observed during interactions of programming pairs. The adapted RCISS used for this study consisted of 13 speaker codes and 10 listener codes (see table 3.3).

	Speaker Codes	Listener Codes
	Interest	Backchannels present
	Validation	Facial movement present
	Excitement	Looking at each other
	Humor/Laugh	Positive facial expression
	Empathy	Responsive facial movement
	Other Positive	
	Complain	No backchannels
	Criticize	No facial movement
	Constrained Anger	No looking at each other
	Defensive	Negative facial expression
	Yes-but	Stonewalling
	Put down	
	Tension/Tense humor	
	Other negative	

TABLE 3.3
Modified RCISS coding scheme used in the pair-programming study.

The coding itself was performed in two steps using VCode, a behavioral coding software (Hagedorn, Hailpern, & Karahalios, 2008). In a first step each video was segmented into speaker-turns (81 speaker turns on average). The second step then consisted of watching one speaker-turn at a time and checking all codes that applied from the list of available RCISS codes. An average 5-minute segment had a total of 146 codes with 22 speaker codes and 124 listener codes. Figure 3.2 shows a 16-second window of a coded video. I initially coded all videos, and 50% of the videos were randomly selected and coded by a second observer. It took approximately 20 minutes to code speaker turns for each video in a first pass and 30 minutes to do the emotion coding using RCISS for each video in the second pass. At the time the videos were coded, both observers were blind to the performance of the teams. Additional data was only accessed after the coding was completed.

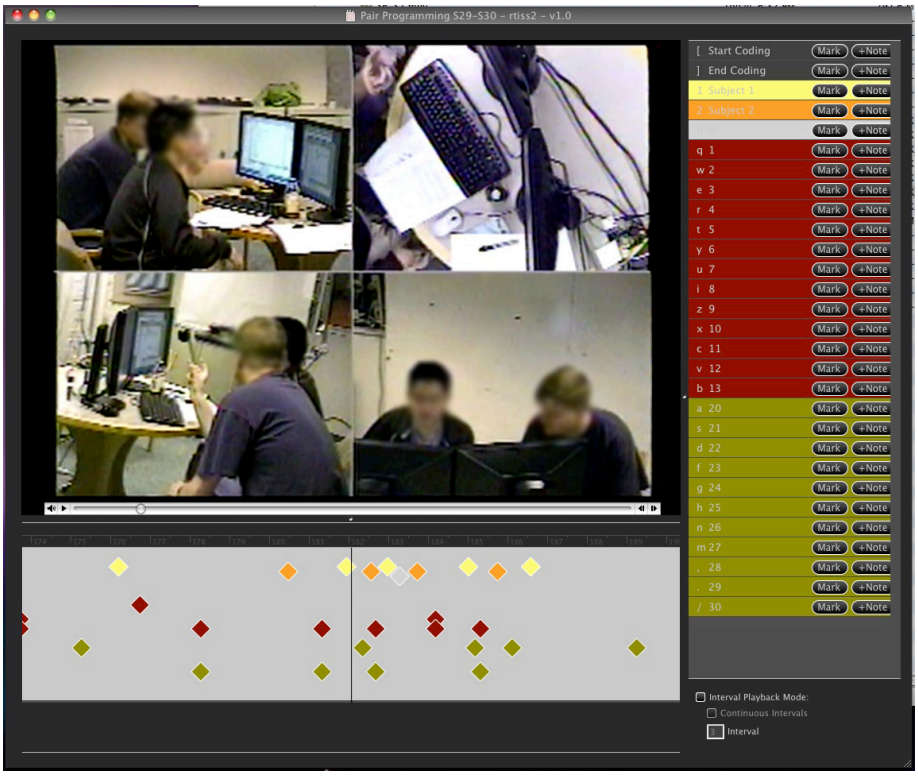


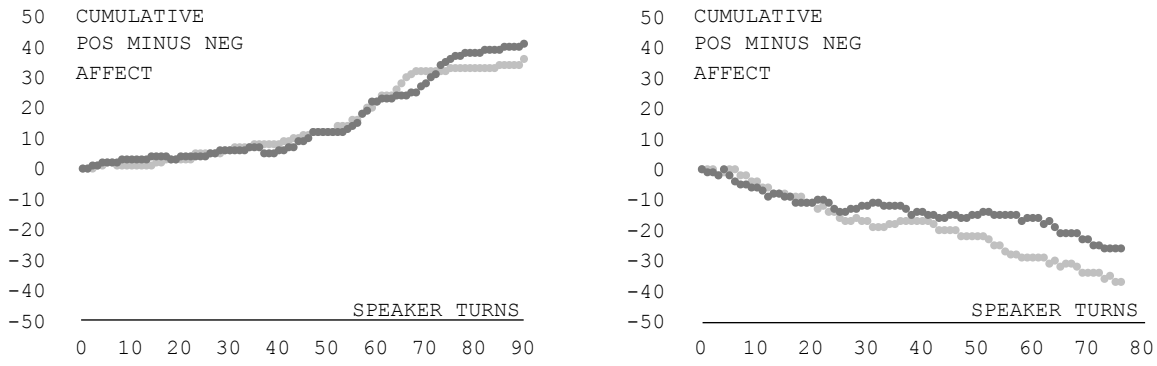
FIGURE 3.2
VCode interface showing a 16-second section coded in RCISS. The orange and yellow squares in the upper rows indicate speaker turns. The red and green squares in the lower part mark occurrences of negative and positive behaviors respectively.

RESULTS

The average Kappa (Cohen, 1960) for the modified RCISS coding was computed using GSEQ, a software for sequential data analysis (Bakeman & Quera, 1995). The coder reliability analysis for the eight double-coded videos revealed a Kappa of κ=0.50 for the speaker codes and a Kappa of κ=0.44 for the listener codes. According to Altman (Altman, 1991) and Landis (Landis & Koch, 1977), these are considered moderate levels of agreement. From the videos that were coded twice, I randomly selected one file to be included in the analysis.

To analyze the data I compared measures of affective balance of each pair with the respective performance relevant outcomes. Based on the RCISS codes, I classified the programming pairs into five regulated (R) and eleven non-regulated (NR) pairs as the main predictor variable. In studies of marital interaction this distinction was highly predictive of subjective (e.g. marital satisfaction) as well as objective (e.g. divorce) marital outcomes (Gottman, 1994). For each conversational turn the total number of positive RCISS codes minus the total number of negative codes was computed for each programmer. Then the cumulative total of the resulting scores were plotted over the number of speaker turns for each programmer.

FIGURE 3.3
Point graphs of a regulated and a non-regulated programming pair. The graphs always show the emotion trajectories for each programmer separately. The left graph is drawn from a pair that scored amongst the lowest in the sample and the right graph is drawn from a pair that scored amongst the highest of the pairs studied.



I determined the slopes of each point graph using linear regression analysis. The average r-squared for a linear fit to the point graphs was 0.87, indicating a stable trend in each pair's affective style. As described in (Gottman & Levenson, 1992), I then classified a pair as regulated if both slopes were significantly positive and non-regulated if they did not both have significant slopes. Figure 3.3 shows an example of a regulated and a non-regulated programming pair. Additionally I calculated a second continuous predictor variable by taking the mean slope of the two regression lines as a measure of the group hedonic balance.As subjective outcome variables I used the satisfaction with the code and the satisfaction with the programming experience. As objective outcome variables I used the Code Performance Score and the Duration. A t-test was performed to test the relationship between the binary (regulated vs non-regulated) predictor variable and the outcome variables. Pearson correlations were used to test the relationship between group hedonic balance and performance. Because the Code Performance score, $D(17) = 0.25$, $p < .05$, was significantly non-normal, I examined Kendall's tau correlation coefficients to test the relationship between group hedonic balance and Code Performance.

My hypothesis was that group hedonic balance would be predictive of performance-relevant outcomes and indeed the hedonic balance based classification into regulated and non-regulated pairs was predictive of several performance relevant outcomes.

Subjective Performance:

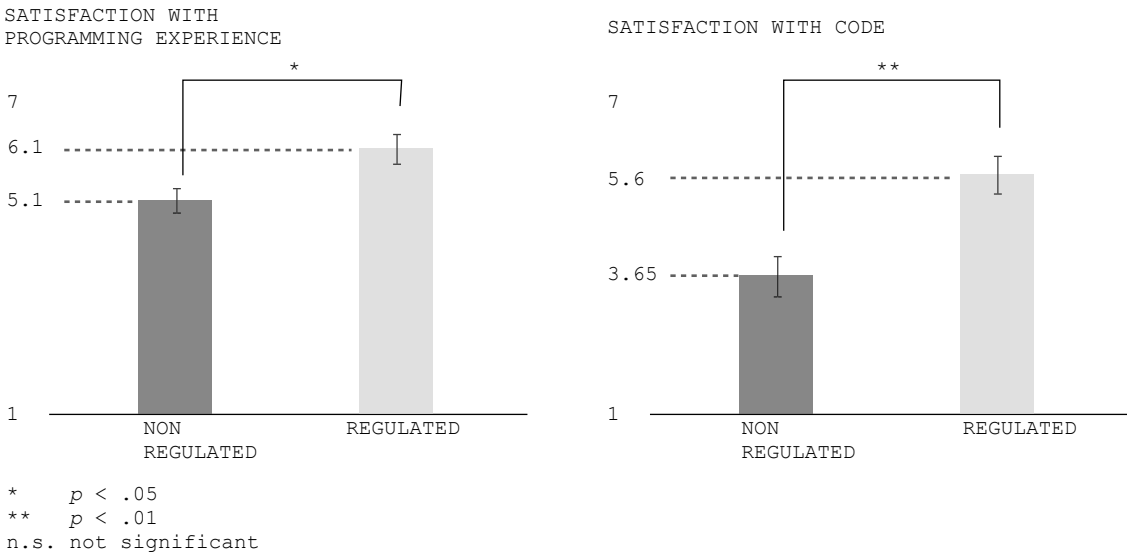
On average, regulated pairs (M = 6.10, SE = 0.33) were significantly more satisfied with the overall programming experience than the non-regulated pairs (M = 5.10, SE = 0.27), $t(13) = 2.25$, $p < 0.05$ (one-tailed), $r = .53$. In addition, regulated pairs (M = 5.60, SE = 0.40) were significantly more satisfied with the developed code than the non-regulated pairs (M = 3.65, SE = 0.43), $t(13) = 2.85$, $p < 0.01$ (one-tailed), $r = .62$.

I also noted that the continuous Group Hedonic Balance score was significantly related to the satisfaction with the developed code, $r = .73$, p (one-tailed) < .01, and to the satisfaction with the overall programming experience $r = .63$, p (one-tailed) < .01.

Objective Performance:

In terms of the code performance scores, regulated pairs (M = 18.20, SE = 0.58) wrote, on average, significantly better code than the non-regulated pairs (M = 10.18, SE = 2.35), $t(14) = 2.25$, $p < 0.01$ (one-tailed), $r = .66$. In addition, regulated pairs (M = 235.20, SE = 20.01) took, on average, less time to finish the task than the non-regulated pairs (M = 317.55, SE = 31.23). This difference was not significant, $t(14) = -1.68$, $p > 0.05$ (one-tailed); however it did represent a medium-sized effect $r = .41$.

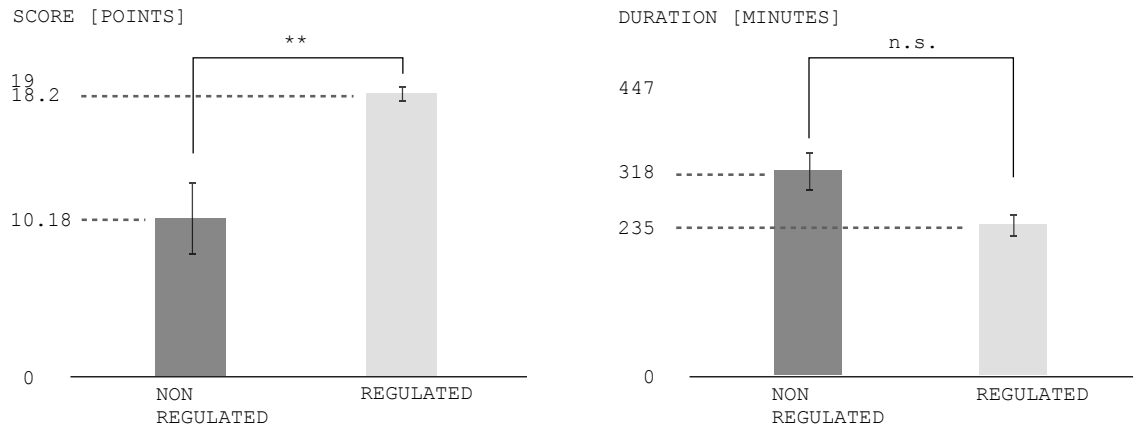
FIGURE 3.4
Bar charts comparing subjective performance between regulated and non regulated pairs.The figures above show the differences in subjective performance measures for regulated vs. non-regulated programming pairs. All error bars are +- 1 standard error.



* $p < .05$
** $p < .01$
n.s. not significant

I also noted that the continuous Group Hedonic Balance score was significantly related to the Code Performance, $\tau = .40$, p (one-tailed) < .05. Group Hedonic Balance was not significantly related to the Duration it took a pair to complete the task. However, pairs that scored high in terms of their code performance also took significantly less time to complete the task, $\tau = -.40$, p (one-tailed) < .05. As can be seen from the error bars, the variance of scores in each group (regulated vs. non-regulated) varied greatly. While the regulated pairs all performed well, the performance of the non-regulated pairs covered a wider range of higher and lower scores. The large variance in scores is in line with the data collected in the original study that compared single programmers with pairs. The study showed that in this task pairs did not show any significant performance advantages over solo programmers. Therefore a possible explanation might be that in the non-regulated pairs, one person might have just developed the solution alone. In one of the non-regulated pairs for example that still reached the highest possible code performance score, a highly experienced programmer took on most of the work by himself, leaving his partner to just watch most of the time. This explanation would suggest that in more complex and interdependent tasks less of the non-regulated pairs would be able to perform that well. Further research on affective dynamics should therefore attempt to create task settings that are complex enough, highly relevant, and engaging.

FIGURE 3.5
Bar charts comparing objective performance between regulated and non regulated pairsThe figures above show the differences in objective performance measures for regulated vs. non-regulated programming pairs. All error bars are +- 1 standard error.



* $p < .05$
** $p < .01$
n.s. not significant

Exploratory Analyses and Results

In addition to testing the correlation between Group Hedonic Balance and performance, I performed an exploratory analysis to generate insights into specific affect behaviors and their relationship with pair programming performance. For each of the speaker and listener code categories, I calculated the average number of codes per turn and compared it with the subjective and objective performance measures (see table below). Additionally I explored correlations between the number of speaker turns and outcomes. Due to the exploratory nature of this analysis, correlations at the 10 percent level are flagged as significant as well.

The number of speaker turns in the five-minute slice was negatively correlated with the time it took for a pair to finish the programming task, $r = -.61$, $p < .05$, (two-tailed). Several positive and negative speaker behaviors were significantly correlated with subjective and objective performance metrics as shown in the table below. The strongest correlate with performance among the listening behaviors was the mere presence or absence of any listening behavior. The presence of back-channeling behavior (BCP) to take one, was significantly correlated with objective code performance ($r = .47$, $p < .05$, two-tailed), programming duration ($r = -.53$, $p < .05$, two-tailed), satisfaction with the code ($r = .69$, $p < .01$, two-tailed), and satisfaction with the overall programming experience ($r = .66$, $p < .01$, two-tailed).

TABLE 3.4
Results of the exploratory analysis for the pair-programming study.

	Objective Performance		Subjective Performance	
	Code Performance	Duration	Satisfaction with Code	Satisfaction with Experience
Interactivity				
Speaker Turns	.26	**-.61***	.14	.24
Positive Speaker Behaviors (Counts per Turn)				
Interest	-.16	.40	-.06	-.03
Validation	.08	.19	.21	.03
Excitement	**.41***	-.31	**.50†**	.43
Humor	**.43***	-.24	.31	.44
Negative Speaker Behaviors				
Defensiveness	-.02	-.04	-.39	-.41
Yes-But	-.12	.10	.05	-.36
Put-Down	**-.40***	.24	-.17	**-.56***
Complaint	-.20	.37	-.39	-.32
Other Negative	.08	-.37	.04	.02
Tension	-.10	-.10	-.04	.02
Constrained Anger	-.27	.38	-.31	**-.45†**
Positive Listener Behaviors				
BCP	**.47***	**-.53***	**.69****	**.66****
FMP	.13	-.11	**.48†**	.21
LSP	.10	-.01	**.55***	.42
PFE	-.29	.21	**-.57***	-.26
RFP	.05	-.12	.18	.28
Negative Listener Behaviors				
NBC	**-.42***	**.50***	**-.71****	**-.64***
NFM	-.05	-.11	.21	-.15
NLS	-.10	-.10	.21	-.15
NFE	-.27	-.10	-.41	-.20
STW	-.20	.00	-.43	-.23
Aggregated Behaviors				
Positive Speaker Behaviors	**.38***	.03	**.46***	.28
Positive Listener Behaviors	**.34†**	-.27	**.58****	**.54****
Positive Behaviors	**.38***	.03	**.46***	.28
Negative Speaker Behaviors	-.15	.03	-.12	-.18
Negative Listener Behaviors	**-.38***	.25	**-.56****	**-.46***
Negative Behaviors	-.15	.03	-.10	-.14

Note: All correlations are Pearson correlations, only correlations with Code Performance (First Column) are Kendall's Tau correlations. The codes "Criticism" and "Empathy" are dropped because no occurrences were observed in the data. Variables are mean counts of behavior occurrences per turn.
Negative Listener Codes: NBC = no backchannel behavior, NFM = no facial movement, NLS = no looking at speaker, NFE = no facial expression, STW = stonewalling. Positive Listener Codes: BCP = backchannel behavior present, FMP = facial movement present, LSP = looking at speaker, PFE = positive facial expression, RFP = responsive facial movement.
N = 16
$^{\dagger}p < .1$, $^{*}p < .05$, $^{**}p < .01$ (two-tailed)

DISCUSSION

With this study, I introduced the Group Hedonic Balance a useful metric for describing the affective quality of pair programming interactions. The group hedonic balance assessed during the first five minutes of a software engineering interaction significantly correlated not only with subjective performance measures but also with the objective quality of the developed code. In addition, I found that it is possible and meaningful to draw a distinction between regulated and non-regulated programming pairs based on differences in their affective interaction dynamics. This classification of programming teams into regulated and non-regulated pairs was predictive of subjective as well as objective performance measures. The usefulness of the distinction between regulated and non-regulated pairs is also supported by qualitative observations. What characterized non-regulated pairs were generally low levels of energy, with little interactivity, evidence of non-listening, and a constant presence of friction between the programmers. There was rarely any humor, or curiosity in those pairs. Regulated pairs, on the other hand, exhibited high levels of energy and interactivity, participants often laughed or told jokes, and there were ample signs that the participants were listening to each other and visibly seemed to enjoy working with each other. A limitation of this study is that the resulting sample size of 16 pairs can be considered low. However, a meta analysis by Hannay and colleagues (Hannay, Dyba, Arisholm, & Sjoberg, 2009), reviewing 18 empirical studies of pair programming, revealed that 17 pairs is a relatively large sample in comparison with current studies on pair programming: The average sample size among the studies investigated was only 18 pairs, with only six studies having used larger samples than the 16 I used here.

In sum, this study was successful in taking methods from the study of affect in marital interactions and applying them to gain insight into technical work. A positive Group Hedonic Balance is not only an important indicator for stable marriages but also a critical indicator of successful engineering-teams.

STUDY 2 – NEGOTIATION

INTRODUCTION: DISPUTE RESOLUTION

The goal of this study was to explore the relationship between affective interaction dynamics and team performance and specifically to examine the relationship between Group Hedonic Balance and performance by focusing on a central activity in engineering design: negotiation. Negotiations play a crucial role in the interactions of engineering teams (Bucciarelli, 1988, 2002) and affective interaction dynamics influence their process and outcome (Curhan & Pentland, 2007). Even though most negotiations involve two or more persons interacting, the majority of past research on negotiation and emotion has been dominated by a focus on the intrapersonal relationships between affect and cognition (Morris & Keltner, 2000). In this line of research people have explored things such as the role of affect on problem-solving, decision making, or risk-taking (Isen & Baron, 1991; Johnson & Tversky, 1983). It is only recently that the focus has partially shifted towards an interest in the interpersonal and social functions of emotions (Morris & Keltner, 2000; Van Kleef, 2007; Van Kleef, De Dreu, & Manstead, 2004). Most studies focusing on the social functions of emotions have looked at a few specific emotions such as anger or happiness and their isolated effects on the opposite party in a negotiation (Van Kleef, De Dreu, & Manstead, 2004). However to the author's knowledge, at the point of this study nothing has been done to study the role of a negotiation pair's balance of positive and negative affect as a correlate of negotiation performance. I therefore hypothesized that:

> *H: The Group Hedonic Balance assessed from a thin slice of a two-party negotiation will be positively correlated with subjective and objective negotiation outcomes.*

In addition to exploring this hypothesis, this study built on the previous study in several ways: First, I wanted to increase the number of subjects to allow for more statistically robust conclusions. Second, I wanted to include more detailed and established subjective measures of performance. Third, because the pair programming interactions were so low in their expressiveness of emotions I wanted to select a task context that would allow for more emotionally engaging interactions. Finally I wanted to generate higher quality videos, to be able to look at facial expressions, which had not been possible in the previous study.

METHOD

Participants

Participants were drawn in pairs from a graduate level negotiation class at Stanford University. In this class students are introduced to a particular negotiation framework that originates from the Program of Negotiation at Harvard University. In the class students have to negotiate cases every week as homework assignments and the experience and insights gained from the cases are then discussed later in class. For this study I recruited students by asking them to negotiate their homework case, a two party dispute resolution, in a laboratory environment while being videotaped, so that their negotiation process would be available for detailed analysis. I recruited students from a specific negotiation case that was chosen for its task qualities which I will describe later and because the case was given at an early stage of the class (3rd week) and I was hoping that at that point the participants would not have converged onto the specific negotiation style taught in the class. 52 students (26 pairs) agreed that their negotiation interaction could be recorded as part of a study. Participants ranged in age from 19 to 32 ($M = 24.2$, $SD = 3.4$). Detailed participant statistics are displayed in the table below. Participants were randomly grouped into pairs resulting in 4 all female pairs (15.4%), 7 all male pairs (26.9%), and 15 mixed pairs (57.7%).

		Frequency	Percent
Sex	**female**	23	44.2
	male	29	55.8
	total	52	100
Ethnicity	**Caucasian**	33	63.5
	Asian-American	12	23.1
	Hispanic/Latino-American	2	3.8
	Other / Multi-Ethnic	5	9.6
	total	52	100
Major	**Engineering**	29	55.8
	Business	8	15.4
	Other	15	28.8
	total	52	100

TABLE 4.1
Participant statistics for negotiation study.

Materials and Measures

Rapid Couples Interaction Scoring System (RCISS): As in the previous study, I used an adapted version of the Rapid Couples Interaction Scoring System (RCISS) to quantify affective behavior (Krokoff, Gottman, & Hass, 1989). This choice again was driven by the amount of time that would be required to code a video and by the high granularity in capturing listener behaviors such as whether the listener looks at the person who is speaking, or nods.

TABLE 4.2
Modified RCISS coding scheme used in the negotiation study

	☹	☺
Speaker Codes	Negative Task Talk	Neutral or Positive Task Talk
	Tension	Validation
	Tense Humor	Excitement
	Constrained Anger	Humor / Laugh
	Defensiveness	Affection
	Yes-But	
	Criticize	
	Put Down (Domineering / Contempt / Belligerence)	
Listener Codes	No Backchannels	Backchannels Present
	No Facial Movement	Facial Movement Present
	No Looking at Speaker	Looking at Speaker
	Negative Facial Expression	Positive Facial Expression
	Stonewalling	Responsive Facial Movement

Performance: I assessed subjective and objective measures of group performance after the conclusion of each negotiation. As the main subjective performance measure I used the Subjective Value Inventory (SVI) (Curhan, Elfenbein, Xu, Cambridge, & Drive, 2006) and averaged the scores for each pair. The Subjective Value Inventory is a survey instrument to assess the subjective value of a negotiation. The overall subjective value (SVI Global) is comprised of four factors, namely feelings about the instrumental outcome (SVI Instrumental), feelings about the self (SVI Self), feelings about the negotiation process (SVI Process), and feelings about the relationship (SVI Relationship). As objective performance measures I assessed the duration it took for a group to conclude the negotiation, the monetary agreement each group made, and the complexity of an agreement as operationalized by counting the number of additional agreements beyond the monetary one.

Procedure

This study was conducted in the context of a graduate level negotiation class at Stanford University. The task was a dispute resolution about the cost of an auto repair (Patton, 1982):

> *Dr. Susan Garfield has a billing dispute with John Eazer, owner of a local garage, over some repair work done on Garfield's car. Finding the bill significantly higher than the original informal estimate, Garfield angrily confronted Eazer. Eazer prepared a second bill at an even higher figure. Frustrated, Garfield returned to the garage after closing time with a spare key and drove her car home, without paying anything. Eazer turned to his child-in-law, an attorney, wishing to file a criminal complaint. When phoned, Garfield referred the attorney to her father, a senior partner in a local law firm. Garfield's father is letting one of his young associates handle the case.*

This task was selected because each participant takes on the role of a representative (attorney). It was assumed that this would make it easier to identify with the given role and therefore create a more emotionally engaged interaction. Additionally this case was chosen because there were no status differences between roles. Participants were randomly formed into one of 26 dyads. Within each dyad it was then determined at random who would represent which role (Garfield's attorney or Eazer's attorney) and then these specific pairings and role assignments were presented to the students in class. Each participant received the specific instructions for their role in the dispute resolution one week prior to the negotiation. Certain aspects of the case were shared between participants, others were confidential.

Negotiation: Students arrived at the lab, were greeted by the experimenter, escorted into a small room, and seated in chairs across a round table facing each other. A microphone was placed on the table and two cameras were placed so that they recorded each person's upper body and face and an additional camera recorded the entire scene. After giving general instructions and obtaining participants' consent to be audio and videotaped, the experimenter started the video recording equipment and

asked the participants to start their negotiation. No guidance was given as to how to conduct the negotiations and no time-limit was set for finding a conclusion. Immediately after finishing the negotiation, the experimenter entered the room and asked the participants to fill out a paper version of the subjective value inventory (Curhan, Elfenbein, Xu, & Drive, 2006) which asked them to report on their feelings about the instrumental outcome (e.g. "How satisfied are you with your own outcome?"), feelings about the self (e.g. "Did you 'lose face' in the negotiation?"), feelings about the negotiation process (e.g. "Did you characterize the negotiation process as fair?'), and feelings about the relationship (e.g. "Did the negotiation make you trust your counterpart?").

Systematic Observation of Behavior: From each of the 26 the video-recordings I extracted a five-minute slice at five minutes into the negotiation. The videos were then loaded into VCode (Hagedorn, Hailpern, & Karahalios, 2008), and coded according to the modified RCISS coding scheme introduced in the previous section. Two coders who were blind to the hypotheses and to the outcomes of the negotiations coded the videos. Eight videos (30%) were double coded to assess inter-rater reliability.

The method described here allowed me to collect data about a group's affective interaction dynamics occurring in a thin slice early during their interaction process and about their subjective and objective performance at the end of their process. The goal was to get data that would allow me to test my primary hypothesis and that would allow me to explore other possible correlations between affective interaction characteristics and performance.

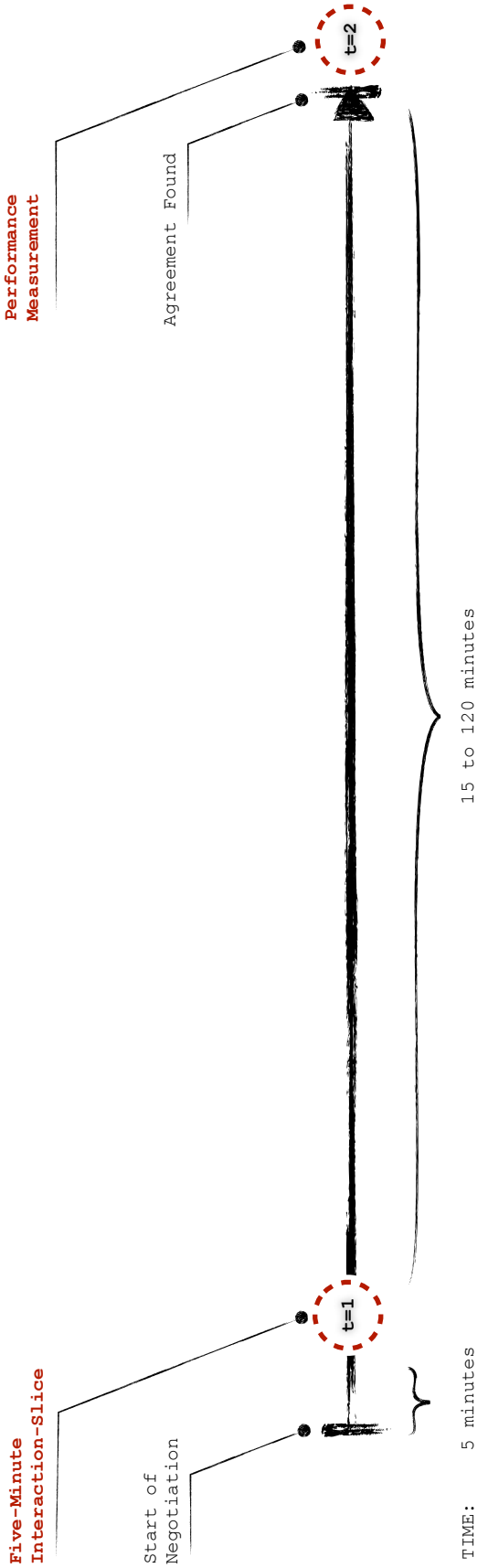


FIGURE 4.1
Negotiation study: Procedure overview

RESULTS

None of the groups had to be dropped from the analysis. Each group came to an agreement but the time it took each group to conclude their negotiation varied greatly between 14 to 122 minutes (M = 55.0, SD = 23.9). The groups agreed on monetary settlements between $2282 and $3102 ($M$ = 2739.5, SD = 233.5) and the level of detail of their agreement varied between no additional agreements and 4 additional agreements such as free oil changes for further repairs (M = 1.0, SD = 1.2). In a direct comparison with the pair programming study, the specific study setup seemed to have been successful in setting up more emotionally engaging interactions. Also the quality of the video recordings was greatly increased over the first study which made it possible to identify facial expressions more easily.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Subjective Outcomes							
(1) SVI Instrumental [1-7]	–	.45*	.68***	.67***	.81***	-.29	.12
(2) SVI Self [1-7]		–	.69***	.64***	.78***	-.28	.08
(3) SVI Process [1-7]			–	.84***	.93***	-.40*	-.07
(4) SVI Relationship [1-7]				–	.92***	-.30	.15
(5) SVI Global [1-7]					–	-.37	.12
Objective Outcomes							
(6) Duration [min]						–	.36*
(7) Agreement Complexity [# of items]							–
Descriptives							
M	5.48	5.61	5.49	5.74	5.58	55.0	1.04
SD	0.66	0.58	0.76	0.81	0.61	23.9	1.18
Minimum	4.38	4.38	3.75	3.63	4.16	14	0
Maximum	6.88	6.75	6.63	6.88	6.53	122	4

TABLE 4.3
Correlations and Descriptives for subjective and objective performance metrics.

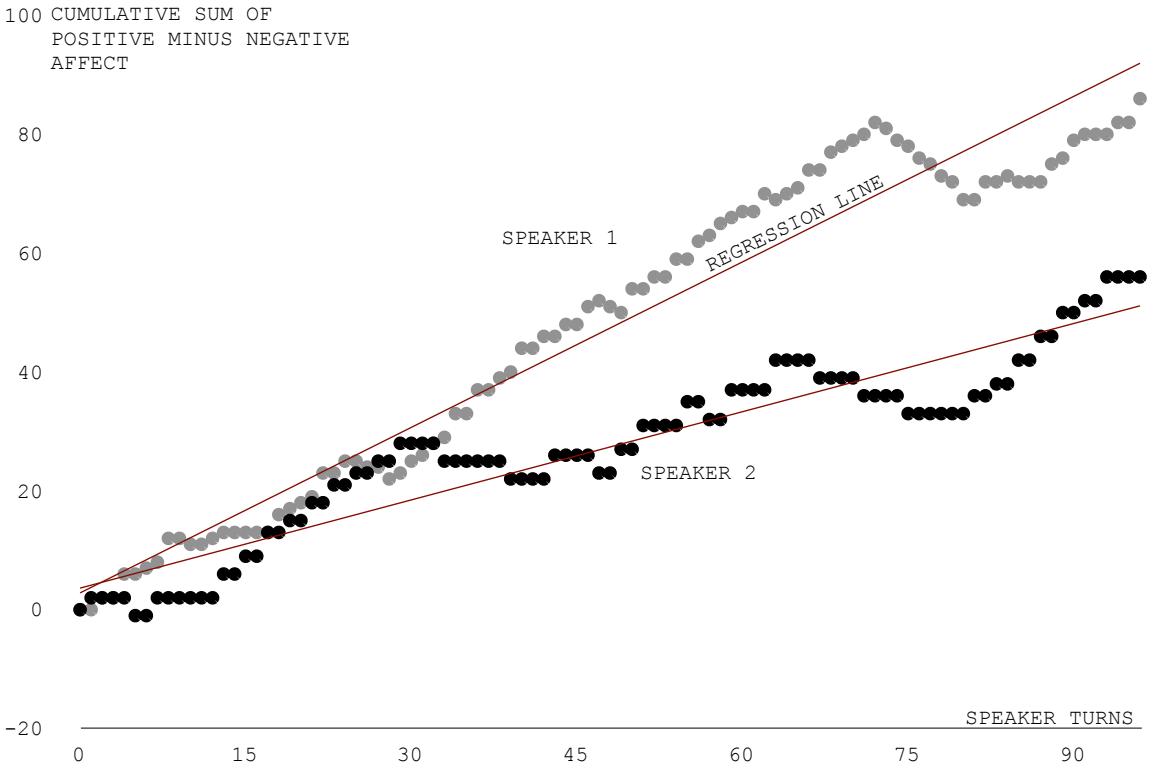
Note: All correlations are Pearson correlations except correlations in the last column (8) are Kendal's tau correlations.
N = 26
*p < .05, **p < .01, ***p < .001 (two-tailed)

Coder agreement for the modified RCISS coding was assessed using Cohen's Kappa (Cohen, 1960) and ranged between $\kappa = 0.28$ and $\kappa = 0.58$ for the listener codes (M = 0.42, SD = 0.1) and between $\kappa = 0$ and $\kappa = 0.49$ for the speaker codes (M = 0.2, SD = 0.22). According to (Landis & Koch, 1977) this is a moderate level of

agreement for the listener codes. Due to the extremely low level of agreement for the speaker codes, the speaker codes were dropped from the analysis and only the listener codes were used to construct the Group Hedonic Balance score.

To construct the measure for a group's hedonic balance, I adapted Gottman's method of using slopes of point graphs as indicators of a couple's affective interaction quality (Gottman, 1994): The Group Hedonic Balance measure was therefore operationalized by (1) taking a five minute video slice from the beginning of the negotiation, (2) coding each interaction with an adapted version of the Rapid Couples Interaction Scoring System (RCISS) (Jung, Chong, & Leifer, 2010; Krokoff, et al., 1989), (3) plotting point-graphs displaying the cumulative positive minus negative affect over the speaker turns for each individual according to the procedures outlined in (Gottman, 1994), (4) assessing the slope of each point graph through linear regression, and (5) averaging the point-graph slopes for each negotiation pair. Group Hedonic Balance ranged between 0.41 and 1.3 ($M = 0.74$, $SD = 0.24$) all being positive.

FIGURE 4.2
Point graphs for two participants in a negotiation interaction.



Since the hypotheses are stated at the group level, the group mean was taken for each variable. Pearson correlations were used to test correlations between Group Hedonic Balance and all performance variables except detail score. Because the detail score score, $D(26) = 0.24$, $p < .001$, was significantly non-normal, I examined Kendall's tau correlation coefficients for the relationship between Group Hedonic Balance and detail scores.

I expected that the affective interaction dynamics would be correlated with performance relevant outcomes. Specifically I was asking whether the joint performance of negotiation pairs is higher for pairs with a high Group Hedonic Balance score than for those with a low Group Hedonic Balance score.

Table 4.4 shows that there is only very limited support for the hypothesis that relates Group Hedonic Balance with performance. Group Hedonic Balance correlated significantly in the predicted way only with participants feelings about the negotiation process, $r = .35$, p (one-tailed) $< .05$. However Group Hedonic Balance also correlated significantly negative with the level of detail in the agreement, thus disconfirming the hypothesis, $\tau = -.32$, p (one-tailed) $< .05$.

TABLE 4.5
Results of the exploratory analysis for the negotiation study.

	SVI Instrumental	SVI Self	SVI Process	SVI Relationship	SVI Global	Duration	Complexity of Agreement
Group Hedonic Balance	-.39	.21	.35*	.17	.20	-.03	-.32*

Note: All correlations are Pearson correlations. The correlation with Complexity of Agreement was tested with Kendal's tau.
N = 26
*$p < .05$, (one-tailed)

Results from Exploratory Analyses

In addition to testing the correlation between Group Hedonic Balance and performance, an exploratory analysis was performed to generate insights into specific affect behaviors and their relationship with performance relevant outcomes. For each of the 5 positive and 5 negative listener code categories, I calculated the average number of codes per turn and compared it with the subjective and objective performance measures (see table below). Additionally I explored correlations between the number of speaker turns and outcomes. Due to the exploratory nature of this analysis, correlations at the 10 percent level are flagged as significant as well.

TABLE 4.4
Results for correlation analysis between Group Hedonic Balance and negotiation performance.

	Subjective Performance					Objective Performance	
	SVI Instrumental	SVI Self	SVI Process	SVI Relationship	SVI Global	Duration	Agreement Complexity
Interactivity							
Speaker Turns	-.06	.33	.08	.13	.13	-.03	.26†
Negative Listener Count per turn							
NBC	-.01	-.21	-.03	-.02	-.07	.13	-.09
NFM	.32	-.22	.04	.23	.12	-.14	.14
NLS	-.09	-.04	-.19	-.20	-.16	.16	.12
NFE	-.42*	.00	-.16	-.18	-.22	.06	-.06
STW	–.–	–.–	–.–	–.–	–.–	–.–	–.–
All NEG Codes	-.18	-.21	-.22	-.15	-.22	.11	-.04
Positive Listener Count per Turn							
BCP	-.19	.08	.16	.03	.03	.06	-.31*
FMP	-.38†	.11	-.01	-.21	-.15	.23	-.20
LSP	.01	.02	.29	.22	.17	.00	-.19
PFE	.03	.33†	.35†	.20	.26	-.08	-.25
RFM	.16	.09	.26	.21	.22	.09	.19
All POS Codes	-.09	.19	.31	.15	.17	.10	-.20

Note: All correlations are Pearson correlations, only correlations with Agreement Complexity (Seventh Column) are Kendall's Tau correlations.
Negative Listener Codes: NBC = no backchannel behavior, NFM = no facial movement, NLS = no looking at speaker, NFE = no facial expression, STW = stonewalling.
Positive Listener Codes: BCP = backchannel behavior present, FMP = facial movement present, LSP = looking at speaker, PFE = positive facial expression, RFP = responsive facial movement.
N = 26
†$p < .1$, *$p < .05$, **$p < .01$ (two-tailed)

The exploratory analysis revealed that the more frequent a pair switched between speakers the more complex their agreement became, $\tau = .26$, p (two-tailed) < 0.1. The less facial expressions (NFE) of affect a group displayed per turn, the worse it felt about the instrumental outcome $r = -.42$, p (two-tailed) < .05. The occurrence of backchannel behavior (BCP), such as nods, or any movement indicating an engagement in what was being said is negatively correlated with the complexity of the final agreements $\tau = -.31$, p (two-tailed) < 0.05. The presence of facial movement during a speaker turn, no matter if positive or negative (FMP) is negatively correlated to perceptions about the instrumental outcome $r = -.38$, p (two-tailed) < .1. Finally, the groups who made more positive facial expression when listening (PFE) also reported feeling better about themselves during the negotiation and about the process $r = -.33$ and r = .35, p (two-tailed) < .1.

DISCUSSION

I found that a Group's Hedonic Balance, measured during an early five-minute sample of a negotiation, correlates with a group's perceptions about the negotiation process. This finding is particularly exciting in the light of Curhan's and Pentland's (Curhan & Pentland, 2007) work, in which they were able to find correlations between interpersonal dynamics and outcomes on the individual level but not on the group-level. How people feel about a negotiation is critical in teamwork as it has direct implications upon a group's continued collaboration. Curhan and colleagues (Curhan, Elfenbein, & Eisenkraft, 2010), for example, found that subjective outcomes are a better predictor of continued negotiation than objective outcomes.

Contradictory to my leading hypothesis, I also found that Group Hedonic Balance was negatively correlated with the complexity of an agreement. This is surprising, as many researchers have related positive emotions with higher creativity (Isen, Daubman, & Nowicki, 1987), and a broadened repertoire of ideas (Fredrickson, 2001). However, it seems plausible that a higher Group Hedonic Balance led to an improved problem solving performance (Estrada, Isen, & Young, 1997) leading teams to agree on simple solutions. Additionally, it seems plausible that interactions with high amounts of positive emotions led to groupthink-like effects (Janis, 1982), resulting in groups making simple decisions rather than weighing in all the concerns or options.

A key limitation of this study is the lack of reliable speaker coding data. This is particularly disappointing as there were many differences between negotiations pairs on a qualitative level as to how much constrained anger or contempt they showed during their interactions. Display of anger, in particular, has been shown to correlate with negotiation outcomes (van Kleef, De Dreu, & Manstead, 2004). The differences in listener behavior alone were not large enough to conclusively confirm or disconfirm the hypotheses about most of the performance variables used. This might be due to a lack of sensitivity of the employed listener codes but it could also have been due to the specific populations the subjects have been drawn from. All

subjects were part of a class that follows a negotiation approach which puts relationship building at the center of its method (Fisher & Ury, 1991). The particular set of values taught in the class might have led the participants to put particular effort into building positive interactions. This focus on positive interactions was reflected in the data as there were no groups with a negative hedonic balance.

A team's willingness for continued collaboration is an important characteristic of high performance teams (Hackman, 1987). The subjective value of a negotiation is an important determinant of people's willingness for continued collaboration and loyalty to an organization (Curhan, Elfenbein, & Kilduff, 2009; Curhan, Elfenbein, & Eisenkraft, 2010).

STUDY 3 – ENGINEERING DESIGN

INTRODUCTION: MECHATRONICS DESIGN

The goal of the third study was to further explore the relationship between affective interaction dynamics and team performance in engineering design teams. More specifically I wanted to to examine the relationship between Group Hedonic Balance and engineering design team performance in the context of a complex longitudinal mechatronics design project. In line with the previous studies, I hypothesized that:

> *H: The Group Hedonic Balance assessed from a thin slice of a group's interaction will be positively correlated with subjective and objective team performance measures assessed at the conclusion of a project.*

Both previous studies examined the relationship between affective dynamics and team performance in dyadic interactions. The first study gave support for Group Hedonic Balance as an indicator of team performance. This was encouraging especially since the intensity of emotional engagement was extremely low in the pair programming interactions. With the choice of a negotiation as the context for the second study I wanted to create a more intense level of emotional engagement. I hoped that the dispute resolution task would surface some of the dynamics I had seen and experienced during conflict discussions in engineering teams. For the second study I had also carefully redesigned the recording setup in order to give me better recordings of the interactions I hoped to witness. However, as in the first study, the laboratory based negotiation task failed to deliver a real-world level of emotional engagement. The negotiation interactions seemed more animated than the pair programming interactions, but they were still far from the engagement levels I had witnessed in real work teams. My third study therefore had to move beyond teams that only existed for the duration of a laboratory task. I wanted to observe and record engineers in teams for whom it really mattered what they were working on, who had

a high level of investment into the given task. I also wanted to go beyond dyads and study affective interaction dynamics as they occur in larger groups of three or even four engineers. Therefore this study builds on the previous ones in three important aspects:

(1) Shift from dyads to teams of three and four engineering students.
(2) Shift from studying teams that exist only in the laboratory to real engineering teams in the field.
(3) Shift from short tasks taking minutes or hours to long-term projects iver several months.

RESEARCH DESIGN AND METHOD

Participants

Participants were drawn from three consecutive cohorts of a graduate level course in team-based engineering design at Stanford University (Carleton & Leifer, 2009). The course spans three quarters or approximately nine months. Early during the course and after obtaining informed consent, I recruited participants with their respective teams to participate in the interaction-focused laboratory portions of the study. At the end of the course I contacted each participant again individually to report on a range of performance relevant self-report measures.

The Individuals: 100 Students distributed into 30 teams participated in the study. This comprised the entire population of students from three consecutive cohorts. All students, with the exception of one student who was pursuing a graduate level degree in product design, were graduate students in mechanical engineering. Most students had several years of prior industry experience. Students usually spend 20 to 40 hours per week on this course and up to 50 or more before major deadlines.

Number of teams	**total**	30
	all male	11
	all female	0
	mixed	19
Team size	**Mean # of Students per team**	3.37
	Min # per team	2
	Max # per team	4

TABLE 5.1
Participant statistics for the engineering design study

The Teams: Participants were grouped into a total of 30 teams. Teams were formed mid-way through the first quarter of the course and did not change for the remainder of the two and a half quarters. Students were asked to form teams according to their own preferences. Specific encouragement was given to form heterogenous teams. A minority of students who did not self-assemble into teams were assigned into teams by the teaching team of the course. In some teams, students had known each other before taking the course. The teams are entirely self-managed and formally leaderless. Teams do not designate an explicit leader or project manager and leadership is emergent throughout the course.

The Setting: The course the students were recruited from, is laid out as a continuous sequence of three quarters or approximately nine months. Only students are admitted into the course who make a commitment to stay within the course for the duration of the entire sequence. The course starts with a series of warmup and team-building exercises that are designed for students to get to know each other and to get familiar with the specific approach to engineering used in that course. Mid way through the first quarter, students are formed into teams and bid for a project to work on. Each team is then paired with an industry sponsor, who provides a compelling and open-ended task, a company liaison who represents the interests of the company sponsoring the project, and a budget of approximately US $20,000 that can be used for building prototypes and contracting out work. For the remainder of the two and a half quarters, students stay within their designated teams and complete a sequence of assignments that build up to a final functional prototype that is presented at a project fair at the end of the three quarter sequence. Additionally each team writes comprehensive report in which they describe their design process, their decision making path and the final design. The course provides students with a wide range of personal and physical resources. From the personal side the course is supported by two professors, a group of up to three teaching assistants and a set of industry-based coaches recruited from a pool of alumni of that course. Additionally each team has an as-

signed partner team at a remote global site to work with. Partner teams are also embedded in team based design courses and are distributed at other universities around the globe. In addition to these resources each team owns a designated area in a "design loft" used as a communal space for the course. The loft is not only a workspace but also a social hub for all participants. There is for example a weekly drinks and dinner event for students to get together with coaches and professors in a relaxed atmosphere. In addition to providing each team with their own space the loft is equipped with resources ranging from a small machine shop area, to a library with past project documentations, and extensive video conferencing equipment. More extensive descriptions of this course, it's structure and history can be found in overviews by Carleton & Leifer (2009) and Stefik & Stefik (2004).

FIGURE 5.1
The ME310 Loft

The Tasks: Each team worked on a one- to two-page project brief that was given by the industry sponsor. The tasks were designed in collaboration with the teaching team to suit the needs of the company and the context of the class. All tasks are open-ended without a set goal. It is part of the learning goals in this class for students to learn how to define a product specification from a highly ambiguous starting point.

Task #	Industry	Abbreviated Problem Description
1	Automotive	Design and build a Human Input Device (HID) that will accommodate driver and passenger needs in the year 2020.
2	Software	Design a website that introduces the various alternative fuel technologies to the consumer.
3	Mechatronics	Design and build a camera-projector prototype sensor system usable in mobile robotics.
4	Government	Develop a solution that protects a person falling down by preventing him/herself from getting hurt.
5	Software	Design and develop a solution that enables a transfer from mechanical design/manufacturing techniques to the design and construction of buildings.
6	Consumer Products	Design a new consumer electronics oral care solution that motivates its user to maintain regular oral hygiene and that provides feedback of how effective hygiene has been.
7	Consumer Products	Develop a platform for blending traditional, physical symbols ("atoms") with what are now relatively separate social practices in the digitally connected world ("bits") as an exciting domain for future wearable, network-capable consumer products.
8	Software	Design and build a system that allows home office work.
9	Telecommunication	The development of new services and products in the health area using mobile telephony as a platform of communications to promote the use and development of 3G technology.

TABLE 5.2
Abbreviated task descriptions of typical engineering design projects in the class studied

Materials and Measures

Data were collected regarding the three main constructs of Affective Quality of Interactions, Interpersonal Trait Perceptions, and Performance.

Affective Interaction Dynamics: The measure of the affective quality of an interaction was operationalized with two procedures – one focusing on the behavioral aspects and the other one focusing on the experiential aspects of an interaction. With

experience and behavior I am focusing on two out of the three common ways in distinguishing different types of affect based on their physiological, behavioral, or experiential properties (Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005). There is evidence for a coherence between those three properties (Mauss, et al., 2005) but also evidence that a closer look at each property reveals insights that cannot be gained from the other properties alone (Levenson & Gottman, 1983). Since the focus of my research was on the behavioral attributes of emotion and since I did not have access to instruments for the recording of physiological data, I only captured data regarding experience and behavior.

The operationalization of the interaction measure by focusing on behavior was done by coding the video of the Interaction Session (see below) using a modified version of the Specific Affect Coding System (SPAFF) (Giese-Davis, 1995; Giese-Davis, Piemme, Dillon, & Twirbutt, 2005). With this study I departed from using the RCISS coding scheme that I had employed in the two previous studies. RCISS fails to be applicable in groups larger than two because of its dependence on speaker turns as the time-unit for coding. Speaker turns are difficult to determine for teams of three and four members. Additionally many of the listener codes of the RCISS are only applicable in a dyadic context. An important point was that RCISS does not capture duration and specific emotion occurrences. SPAFF was also chosen for two specific qualities that distinguish it from other behaviorally based categorization systems for emotions. First, SPAFF (Coan & Gottman, 2007; Gottman, 1996) captures four quadrants of behavior: Facial muscle movement, speech prosody or tone of voice, verbal content, and body posture and movement. This distinguishes it from other coding systems for emotions such as the commonly used Facial Action Coding System (FACS) (Ekman & Friesen, 1978). Second, in comparison to other behavioral emotion coding systems that distinguish affective behaviors on the level of movement, SPAFF distinguishes between affective behaviors on the level of the meaning a particular behavior has in a specific cultural context. The second characteristic makes SPAFF extremely powerful but also highly difficult to apply and to obtain high inter-

rater reliability. SPAFF was used in this study to code the videos obtained from the interaction session. Videos were coded by students who were trained in both FACS and SPAFF. Coders were blind to the hypotheses and to the performance of the respective teams. Coding was done for each person shown in the video separately. Coding took about six to eight hours per person (~20 to 30 hours per team) and 9 videos (25%) were double coded to check for reliability of the coding.

The operationalization of the interaction measure by focusing on subjective experience was done using the affect rating dial procedure (Levenson & Gottman, 1983; Gottman & Levenson, 1985; Ruef & Levenson, 2007). With this method I wanted to capture each participant's continuous moment-to-moment changes in affective experience as they went through the problem discussion session. An additional aim was to add a self-report measure of affect to the behavioral one and to have a measure of affect that would allow a "cheaper" collection of data than it was possible by coding behavior. This idea of an alternative affect measure was supported by prior findings that marital outcomes can be predicted from the ratio of positive to negative affect independent of whether the data was obtained from physiological, self-report, or behavioral characteristics (Gottman, 1994). To accomplish this, directly after going through the problem discussion, each subject was shown a video record of the previous interaction on a separate monitor, and they were asked to provide affect ratings. The rating were obtained by asking the participants to manipulate a dial such that it always indicated what they felt during the interaction session. Previous research suggests that if people recall an episode from memory, they "re-live" the emotions experienced during the actual episode (Ruef & Levenson, 2007). The scale on the dial traversed from "very negative" over "neutral" to "very positive" on a 15 point scale. The dial was connected to a video coding software (Hagedorn, Hailpern, & Karahalios, 2008) and the positions of the dial were logged automatically. During initial tests of this method I noticed that participants got easily distracted by each others' reactions to the video recording. Thus I decided to set up individual headphones for each rater and installed blinds between the

screens to increase focus on the video and avoid distractions. Additionally the blinds helped in allowing subjects to rate negative experiences as such without the fear of their team-mates noticing.

Subjective Team Performance: Subjective Performance was assessed by measuring three variables, that have been identified as important dimensions of team performance (Hackman, 1987): (1) The productive output of the work group as it meets, exceeds, or disappoints the performance standards of the people who receive and/or review the output. (2) The state of the group as a performing unit. This dimension focuses on the social processes used in carrying out the work as they maintain, enhance or worsen the capability of members to work together on subsequent team tasks. (3) The impact of the group experience on individual members as it balances, satisfies, or frustrates the personal needs of group members. These three measures of performance were operationalized using portions of the Team Diagnostic Survey (Wageman, Hackman, & Lehman, 2005). The actual output measure was operationalized additionally by assessing grades, and evaluations that students received for their deliverables throughout and at the end of the course.

Objective Team Performance: As objective measures of team performance I took the grade each team got for the two most important deliverables of the course: (1) The final hardware implementation, and (2) the final 200 to 300 page research report. Both deliverables are not just for the teaching team but are also given to the industry project sponsor. Each grade represented the average of scores from an evaluation by the teaching team of the course.

Additional Measures: Interpersonal Perceptions: Interpersonal trait perceptions were assessed by measuring perceptions of closeness in between members of a group. This closeness measure was operationalized through the Inclusion of the Other in the Self

Scale (IOS) (Aron, Aron, & Smollan, 1992), a single-item, pictorial self report measure of closeness.

Procedures

The procedures used in this study include (1) an Interaction Session with a series of tasks, (2) a Recall Session, and (3) a Follow Up. These procedures were adapted from those developed by Levenson and Gottman (1983) to study the role of affect in the interactions of married couples. The goal of the interaction session was to produce a sample of a team's interaction that was representative of their behavior when working together outside the lab. At the same time I wanted to get a sample of behavior that was comparable between teams and for that a laboratory session seemed most suitable. The design of the Interaction Session used here was based on Gottman's Interaction Session (Gottman & Levenson, 1992). In the original design, couples were asked to first engage in a discussion about the events of the day, second, to discuss a problem area of continuing disagreement in their marriage, and, third, to discuss a mutually agreed on pleasant topic. For each topic 15 minutes were given and each session emphasized a different type of affect: neutral in the first session, negative in the second, and positive in the third. Due to the particular context of my study and specific time constraints in keeping the session at a length of one hour, I made several modifications to the original design. The changes will be described as I lay out the procedures. Figure 5.2 provides an overview of the study procedures and their deployment over time. The procedures described here were run during winter quarter at time t=1 and they were repeated during spring quarter at time t=2. As a final performance assessment I asked every participant to fill out a survey after the end of the class (time t=3) and after all course deliverables had been turned in.

Interaction Session: Each team in the class was given the opportunity to sign up for a one-hour "Team Reflection Exercise" during one of the first weeks of the second quarter of the class. To ensure a high-stakes interaction, the week chosen was a particularly busy one as it directly led up to one of the major deliverables-deadlines of

the course. Upon arrival at the lab, students were led into the laboratory, and each student was asked to sit in one of the four chairs that were evenly spaced around a circular table in the center of the room. After briefing the students with a rough outline for the next hour they were spending in our lab, I asked for informed consent to be videotaped and for consent for the data of the exercise to be used for my research. In the hour following the introduction, each team went through the same sequence of four 15-minute tasks. Audio and video were recorded for both the Requirements Task and the Problem Discussion Task. Four video cameras were installed in the room such that they clearly captured each person's face and upper body. The cameras could be adjusted remotely to adapt to subjects changing positions. A microphone was placed in the center of the table to capture speech at high quality.

In the first task the team was asked to discuss project goals and come to an agreement on the three most important requirements their design would have to fulfill. With this task I wanted to give students an opportunity to get familiar with the lab and bring their thoughts into the context of the course project they were working on. There were several reasons for replacing Gottman's "Events of the Day" discussion with the requirements discussion used here. First, due to the high pressure in working on the class deliverables, I could not ask students to not engage for 8 hours before coming to the lab. Second, since the interaction session was embedded in a class, every task had to contribute directly to a team's project. Third, I wanted a task that would allow the students to recall the context of the projects they were working on. This was necessary because when assembling in the lab, students partially came from other classes. The requirements task allowed for students to quickly get into a meaningful discussion about their project, to surface different values and disagreements and to directly contribute to the progress of a team's specific project.

The second task was modeled after procedures to set up a dyadic problem discussion task in couples (Levenson & Gottman, 1983; Roberts, Tsai, & Coan, 2007). First, each team member was given five minutes to individually fill out a questionnaire that asked for a list of issues of disagreement within the team. Any

issue could be listed whether it was something small such as the font used for a presentation or something major such as the general direction of the project, or specific interpersonal clashes or disagreements. This task was adapted from the Couples Problem Inventory task (Gottman, Markman, & Notarius, 1977). Due to the broad range of problem topics and the non-existence of previous problem inventories we chose a free-form structure for this task without predefined problem categories. After this problem issue assessment I entered the laboratory, joined the team at the table, and facilitated an interaction that had the goal for each participant to share an issue they deemed important, and then to come to an agreement as a team about which topic to discuss during the following 15-minute task. Once an agreement was formed, the task was given to the team to discuss the selected issue with the goal to come to an agreement and I left the room. The ensuing 15-minute "conflict" interaction was the main focus of my laboratory procedure and teams exhibited a wide range of styles as they discussed their problem topic. The decision to focus on the problem discussion was based on previous studies, which found that data from the problem discussion worked best in predicting performance relevant outcomes in marriages (Gottman & Levenson, 1992; Levenson & Gottman, 1985). I omitted the discussion of a “pleasant topic” because I had only limited time to conduct the session and because I found from earlier experimentations with the procedure that teams were not overly distressed from the problem discussion.

Recall Session: For the final task I asked the team to follow into an adjacent room. There, each student was asked to watch a recording of their problem discussion interaction and to indicate how they were feeling during the interaction through a rating dial that was placed in front of them. This task was modeled after the “Recall Session” used in studies on couples (Levenson & Gottman, 1983; Ruef & Levenson, 2007). Before being released and debriefed the participants were asked to fill out selected portions of the Team Diagnostic Survey (TDS), the Inclusion of the Other in

the Self Scale (IOS), and a specific release form that asked for permission to use the video beyond mere analysis.

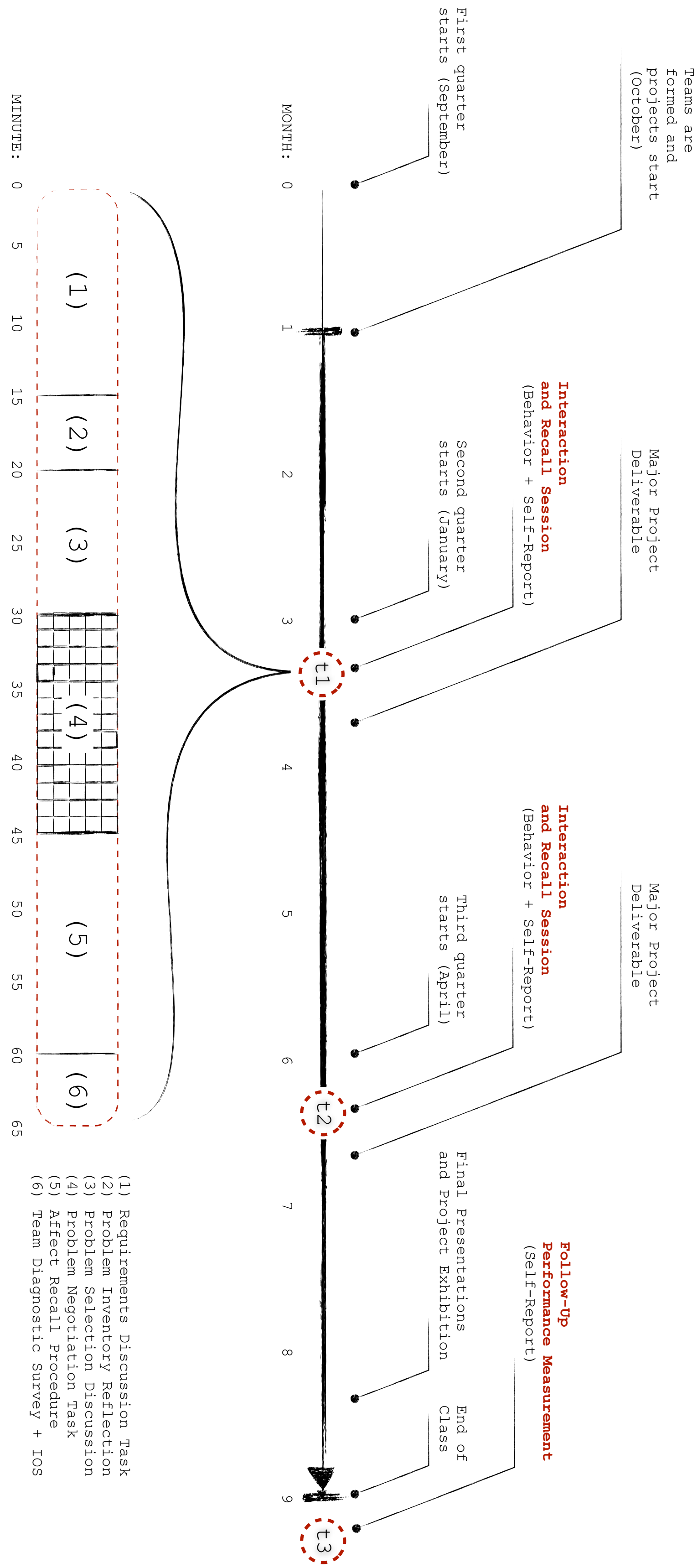


FIGURE 5.2
Engineering Design study: Procedure overview.

RESULTS

All groups were included in the analyses. The procedures were successful in establishing a high level of emotional engagement during the main interaction session that fax exceeded what I was able to observe in the previous two studies. Despite the four cameras and the microphone on the table, some groups engaged in interactions that were surprisingly vulnerable. The teams utilized the interaction session to discuss a wide range of different topics. Most teams reported the interaction session to be valuable and several teams asked whether they could continue their discussion for an extended amount of time.

TABLE 5.3
Correlations and Descriptives for subjective and objective performance metrics.

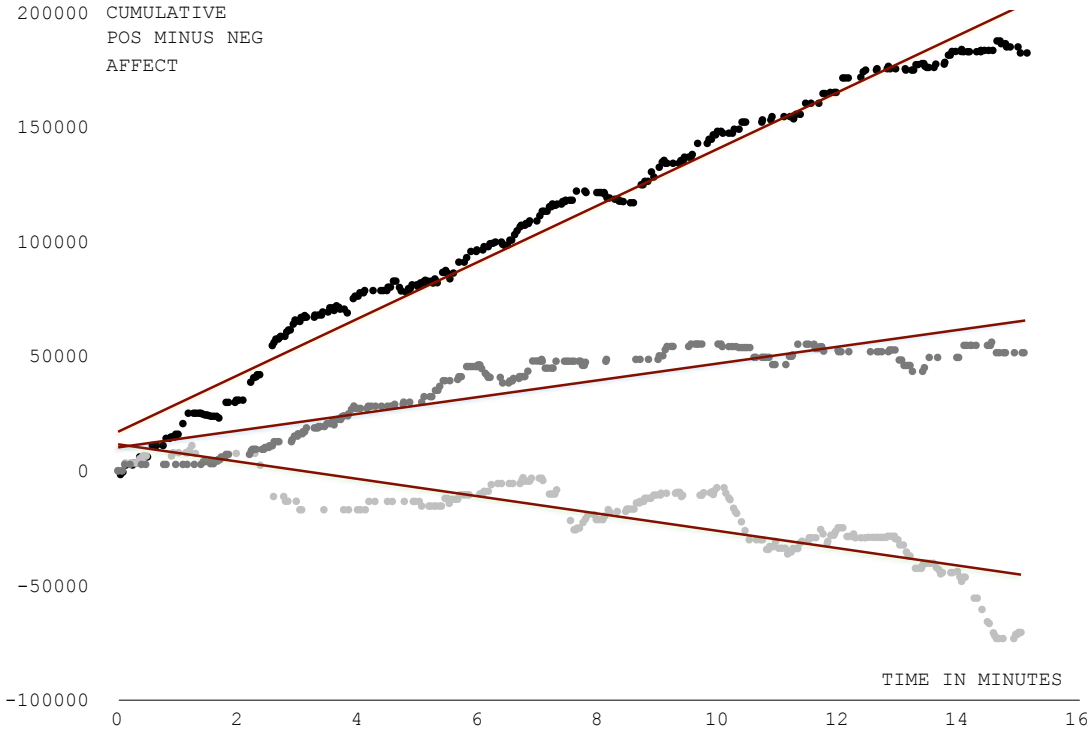
	(1)	(2)	(3)	(4)	(5)	(6)
Subjective Team Performance						
(1) TDS Process [1-5]	–	.86***	.83***	.95***	-.08	.25
(2) TDS Interpersonal [1-5]		–	.79***	.96***	-.10	.30
(3) TDS Learning [1-5]			–	.90***	-.08	.34
(4) TDS Global [1-5]				–	-.09	.31
Objective Team Performance						
(5) Final Prototype Grade					–	.51**
(6) Final Documentation Grade						–
Descriptives						
M	3.55	3.73	4.02	3.77	3.86	4.32
SD	0.56	0.86	0.46	0.59	0.37	0.28
Minimum	2.48	1.56	3.05	2.36	2.59	3.65
Maximum	4.59	4.81	4.58	4.65	4.50	4.70

Note: All correlations are Pearson correlations
N = 30
*p < .05, **p < .01, ***p < .001 (two-tailed)

Coder agreement for the modified SPAFF coding was assessed using Cohen's Kappa (Cohen, 1960) and ranged between κ = 0.51 and κ = 0.67 (*M* = 0.59, *SD* =0.07). According to (Landis & Koch, 1977) this is a moderate to substantial level of agreement. Two measures of Group Hedonic Balance measure were constructed: One based on each individual's stream of affective behavior over time (SPAFF) and one based on each individual's subjective experience over time (Recall). The behavior based Group Hedonic Balance score (SPAFF) was constructed by (1) multiplying

the millisecond-duration of each SPAFF code with "+1" for positive codes, "-1" for negative codes, and "0" for the neutral and the tension codes and plotting the resulting values cumulatively as a point graph over time. Then (2), linear regression analyses were performed on all point graphs to determine their slope. Finally (3) the average slope of the point graphs for each team was taken as their Group Hedonic Balance score. The subjective experience based Group Hedonic Balance Score (Recall) was determined by (1) counting the number of instances the rating dial was turned to one of the upper five positive points (on the 15 point scale) and subtracting the number of instances the rating dial was turned onto one of the lower five negative points. Then (2) the mean of these scores was calculated for each group. The behavior based and the subjective experienced based Group Hedonic Balance scores were significantly correlated, $\tau = .76$, p (one-tailed) < 0.01, allowing the subjective experienced based scores to be used as proxy for the behavior based scores. Subjective and objective measures of group performance were constructed by averaging individual scores on the team level.

FIGURE 5.3
Behavioral point graphs for three members of a design team during the conflict interaction session.



All data were analyzed using Pearson correlations except when testing correlations with the Final Prototype Grade. The Final Prototype Grade, $D(30) = 1.66$, $p < .05$, was significantly non-normal, and therefore Kendall's tau correlation coefficients were used to analyze correlations with this objective performance measure. Descriptives and correlations between performance scores are displayed in the table below. Interestingly, no significant correlations could be found between subjective and objective performance scores.

This study set out to explore whether affective interaction dynamics in a thin lice of a team's interaction is correlated with performance relevant outcomes. In particular I wanted to test whether Group Hedonic Balance is correlated significantly with subjective and objective team performance metrics. I had measured Group Hedonic Balance in January (time t = 1) and March (time t = 2) and correlated these measures with team performance measured in July (time t = 3) after the class was over. Table 5.4 and 5.5 reveal that, in fact, Group Hedonic Balance assessed from only a 15-minute slice of a team's interaction is correlated significantly with several subjective and objective measures of team performance.

TABLE 5.4
Results of the correlation analysis between Group Hedonic Balance measured at time t=1 and Team Performance.

	Subjective Team Performance at T=3				Objective Team Performance at T=3	
	TDS Process	TDS Interpersonal	TDS Learning	TDS Global	Prototype Grade	Documentation Grade
Group Hedonic Balance (SPAFF) at T=1	**.59***	.57	**.59***	**.61***	.39	.11
Group Hedonic Balance (Recall) at T=1	.37	**.59***	.48	.52	**.50***	.22

Note: All correlations are Pearson correlations except the correlations with Prototype Grade were tested with Kendal's tau. N = 9
*p < .05, (one-tailed)

At time t = 1, subjective experience and behavior based measures of Group Hedonic Balance were available for N = 9 teams. The behavior-based Group Hedonic Balance measure at time t = 1 was correlated significantly with the teams' overall self-reported assessment of their effectiveness (TDS Global), $r = .61$, $p < .05$

(one-tailed). Further I found significant correlations with two sub-scales of global team effectiveness on the Team Diagnostic Survey: First, with the teams' self-reported quality of their task processes, $r = .59$, $p < .05$ (one-tailed), and second, with the quality of their learning experience, $r = .59$, $p < .05$. No significant correlations could be found at t = 1 between the behavior based Group Hedonic Balance measure and any of the two objective team performance measures. The subjective experience based Group Hedonic Balance measure taken at t = 1, however, was correlated significantly not only with the teams' quality of interpersonal relationships $r = .59$, $p < .05$ (one-tailed), but also with the team's objective quality of their delivered final prototype $r = .50$, $p < .05$ (one-tailed). These results are particularly exciting given that at time t=1, data were only available for N=9 teams. Overall these results provide support for the hypothesis that Group Hedonic Balance and team performance correspond.

TABLE 5.5
Results of the correlation analysis between Group Hedonic Balance measured at time t=2 and Team Performance.

	Subjective Team Performance				Objective Team Performance	
	TDS Process	TDS Interpersonal	TDS Learning	TDS Global	Prototype Grade	Documentation Grade
Spring Group Hedonic Balance (Recall)	.50**	.61***	.53**	.59***	-.06	.00

Note: All correlations are Pearson correlations except the correlations with Prototype Grade were tested with Kendal's tau.
N = 28
*$p < .05$, (one-tailed)

At time t = 2, only the subjective experience based measure of Group Hedonic Balance was available for N = 30 teams. The subjective experience based Group Hedonic Balance measure at time t = 2 was correlated significantly with the teams global team effectiveness ($r = .59$, $p < 0.001$, one-tailed) and all three sub-scales of team effectiveness: The quality of task processes ($r = .50$, $p < 0.01$, one-tailed), the quality of interpersonal relationships ($r = .61$, $p < 0.001$, one-tailed), and the quality of the team's learning experiences ($r = .53$, $p < 0.01$, one-tailed). None of the correla-

tions between Group Hedonic Balance and any of the objective team performance scores were found significant. Despite the lack of significant correlations with the objective team performance measures, the hypothesis that Group Hedonic Balance and team performance are related receives strong support.

Results from Exploratory Analyses

In addition to testing the correlation between Group Hedonic Balance and team performance, an exploratory analysis was performed to generate insights into specific affect behaviors and their relationship with performance relevant outcomes. For each of the 23 categories of the modified SPAFF coding scheme, I calculated the team-average count of occurrences and duration, and compared these values with the subjective and objective performance measures (see table below). Additionally I explored correlations between average occurrence count and duration of the overall negative, neutral, and positive affect and team performance. Due to the exploratory nature of this analysis, correlations at the 10 percent level are flagged as significant as well.

The exploratory analysis into the relationship between specific emotions and team performance revealed significant correlations on a more fine-grained level. The analysis revealed strongly significant correlations between the prototype grade and particularly the number of expressions of verbal contempt ($r = -.70$, $p < .01$, two-tailed), belligerence ($r = -.63$, $p < .05$, two-tailed), and defensiveness ($r = -.70$, $p < .01$, two-tailed). These results are particularly interesting because contempt and stonewalling have been found to be two of the four most corrosive behavior pattern in marital interactions (Gottman, 1994). Finally the average durations of constrained anger ($r = -.72$, $p < .05$, two-tailed) and domineering ($r = .69$, $p < .05$, two-tailed) were significantly correlated with a team's average perception of the quality of their task processes.

TABLE 5.6
Results of the exploratory analysis for the engineering design study.

		Subjective Team Performance				Objective Team Performance	
		TDS Process	TDS Interpersonal	TDS Learning	TDS Global	Prototype Grade	Documentation Grade
Pos	Low Affection	-.33 (-.51)	-.21 (-.36)	-.17 (-.52)	-.24 (-.46)	-.20 (-.20)	.18 (-.01)
	High Affection	.34 (.34)	.17 (.17)	.17 (.17)	.22 (.22)	-.12 (-.12)	.12 (.12)
	Validation	.35 (.52)	.39 **(.59†)**	**.59†** (.50)	.45 (.57)	.28 (.28)	.17 (.27)
	Interest	-.09 (-.10)	-.27 (-.21)	.10 (.12)	-.14 (-.10)	.06 (.11)	-.44 (-.35)
	Excitement	.36 (.37)	.39 (.55)	.40 (.45)	.40 (.50)	.15 (.22)	.22 (.39)
	Humor	.38 (.56)	.47 **(.59†)**	.54 (.57)	.49 **(.61†)**	.37 (.39)	.23 (.36)
Neutral		.43 (.22)	.35 (.20)	.58 (.20)	.45 (.21)	-.06 (-.11)	-.41 (-.28)
Neg	Tense Humor	.48 (.49)	**.61† (.66†)**	.39 (. 38)	.54 (.57)	.09 (.06)	.29 (.35)
	Tension	.52 (-.49)	.57 (-.55)	**.66† (-.59†)**	.61† (-.57)	.00 (-.06)	.17 (-.01)
	Low Fear	.46 (.44)	.20 (.18)	.35 (.36)	.32 (.31)	-.45 (-.47)	-.10 (-.15)
	High Fear	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)
	Low Sadness	.13 (.16)	.10 (.19)	.37 (.39)	.18 (.25)	-.15 (.09)	.37 (.39)
	High Sadness	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)
	Direct Anger	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)
	Const. Anger	-.16 **(-.72*)**	-.29 (-.54)	.03 (-.55)	-.18 **(-.62†)**	-.09 (.06)	-.40 (.17)
	Micr. Contempt	-.29 (-.35)	-.01 (-.08)	-.18 (-.26)	-.14 (-.21)	.09 (.11)	-.12 (-.31)
	Verbal Contempt	.22 (-.12)	.04 (-.23)	.15 (-.23)	.12 (-.21)	**-.70** (-.44†)**	-.45 (-.40)
	Domineering	**.60† (.69*)**	.44 **(.60†)**	.60† (.57)	.55 **(.65†)**	-.34 (-.28)	.08 (.34)
	Belligerence	-.20 (-.24)	-.08 (-.07)	-.42 (-.42)	-.21 (-.22)	**-.63* (-.65*)**	-.13 (-.17)
	Defensiveness	.20 (.41)	.13 (.33)	.08 (.24)	.14 (.35)	**-.70** (-.50†)**	-.30 (-.15)
	Whining	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)
	Disgust	-.04 (-.10)	-.18 (-.21	.26 (.21)	-.03 (-.08)	.04 (.04)	-.31 (-.28)
	Stonewalling	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)	–.– (–.–)
	Overall Negative	-.44 (-.44)	-.36 (-.44)	**-.60†** (-.49)	-.47 (-.47)	-.31 (-.11)	-.04 (-.16)
	Overall Neutral	.28 (.24)	.16 (.21)	.422 (.24)	.22 (.24)	-.37 (-.10)	**-.64†** (-.27)
	Overall Positive	.25 (.48)	.24 (.54)	.42 **(.58†)**	.30 (.56)	.47 (.48)	.37 (.24)

Note: All correlations are Pearson correlations except the correlations with Prototype Grade were tested with Kendal's tau.
Correlations are for average counts of affect category per team (with average durations in brackets)
N = 9
*$p < .05$, $p < .1$ (one-tailed)

DISCUSSION

This study provides evidence that the affective interactions dynamics occurring during a thin slice of a team's interaction, even when taken early during a longitudinal project, are strong indicators of a team's overall performance. The hypothesis that Group Hedonic Balance is correlated positively with team performance received ample support. Group Hedonic Balance accounted for more than 25% of the variance of one objective and all of the subjective team performance metrics.

At this point it is still an open question whether the findings presented here can be generalized to different types of groups across a variety of contexts. The teams studied here were relatively small with only three or four members. They also had no formal leadership structure and they were working on highly ambiguous engineering tasks. Any of these factors influence how group members regulate their emotional life. A clear goal and task assignments by a designated project leader can keep a group progressing towards a project deliverable despite negative emotions and deep resentments between members of a team. Therefore if strong external structures exist for a team's tasks, communication, and goals, interpersonal affect might have less impact on a team's productive outcome.

One of the most interesting questions emerging from this study is what it means for a group to be regulated or non-regulated. This distinction, which has been useful to distinguish functional from dysfunctional interaction styles of couples as well as programming pairs might not be easily transferrable to larger groups. Simmel, for example pointed out that dyads are fundamentally different from triads because they are the only groups that get immediately destroyed when a member is removed (Coser & Merton, 1971). When a third person is added to a group, according to Simmel (1908), the dynamics drastically change as there is no possible configuration between the three in which one member is not sometimes seen as an intruder regarding the other two members. According to Simmel, three person groups are in constant conflict and out of this dynamic, three configurations emerge: The mediator, the holder of balance of power, and the constant disturber of the solidarity

enjoyed by the other two (Mills, 1953). In spite of these differences between dyads and larger groups, researchers have indeed suggested that it only takes *"one bad apple to spoil the barrel"* (Felps, Mitchell, & Byington, 2006). People have a *"tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person and, consequently to converge emotionally"* (Hatfield, Cacioppo, & Rapson, 1992). This effect, called Emotional Contagion, has been successfully applied to the study of emotions in work teams (Barsade, 2002). A distinction of groups into regulated and non-regulated ones based on one negative group member, however, did not reveal any significant correlations with performance. One single negative group member did not indicate a low performance team. This might be due to the relatively low number of subjects, or the data collection point might have been too early for converging effects to occur. Further studies should explore the behavioral dynamics emerging dependent on the ratio of positive and negative members in a group.

This study presented an initial exploration into the relationship between behavioral affective dynamics and performance in engineering teams. It showed how methods from the study of marital interactions can be applied to gain insights into affective interaction dynamics in small groups. Specifically this study provided evidence that affective balance and a group's ability to regulate their emotions is not only critical for married couples but also for complex organizational groups.

A BALANCE THEORY OF INTRA-GROUP CONFLICT AND PERFORMANCE

INTRODUCTION

In the previous three studies I presented evidence that affective interaction dynamics and specifically Group Hedonic Balance are critical indicators not only of successful marital interactions but also of successful task oriented work groups and teams. Particularly the last study demonstrated how affective interaction dynamics during a 15-minute intra-group conflict discussions distinguished between more successful and less successful teams.

The phenomenon of conflict in teams and its impact on performance has been a longstanding focus of interest for many researchers of organizational groups and teams. The work around the topic of intra-group conflict is probably the largest coherent body of work linking interaction characteristics with performance in small groups. Yet, researchers of groups and teams in organizations have studied conflict almost exclusively from a topical perspective. Consequently the categorizations of conflict into functional or dysfunctional have been made by topic or what conflict is about. In contrast, researchers of marital interactions have studied conflict from an affective perspective. Distinctions about the type of conflict and whether a type is functional or dysfunctional have been made based on emotional differences and not topical ones. Marital conflict types are distinguished by the emotional dynamics occurring in them, or by how conflict is actually carried out. The previous studies suggest that emotional dynamics might distinguish not only functional from dysfunctional marital conflict but also functional from dysfunctional intra-group conflict.

Encouraged by my empirical findings on affective interaction dynamics and Group Hedonic Balance, I therefore want to call into question previous theories about intra-group conflict and performance. In this chapter I will put forth an argu-

ment that the current one-dimensional focus on topic has critical limitations, because it neglects the underlying emotional dynamics. This limitation is particularly important when theorizing the relationship between conflict and performance. I will show in this chapter that by combining theory from marital interaction and organizational behavior, I can propose a theoretical understanding of conflict in teams, that can explain previously observed phenomena in a more coherent way than current theories of intra-group conflict.

CURRENT THEORY: ORIGINS AND LIMITATIONS

Through direct observation and interviews, Jehn (1997) was able to show that people in organizational contexts retroactively and in the moment conceptualize intra-group conflict as either task, relationship, or process oriented. People perceive causes, displays, and consequences of each conflict type as unique and identifiably different from each other. The perception of intra-group conflict in these categories is reflected in the categories researchers have used predominantly to distinguish conflict types when theorizing the relationship between intra group conflict and performance. Consequently, most current theorizations about the relationship between intra-group conflict and performance rely on the particular conceptualization of conflict as either task, relationship, or process oriented. The table below list frequently cited papers on intra-group conflict and illustrates the strong topical focus when distinguishing between task, relationship, and process conflict.

TABLE 6.1
Overview of definitions for task-, process-, and relationship-conflict.

Reference	Task Conflict	Relationship Conflict	Process Conflict
(Guetzkow & Gyr, 1954)	Substantive conflict - conflict rooted in the substance of the task which the group is undertaking. It is associated with intellectual opposition among participants, deriving from the content of the agenda.	Affective conflict - conflict deriving from the emotional, affective aspects of the group's interpersonal relations. It is tension generated by emotional clashes aroused during the interpersonal struggle involve in solving the group's agenda problems.	
(Priem & Price, 1991)	Cognitive conflict is task related, involving the degree of disagreement over the interpretation of a common stimulus.	Social-emotional conflict is interpersonal, involving competition for payoffs or personal disagreements.	
(Jehn, 1995)	Task conflict exists when there are disagreements among group members about the content of the tasks being performed, including differences in viewpoints, ideas, and opinions.	Relationship conflict exists when there are interpersonal incompatibilities among group members, which typically includes tension, animosity, and annoyance among members within a group.	
(Amason, 1996)	Functional cognitive conflict - When conflict is functional, it is generally task oriented and focused on judgmental differences about how best to achieve common objectives.This type of conflict is called cognitive conflict.	Dysfunctional affective conflict - When conflict is dysfunctional, it tends to be emotional and focused on personal incompatibilities or disputes. This type of conflict is called affective conflict.	

Reference	Task Conflict	Relationship Conflict	Process Conflict
(Jehn, 1997)	Substantive conflict is conflict involving the group's task.	Affective conflict refers to conflict in interpersonal relations	Process conflict - conflict about how task accomplishments should proceed in the work unit, who's responsible for what, and how things should be delegated. Process conflict includes disagreements about assignments of duties and resources.
(Eisenhardt, et al., 1998)	Substantive (or cognitive or issue-oriented) conflict: conflict that is centered on alternative courses of action and interpretation of facts, and not conflict that is centered on interpersonal friction and dislike.		
(Pelled, et al., 1999)	Task conflict is a condition in which group members disagree about task issues, including goals, key decision areas, procedures, and the appropriate choice for action.	Emotional conflict is a condition in which group members have interpersonal clashes characterized by anger, frustration, and other negative feelings.	
(Janssen, et al., 1999)	Task conflict in team decision making refers to disagreements about the work to be done including issues such as the allocation of resources, application of procedures, and the development and implementation of policies.	Person conflict in team decision making refers to the occurrence of identity- oriented issues, whereby personal or group beliefs and values come into play. Such personalized disagreement "typically includes tension, animosity, and annoyance among team members" (Jehn, 1995: 258) not directly related to the task being performed.	
(Simons & Peterson, 2000)	Task conflict, or cognitive conflict, is a perception of disagreements among group members about the content of their decisions and involves differences in viewpoints, ideas, and opinions.	Relationship conflict, or emotional conflict, is a perception of interpersonal incompatibility and typically includes tension, annoyance, and animosity among group members.	
(Jehn & Mannix, 2001)	Task conflict is an awareness of differences in viewpoints and opinions pertaining to a group task. Similar to cognitive conflict, it pertains to conflict about ideas and differences of opinion about the task. Task conflicts may coincide with animated discussions and personal excitement but, by definition, are void of the intense interpersonal negative emotions that are more commonly associated with relationship conflict.	Relationship conflict, an awareness of interpersonal incompatibilities, includes affective components such as feeling tension and friction. Relationship conflict involves personal issues such as dislike among group members and feelings such as annoyance, frustration, and irritation. This definition is consistent with past categorizations of conflict that distinguish between affective and cognitive conflict	Process Conflict is defined as an awareness of controversies about aspects of how task accomplishment will proceed. More specifically, process conflict pertains to issues of duty and resource delegation, such as who should do what and how much responsibility different people should get. For example, when group members disagree about whose responsibility it is to complete a specific duty, they are experiencing process conflict.

Reference	Task Conflict	Relationship Conflict	Process Conflict
(De Dreu & Van Vianen, 2001)	Task conflict - conflict about the distribution of resources, about procedures and policies, and about judgments and the interpretation of facts. Task conflict is less threatening to one's personal identity, involves less intense, negative emotions, and tends to motivate team members to search for optimal judgements and decisions	Relationship conflicts concern insights and information that are unrelated to the task, involve negative emotions and threaten one's personal identity and feelings of self-worth.	
(Hinds & Bailey, 2003)	Task conflict refers to disagreements focused on work content.	Affective conflict (sometimes referred to as relationship or emotional conflict) refers to team disagreements that are characterized by anger or hostility among group members.	Process conflict refers to disagreements over the team's approach to the task, its methods, and its group processes.
(Garcia-Prieto, Bellard, & Schneider, 2003)	task-related conflict is assumed to involve cognitive processes	person-related conflict is assumed to involve affective processes.	
(Thatcher, et al., 2003)	Task conflicts are disagreements among group members' ideas and opinions about the task being performed, such as disagreement regarding an organization's current strategic position or determining the correct data to include in a report.	Relationship conflicts are disagreements and incompatibilities among group members about personal issues that are not task related, such as social events, gossip, and world news.	Process conflicts are disagreements about how a task should be accomplished.
(De Dreu & Weingart, 2003)	Examples of task conflict are conflicts about the distribution of resources, procedures and policies, and judgments and interpretation of facts.	Examples of relationship conflict are conflicts about personal taste, political preferences, values, and interpersonal style.	
(Yang & Mossholder, 2004)	Task conflict exists when group members differ in views and opinions regarding the tasks being performed and interpretation of task-related information. Cognitive at its core, task conflict is evidenced when there are disagreements among group members about how particular aspects of tasks are to be accomplished, especially in regard to procedural, policy, and resource distribution issues.	Relationship conflict involves perceived tension and frustration about personal differences such as interpersonal style, attitudes and preferences, and personality. Relationship conflict is centered on group members per se and has an affective element. It is exemplified by friction and clashes over personal values mannerisms.	

The relationship between conflict types and performance

The distinction of conflict types into task and relationship conflict in the research on disagreement in teams is old and dates back to a study in the 50s by Guetzkow and Gyr (1954), who then called it substantive and affective conflict. A similar distinction of group interaction processes into a cognitive task oriented and a emotional relationship oriented category can also be found in other group interaction studies such as Bales' Interaction Process Analysis (Bales, 1950) in which he distinguishes between a task area and a social-emotional area.

In current research, task conflict, also sometimes called cognitive conflict typically refers to conflict originating from disagreements about the content of the task (Jehn, 1995), and thought of as void of intense emotionality (Jehn & Mannix, 2001). Relationship conflict, person conflict, or affective conflict typically refers to conflict in interpersonal relations that exists when there are perceived incompatibilities among group members (Jehn, 1995). It is thought of as high in affective components (Jehn & Mannix, 2001). Process conflict is a younger distinction and is thought of as conflict about how task accomplishments should proceed in the work unit (Jehn, 1997).

The relationship between intra-group conflict and group performance has fascinated researchers for many years. While early conflict theorists suggested that intra-group conflict is always negative to performance (March & Simon, 1958; Pondy, 1967), this notion started to change when Jehn (1995) and Amason (1996) suggested that conflict can actually be good in some situations. The idea that disagreements are necessary and that the absence of conflict can be detrimental to performance has also been proposed earlier by Janis (1982) when introducing the concept of groupthink or by Harvey's Abilene Paradox (Harvey, 1974). Jehn and Amason however suggested that it is not the presence or absence of conflict, but the particular type of conflict that matters. In two highly influential papers on organizational intra-group conflict, Jehn (1995) and Amason (1996) suggested that task-conflict is beneficial whereas relationship conflict is detrimental to a group's performance. Ever

since Jehn and Amason published their highly influential papers on intra-group conflict, almost all following theorizations of intra-group conflict and performance shared to some degree the proposition of helpful task-conflict and harmful relationship conflict (De Dreu & Weingart, 2003).

Challenges of current theories

The relationship between conflict type and performance, however, is not as clear as previously proposed. Jehn's study in '95 (Jehn, 1995) inspired many other studies to explore the relationship between intra-group conflict and performance (De Dreu & Van Vianen, 2001; Eisenhardt, Kahwajy, & Bourgeois, 1998; Janssen, Van De Vliert, & Veenstra, 1999; Jehn, 1997; Pelled, 1996). Even though most of these studies hypothesized task conflict as beneficial and relationship conflict as harmful to positive group performance, their actual findings often do not confirm these hypotheses and instead point to more complex dynamics of intra-group conflict.

TABLE 6.2
Overview of results from previous studies trying to relate conflict type and performance. Each study hypothesized relationship conflict as harmful and task conflict as helpful. The signs show the actual relationship between conflict type and performance as found in the studies.

Reference	Type of Data Collection Method	Subjects	Task Conflict	Relationship Conflict	Process Conflict
(Guetzkow & Gyr, 1954)	behavior coding, interviews, surveys	107 government and business groups	⊕⊖	⊕⊖	
(Jehn, 1995)	semi-structured interviews, surveys	26 management, and 79 groups in industry	⊕⊖	⊖	
(Amason, 1996)	interview, surveys	48 management teams in industry	⊕	⊖	
(Jehn, 1997)	longitudinal field observation	6 management and production teams	⊕	⊖	⊕⊖
(Eisenhardt, et al., 1998)	interviews	12 tech top management teams	⊕		
(Pelled, et al., 1999)	surveys	45 tech industry teams	⊕	○	
(Janssen, Van De Vliert, & Veenstra, 1999)	survey study	106 general team managers	⊕	⊖	
(Jehn & Mannix, 2001)	Surveys	51 Student teams	⊕⊖	⊖	⊕⊖
(De Dreu & Van Vianen, 2001)	Surveys	27 industry teams		⊖	
(Hinds & Bailey, 2003)	Literature review	distributed teams	⊖	⊖	⊖
(Thatcher, et al., 2003)	Surveys	144 MBA student teams	⊖	⊖	⊖
(De Dreu & Weingart, 2003)	Meta analysis	30 publications	⊖	⊖	

⊕⊖ = mixed relationship, ⊕ = positive relationship, ⊖ = negative relationship, ○ = no significant relationship

This discrepancy between hypotheses and actual findings was also pointed out prominently through De Dreu's meta analysis on conflict (De Dreu & Weingart, 2003), in which they were able to show that there is no clear relationship between conflict type and performance. To explain the inconsistencies between the proposed models of conflict and performance, it has been suggested that task conflict can be detrimental as well as beneficial to group performance dependent on the context (De Dreu & Weingart, 2003). Another suggestion explaining the inconsistencies is that beneficial task conflict can gradually transform into detrimental relationship conflict (Yang & Mossholder, 2004), or that they generally occur together (Pelled, Eisenhardt, & Xin, 1999). Other explanations for these discrepancies are often sought in mediation and moderation effects of variables such as task type (De Dreu & Weingart, 2003), trust (Simons & Peterson, 2000), or diversity (Thatcher, Jehn, & Zanutto, 2003). Despite these difficulties, the basic value of the distinction between task and relationship conflict is rarely questioned and the general notion of harmful relationship conflict and helpful task conflict remains largely unchallenged.

Possible origins of current hypotheses

The distinction between task, process, and relationship conflict is a topical one. It is made based on what disagreement is about. Relationship conflict is conceptualized as disagreement *about* personal issues. Task conflict is conceptualized as disagreements *about* the task (Thatcher, et al., 2003). The very use of these distinctions makes certain aspects of conflict more salient than others and shapes the theorizations of the relationship between conflict and performance.

In current definitions of conflict types, task conflict is conceptualized as a cognitive phenomenon and relationship conflict as an affective phenomenon [Table 6.3]. This notion is reflected in the use of conflict terms such as affective conflict for relationship conflict and cognitive conflict for task conflict (Amason, 1996). The origins of current theorizations of relationship conflict as dysfunctional and task conflict as functional can be found in a more than 2000 year old European American cultural suspicion of emotions as being inferior to cognition and inherently harmful.

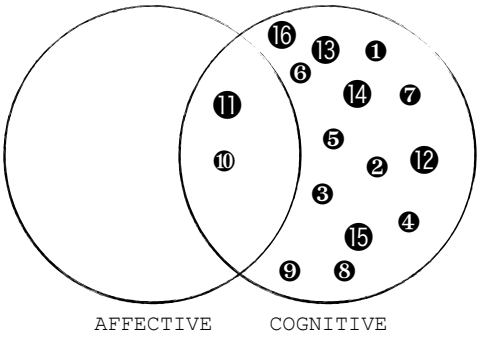
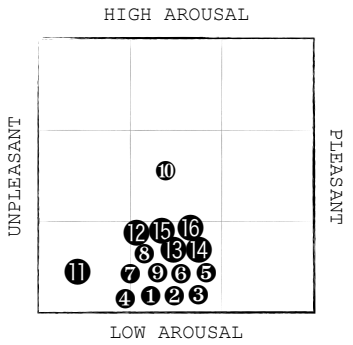
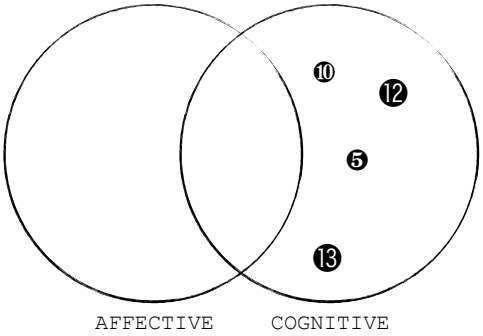
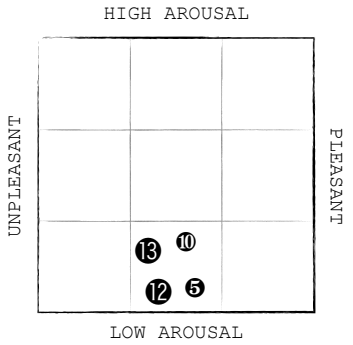
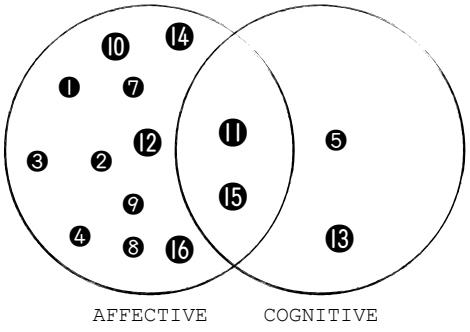
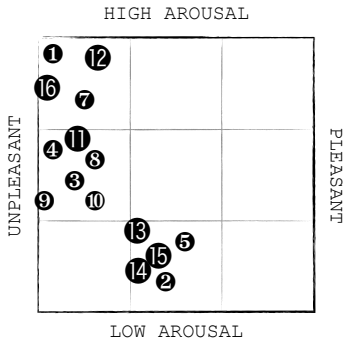
Emotions have long been seen as more harmful and less desirable and thus in constant need of being controlled by reason (Solomon, 1993; Zajonc, 2000). By conceptualizing relationship conflict as an affective phenomenon and task conflict as an emotional phenomenon, it comes by no surprise that the former is theorized as dysfunctional and the latter as functional. Task conflict thus became the Dr. Jekyll and relationship conflict the Mr. Hyde of conflict and we are warned by many conflict researchers to control our emotions in order to keep conflict functional and cognitive.

The same cultural bias might have not only affected how researchers think about conflict but also how subjects respond to surveys on intra group conflict. With only a few exceptions of observational studies (Hobman, Bordia, Irmer, & Chang, 2002; Jehn, 1997), most studies on intra-group conflict are survey-based and rely on retroactive self report measures. Staw (1975) however argued that self report measures are often biased by the respondents' implicit theories about performance. He showed in a series of experiments that people, dependent on their perception of performance, attribute different causes to it. Based on the same cultural bias explained above, it can be assumed that group members when asked in a survey about conflict, attribute negative performance outcome to affective relationship conflict, and positive performance outcomes to cognitive task conflict.

A further examination of conflict type definitions reveals an affective bias between conflict types. With only two exemptions, task conflict and process conflict are conceptualized as neutral from an affective perspective [Table 6.3]. Relationship conflict, on the other hand is associated with different negative emotions of varying intensity. Often terms such as tension, frustration, or anger are used to describe relationship conflict (Pelled, et al., 1999). Interestingly though, positive emotions are not thought to occur during relationship conflict episodes at all. This notion is confirmed by (Barki & Hartwick, 2004) who also observed that negative emotions are typically associated with relationship conflict and even seen as its defining characteristic. Drawing from an extensive body of research that associates negative affect with poor decision making performance, most research on intra-group conflict consequently

hypothesizes affective relationship conflict as detrimental and cognitive, emotionless task conflict as beneficial to performance .

TABLE 6.3
Overview of conflict type conceptualizations. Numbers refer to the references below.

Conflict Type	Affective / Cognitive conceptualization	Affective quality mapping
Task		
Process		
Relationship		

The table places current conflict conceptualizations along their underlying affective and topical dimensions. For each topic category (task, relationship and process) the different definitions are categorized by their conceptualization as either a cognitive or affective phenomenon in the venn diagrams above. Additionally all conflict definitions are placed within an affect grid (Russell, Weiss, & Mendelsohn, 1989) to highlight what kinds of emotions are associated with a particular conceptualization of conflict. If no emotion associations were made in a conflict conceptualization, that particular conceptualization was placed at the center (neutral) of the affect grid. The numbers refer to the following references: 1 (Guetzkow & Gyr, 1954), 2 (Priem & Price, 1991), 3 (Jehn, 1995), 4 (Amason, 1996), 5 (Jehn, 1997), 6 (Eisenhardt, et al., 1998), 7 (Pelled, et al., 1999), 8 (Janssen, et al., 1999), 9 (Simons & Peterson, 2000), 10 (Jehn & Mannix, 2001), 11 (De Dreu & Van Vianen, 2001), 12 (Hinds & Bailey, 2003), 13 (Thatcher, et al., 2003), 14 (Garcia-Prieto, Bellard, & Schneider, 2003), 15 (De Dreu & Weingart, 2003), 16 (Yang & Mossholder, 2004)

Summary

In sum most theories of intra-group conflict and performance rely on a distinction of conflict types along a topical dimension. The topical dimension is a dimension of what conflict is about, e.g. aspects of the task, relationship, or process. This topical dimension is confounded with an affective one. The affective dimension is a dimension of how conflict is carried out, e.g. a positive or neutral emotional tone characterized by expressions interest and excitement, or a negative emotional tone, characterized by expressions of frustration and contempt. Current research systematically associates topical conflict types with specific affective types. Task conflict is generally perceived as cognitive and void of negative emotionality. Relationship conflict is generally perceived as affective and characterized by intense negative emotionality. This affective part, that is underlying current topical conflict distinctions, has influenced current hypotheses about the effects of various conflict types on performance.

I propose that the affective and topical dimension are conceptually and empirically distinct. Affect has been called out as an important moderator by Jehn & Bendersky (2003), but never as an independent descriptor of conflict types itself. There is an opportunity to extend current theory by reconceptualizing conflict as two-dimensional and to distinguish conflict types from an affective in addition to a topical perspective. The groundwork for this affective dimension has been laid out in empirical and theoretical work on conflict and outcomes in marital conflict as well as in the empirical work outlined in the previous chapters.

A NEW THEORIZATION OF INTRA-GROUP CONFLICT AND PERFORMANCE

In the previous sections I laid out current theorizations of intra-group conflict and performance and their reliance on a topical distinction of conflict types. Researchers of organizational behavior have associated outcome predictions such as team member satisfaction and performance with topical features of conflict (De Dreu & Weingart, 2003). In contrast researchers of marital interactions have associated outcome predictions such as marital satisfaction and divorce with affective features of conflict (Gottman, 1994). In addition, the three studies described in this dissertation showed that affective dynamics are not only powerful indicators of successful marriages but also of successful work teams.

In the following paragraphs I propose a theorization of intra-group conflict in organizations that combines the topical with the affective perspective. I will show that a combined affective-topical theory of intra-group sheds some light onto many phenomena that current conflict theories are not able to explain satisfactory.

Affective types of marital conflict

In understanding and predicting how close relationships work or fail, emotions have turned out to be the most productive source of information (Gottman, 2001). Therefore in contrast to researchers of organizational behavior who categorized conflict episodes base on what they are about, researchers of marital conflict have categorized conflict types based how they are carried out emotionally. With the ratio of positive to negative emotions, a distinction is made between regulated and non-regulated couples (Gottman & Levenson, 1992). In a Western European /American context regulated couples are characterized by a 5:1 ratio of positive to negative emotions over time during a conflict episode, whereas non-regulated couples generally exhibit a 0.8:1 ratio of positive to negative emotions over time (Gottman, 1994).

This affect based distinction turned out to have impressive predictive power. Gottman and Levenson (2000), for example, showed that it is possible to accurately (93%) predict the long-term outcome of a marriage based on the affective interaction

quality during a 15-minute video sample of a couple engaging in a problem discussion. It was even possible to predict divorce based on the affective interaction quality during the first 3 minutes of a conflict episode (Carrere & Gottman, 1999). Both the ratios of positivity to negativity in expressed and perceived emotions were predictive of marital outcomes.

Intra-group conflict as an affective process

The impact of conflict on performance is seen as highly contingent on the type of conflict that occurs within a group. In achieving transitions in groups from harmful to helpful conflict, the notion of a "conflict type" is extremely useful and the basis on which that conflict type distinction is drawn becomes critical. Past research has almost always sought this basis in the type of conflict. Jehn recognized the importance of affect as a separate dimension, but did not use it as a basis to draw distinctions between conflict types (Jehn & Bendersky, 2003). Marital models of conflict and performance however suggest that an affective distinction of conflict types are more powerful in establishing useful conflict models. This idea is supported especially by the third study in which I was able to show that the affective dynamics during conflict interactions of small engineering teams are highly indicative of team performance. An affect based distinction of conflict as a basis for an updated conflict-performance model is also supported by the work of Fredrickson, Losada, and Heaphy (Fredrickson & Losada, 2005; Losada & Heaphy, 2004) who proposed that the ratio of positive to negative emotions relates to the outcome of teams, and by (Curhan & Pentland, 2007), who showed that affective interaction patterns in the first five minutes of a negotiation predict its outcome. In figure 6.1 I therefore propose a conflict model that uses affect as the basic for the distinction of conflict types.

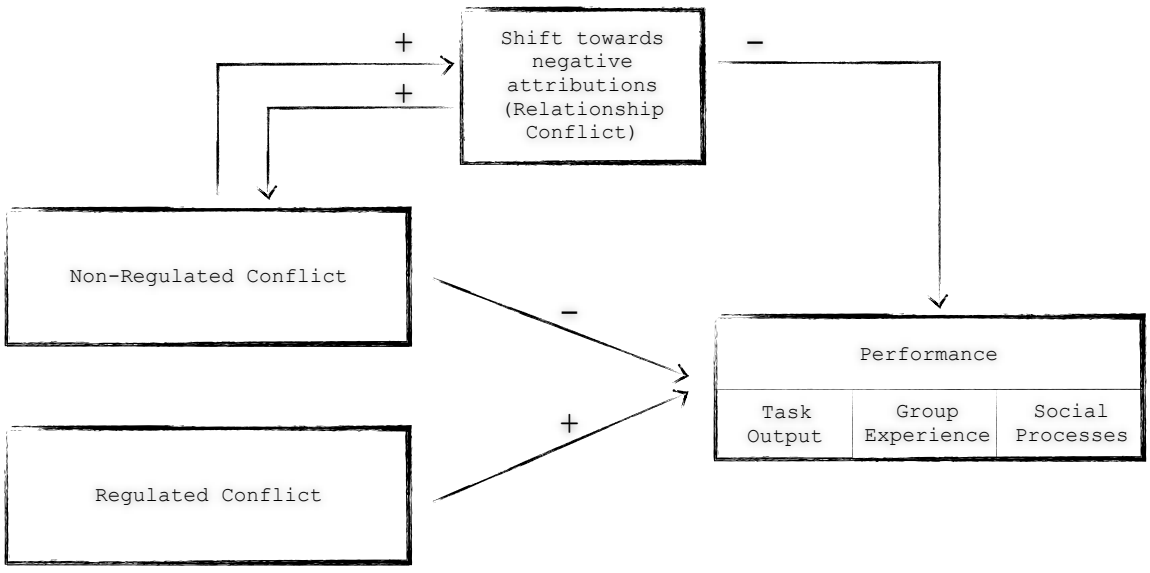
A balance theory of intra-group conflict and performance

With the model depicted in figure 6.1, I want to introduce the notion of intra-group conflict as an affective team process and draw from Gottman's theory of marital dissolution and stability (J. Gottman, 1993) and his Balance theory of marriage Leven-

son & Gottman, 1983) to form an affect based theory of intra-group conflict and performance. I argue that emotions have a central role not only in relationship conflict but in any type of intra-group conflict. The important question is not so much about whether emotions occur or not, but rather what kinds of emotions, in what contexts those emotions occur and how they are responded to.

I use the term intra-group conflict here to refer to interaction episodes that are organized around disagreements, discrepancies, or differing values within a team. By focusing on interactions I refer to conflict only when it is actually carried out in the moment. It is in the context of interactions that conflict surfaces. Intra-group conflict then can be characterized as an interactive process organized around disagreements in which the flow of affective behaviors and perceptions influence each other. The definition of team process as described by Marks et al. (2001) is extended here to include emotional behaviors as acts that convert inputs to outcomes to achieve collective goals.

FIGURE 6.1
Graphical overview of the Balance Theory of Intra-group Conflict and Performance



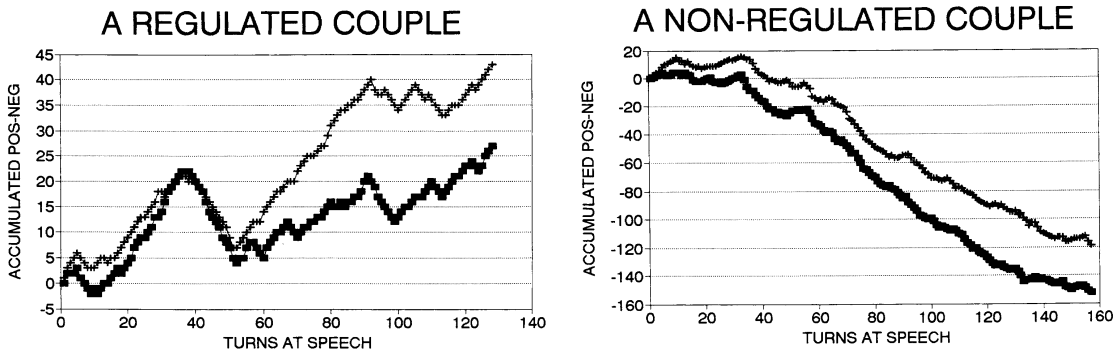
Past theories of conflict and performance have focused on what conflict is about, or the topic such as task, conflict, and relationship when making distinctions about whether conflict is functional or dysfunctional. My goal is to introduce a theory that shifts the focus from topic to affect, or to how conflict is actually carried out

when making distinctions about whether a conflict episode is functional or dysfunctional. This is a shift not only from topic to affect it is also a shift from a retrospective conceptualization of conflict to a focus on the moment-to-moment flow of emotions during conflict. Conflict in teams is constructed in interactions among members of a team. Conflicts do not just come into existence by themselves. They get shaped through interactions. It is those interactions and their quality, that are at the core of my theorization of intra-group conflict.

Distinguishing regulated from non-regulated conflict

The distinction between regulated and non-regulated conflict is based on a Balance Theory of Marriage, which states *"that those processes most important in predicting dissolution would involve a balance, or a regulation, of positive and negative interaction. (Gottman & Levenson, 1992)".* An early method to operationalize the distinction between regulated and non-regulated couples was to codify the occurrence of affective behavior in marital conflict interaction episodes using the Couples Interaction Scoring System (CISS) (Gottman, 1979) and later the Rapid Couples Interaction Scoring System (RCISS) (Krokoff, Gottman, & Hass, 1989).

FIGURE 6.2
Example point graphs of a regulated and a non-regulated couple. Figure from (Gottman, 1994).



"On each conversational turn the total number of positive RCISS speaker codes minus the total number of negative speaker codes was computed for each spouse. Then the cumulative total of these points was plotted for each spouse. [*FIGURE 6.2*] *The slopes of these plots, which were thought to provide a stable estimate of the difference between positive and negative codes over time were determined using linear regression analysis. All couples, even happily married ones, have some amount of negative interaction; similarly, all couples, even unhappily married ones, have some degree of positive interaction (Gottman & Levenson, 1992)."*

This behavior analysis can also be used to distinguish regulated from non-regulated conflict in intra-group conflict. Regulated conflict ( ) is therefore defined as conflict that is characterized by a more or less consistent display of more positive than negative affective behavior. Regulated conflict thus is conflict during which all speaker slopes are significantly positive. Non-regulated conflict ( ) is defined as conflict that is characterized by that at least one of the speaker slopes is not significantly positive. Even though the original distinction between non-regulated and regulated couples was originally behavior-based using RCISS, Gottman was able to show that the categorization of couples can also be operationalized using other affective behavioral coding systems, physiological measures, and experiential measures (Gottman & Levenson, 1992). The applicability of this distinction to work-related contexts was shown in the first study of this dissertation. A point-graph based distinction of programming pairs into regulated and non-regulated pairs was not only possible but also indicative of a pairs subjective and objective performance.

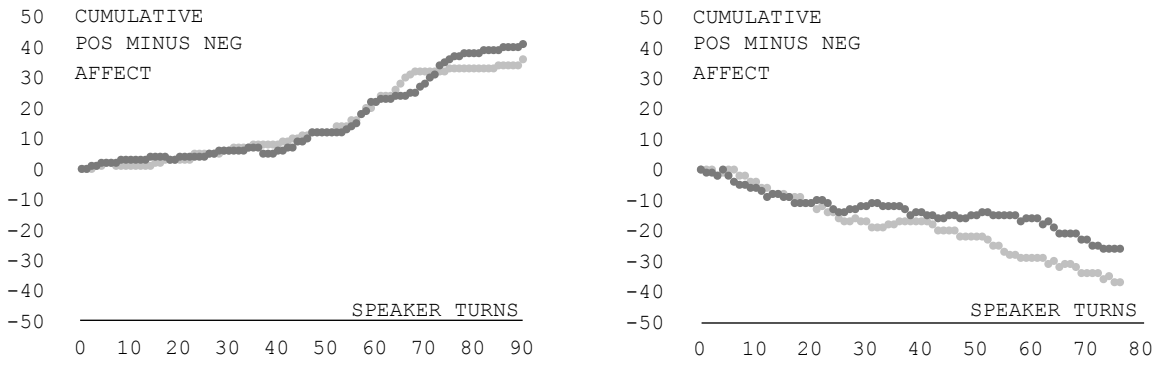


FIGURE 6.3
Point graphs of a regulated and a non-regulated programming pair. The graphs always show the emotion trajectories for each programmer separately. The left graph is drawn from a pair that scored amongst the lowest in the sample and the right graph is drawn from a pair that scored amongst the highest of the pairs studied.

Crucial to the distinction of regulated and non-regulated conflict types is a distinction between harmful and beneficial affective behaviors. Affective behavior is often grouped into positive and negative categories based on valence. It is a categorization based on the subjective experience as pleasant-positive or unpleasant-negative. Many affect categorizations exist, that have valence as a basic dimension (Russell, 1980; Watson & Tellegen, 1985) to categorize emotions. The Specific Af-

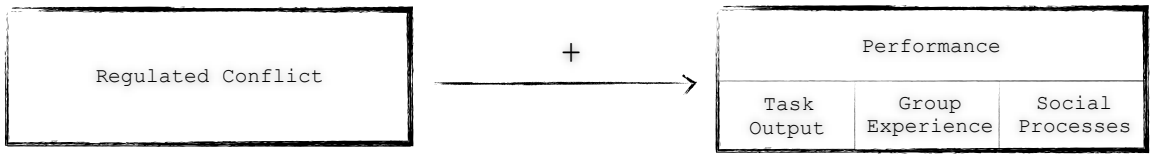
fect Coding System (SPAFF) (Coan & Cottman, 2007), which was used in many studies to classify affective behavior during marital interactions also groups behavior into positive and negative consistent with a valence dimension.

In contrast to the aforementioned valence based approaches, emotions can be categorized as to whether they draw people together or push them apart (Levenson, 1999). Kemper (1984) refers to the first type of emotions as integrative and to the latter type as differentiating emotions. Emotions have relational meaning (Lazarus, 1991), and this is a categorization along that dimension. In the context of this work, I make distinctions between behaviors, based on their affective relational meaning that these behaviors have in a particular cultural context. Distinguishing affective behavior based on relational meaning rather than valence has direct implications for their categorization as either positive or negative. Sadness, for example is usually categorized as negative on many valence based scales (Russell, 1980; Watson & Tellegen, 1985). Gottman, as well, categorizes sadness as a negative behavior (Coan & Cottman, 2007). If, however, we look at the relational meaning of sadness, we can describe it as positive because its expression often draws people together. If we see a person crying, we are more inclined to turn towards that person and help rather than turn away and leave. Giese-Davis refers to sadness as a form of primary negative affect, that is negative affect that has a vulnerable quality which draws people together (Giese-Davis, Piemme, Dillon, & Twirbutt, 2005). She distinguishes primary negative affect from hostile and defensive affect, which is negative affect that pulls people apart. Emotions can also differ in their relational meaning across cultural contexts. Shame, for example, is an emotion that differs greatly in meaning between Japanese and European American cultural contexts. In a European American context shame is seen a withdrawal emotion and negative whereas in a Japanese context shame is seen as an approach emotion (Moore, Romney, Hsia, & Rusch, 1999). This difference reflect the relational meaning shame has in those different cultures. In our cultural context, shame disconnects us from our peers whereas in Japan, shame reconnects people with their community.

Regulated conflict and group performance

Regulated conflict ( ) refers to disagreements about tasks, topics, and relationship (non task) related issues that are carried out with an occurrence of more positive than negative affect. Regulated conflict is characterized by expressions of interest, validation, excitement, and humor and by the absence of hostile emotions such as contempt or belligerence.

FIGURE 6.4
Visualization of relationship between regulated conflict and performance



While positive affect such as positive mood has been directly related to team performance (George, 1995), a growing body of research links the experience of positive affect to improvements in outcomes such as improved creativity (Amabile, Barsade, Mueller, & Staw, 2005), and creative problem solving (Isen, Daubman, & Nowicki, 1987; Isen, Daubman, & Nowicki, 1998). The most extensive empirical support for a link between positive affect and creativity comes from a line of work by Alice Isen, in which she was able to show that inducing positive affect through something as simple as the receipt of a small gift, or a short amusing video can lead to improved creative problem solving (Estrada, Isen, & Young, 1997; Isen, et al., 1987). With the affect and creativity cycle (Amabile, et al., 2005) propose a theory about affect and creativity in organizations. The model describes a cycle in which positive affect promotes creativity which in turn stimulates positive affect. The experience of positive affect has also been shown to lead to more inclusive thinking (Isen & Daubman, 1984) and integrative problem solving (Carnevale & Isen, 1986). Cognitive boundaries seem to become less rigid and more fluid. This integrative notion is consistent with Fredrickson's broaden and build theory of positive emotions, which posits that experiences of positive emotions broaden people's momentary thought-action repertoires and build personal resources (Fredrickson, 2001). Accord-

ing to this theory, individuals who experience specific types of positive affect are more likely to discard everyday behavioral scripts in favor of novel, creative, and often unscripted paths of thought and action (Fredrickson, 1998).

Another major part of research that suggest a positive emotion and outcome performance link is the research on positive affect and decision making. For example (Staw & Barsade, 1993) found that positive affect has beneficial effects on a wide range of decision making tasks, including: more accurate decision making, more careful and thorough considerations of available information, and greater tendency to recognize situational contingencies. Other studies have shown that positive affect supports high decision making performance as it leads to more integrative considerations of available information and a decrease in fixation (Estrada, et al., 1997).

Many studies also support a beneficial relationship between positive affect and social performance outcomes. Staw and Barsade (Staw & Barsade, 1993) for example found that high positive affect leads to increased interpersonal performance and higher contributions to group solutions, in leaderless groups. Through a wide range of experimental settings Alice Isen and her colleagues showed, that the experience of positive affect increases helping behavior (Isen, 1970); Isen, Clark, & Schwartz, 1976; Isen & Levin, 1972, Isen, et al., 1976). This positive relationship between positive affect and helping behavior was found across a wide range of stimuli to induce positive experiences such as receiving cookies, successful completion of a task, finding a coin, or the receipt of stationary as a present. Further (George, 1991) found evidence that positive mood as a state is related to prosocial behavior. Another large body of evidence comes from marital interaction studies in which interaction patterns characterized by positive emotions have been repeatedly related to higher relationship satisfaction (Gottman & Levenson, 1985).

There is also evidence that supports a link between positive affect and positive experience as a performance dimension. Isen and her colleagues (Isen, Shalker, Clark, & Karp, 1978), for example, showed that people experiencing positive affect have a brighter outlook on their circumstances, they seem to see the glass rather half

full than half empty, and they seem to access more positive memories. Staw, Sutton and Pelled (Staw, Sutton, & Pelled, 1994), found that positive affect as expressed or experienced leads to a more supportive social context, greater pay, and higher supervisor evaluations. Baron found a link between experience of positive affect and self efficacy (Baron, 1990). Gottman (1994) showed that an interaction characterized by the expression of positive affect is experienced as a state of well-being. A lack of rejection behaviors, punishing remarks, or statements that cause embarrassment is therefore probably conducive to feeling psychologically safe (Edmondson, 1999) and prone to increased learning performance.

In sum there is a large body of research in support of the notion that the expression and experience of positive affect is related to a broad range of team performance relevant dimensions.

Non-regulated conflict and group performance

The literature supporting a relationship between characteristics of non-regulated conflict interactions and performance is far less extensive as the literature supporting a relationship between regulated affect and performance. Especially in the literature on organizational groups and teams there is almost an absence of research on negative affect and performance. To recall, non-regulated conflict (↘) refers to disagreements about tasks, topics, and relationship (non task) related issues that are carried out with an occurrence of more negative than positive affect. Characteristic emotions of non-regulated conflict are, what Gottman calls the four horsemen of the apocalypse (Gottman, 1994): Defensiveness, contempt, stonewalling, and criticism.

FIGURE 6.5
Visualization of relationship between non-regulated conflict and performance.

There are several mechanisms described in the literature that support a harmful effect of non-regulated conflict on performance. Specifically there are interpersonal effects of negative affect on cognition and problem solving and then there are interpersonal mechanisms on expressed affect and specific emotions.

While there is some support for what is called a depressive-realism effect, that the experience of low levels of negative affect can be beneficial for various decision relevant processes such as accuracy of judgments (Alloy, Abramson, & Viscusi, 1981), or assessments of one's own ability in ambiguous task situations (Tabachnik, Crocker, & Alloy, 1983), there is considerable evidence that these individually advantageous effects might be outweighed by damaging effects of negative affect in interpersonal interactions (Felps, Mitchell, & Byington, 2006). Even though the damaging effects of negative affect in interpersonal settings are often underestimated (Miner, Glomb, & Hulin, 2005), there is evidence that for example, negative affect can poison organizational culture (Aquino, Douglas, & Martinko, 2004), negatively influence perceptions of leaders (Lewis, 2000), negatively influence problem solving performance in family interactions (Forgatch, 1989), and negatively affect health (Begley, 1994). Because of the severity of expressions of negative affect and hostility in interpersonal situations, it has also been referred to as a form of workplace violence (Neuman & Baron, 1998).

In sum, despite of a smaller amount of research into the effects of negative emotions on team-performance relevant mechanisms there is some evidence that negative affect when experienced and especially when expressed, can be harmful to organizational team performance.

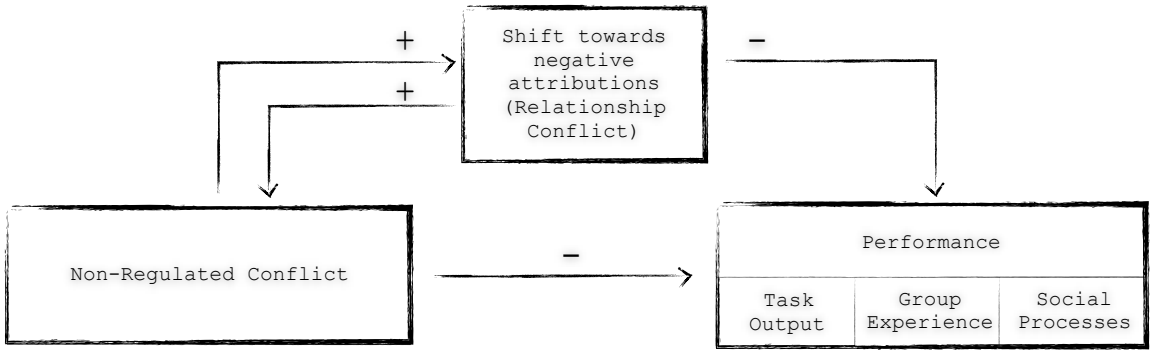
Non-regulated conflict, attributional shifts and group performance - relationship conflict as an emergent state

We only need a thin slice of another person's behavior to make complex assessments about who that person is and what he or she will do in the future (Ambady & Rosenthal, 1992). In a study Ambady and Rosenthal (Ambady & Rosenthal, 1993) showed that from less than a 30 second slice of a teacher's behavior, students formed judg-

ments about the teacher that were in line with the end of semester evaluations by students in the course. In another set of studies Brian Knutson was able to show that people form interpersonal trait inferences from short exposure to facial expressions (Knutson, 1996).

Similarly Gottman, linking interpersonal attributions to expressive behavior, describes a process in which expressive behavior influences perception which in turn affects the formation of global and stable attributions about another person (Gottman, 1994). Felps also theorizes that over time a group member's Negative behavior can be seen as stable and intractable (Felps, et al., 2006). In the current literature on intra-group conflict relationship conflict has been described as a particular type of conflict process. I propose that relationship conflict can be better conceptualized as an emergent state.

FIGURE 6.6
Visualization of relationship between non-regulated conflict, attributional shifts, and performance



According to Marks (Marks, et al., 2001) emergent states are constructs that characterize properties of the team that are typically dynamic in nature and vary as a function of team context, inputs, processes, and outcomes. The notion of relationship conflict as an emergent state draws from Gottman's theory of marital dissolution and stability (Gottman, 1993) when he refers to a sudden, drastic flip in perception. Gottman theorizes that when the ratio of positive to negative behaviors that are displayed in an interaction drop below a certain threshold, the emotional state of the participants change from a state of wellbeing to a state of distress (Gottman, 1994). When this happens it marks a sudden shift in perception from attributing negative

behaviors to situational cases to attributing to dispositional causes. Over time, those negative attributions become global, stable and almost irreversible. Describing relationship conflict as a state change from making positive interpersonal attributions to global negative ones fits quite well with current conceptualizations of relationship conflict. In the relationship conflict examples by Jehn (1997) below, the formation of global negative attributions becomes quite clear.

> *Her attitude just stinks. It's a personality conflict in the first place. I'd rather be working for anyone else but her. I just can't stand her attitude and her voice. We just clash.*
> *Like any situation, there are some of us that don't get along, and so we don't talk at all. (Jehn, 1997)*

In the examples above there are global negative judgments (*Her attitude just stinks*.) and evidence of a shift in perception from situational to dispositional as the very same voice might not always been perceived as negative *(I just can't stand her attitude and her voice)*. Table 6.4 summarizes the characteristics of an attributional shift into the emergent state of relationship conflict. It is important to distinguish the emergent state of relationship conflict from the different types of conflict processes that are focused on relationship related issues as topics of disagreement.

TABLE 6.4
Emergent attributional states.

		State Characteristics
Interpersonal Attributions	**Positive**	- negative interactions are not conceptualizes as relationship conflict - positive interpersonal attributions - perception of well-being during interactions - negative behavior is attributed to situational causes
	Negative	- negative interactions are conceptualized as relationship conflict - global and stable negative attributions that are almost impossible to reverse - perception of distress during interactions - negative behavior is attributed to dispositional causes

SUMMARY AND DISCUSSION

In the preceding paragraphs, I proposed an extension of current organizational theories of intra group conflict by combining them with theories about marital conflict. In sum, the following five ideas are introduced and combined into an updated theory of organizational conflict and performance, a balance theory of intra-group conflict:

(1) A distinction of conflict types into regulated and non-regulated conflict based on affective characteristics as introduced in a Balance Theory of Marriage (Gottman & Levenson, 1992) and a Theory of Marital Dissolution and Stability (Gottman, 1993).

(2) A distinction of conflict types into task, process, and relationship conflict based on topical characteristics according to established conflict theories about organizational intra-group conflict (Jehn, 1997).

(3) A distinction between processes and emergent states (Marks, et al., 2001).

(4) A mechanism linking affective processes with emergent states of global negative attributional shifts towards negative trait inferences as described by Gottman (Gottman, 1993) and Knutson (Knutson, 1996)

(5) A distinction of between regulated and non-regulated conflict based on an approach withdrawal distinction rather than the typical valence based distinction (Kemper, 1984).

The major strengths of the balance theory of intra group conflict come from introducing a shift from focusing on conflict in terms of what it is about (topic) to how it is actually carried out (affect). This shift in focus offers explanations for some phenomena, that previous theories could not explain satisfactory, For example almost all studies on intra-group conflict start out with hypotheses about task conflict being beneficial, and relationship conflict being harmful to performance. However often

results are found that contradict these hypotheses and show task conflict as harmful under some circumstances and relationship conflict as not harmful (Pelled, et al., 1999). This phenomenon is often even referred to as paradoxical (Amason, 1996; Jehn, 1995). Previous theorizations often explained this phenomena by speculations that task conflict can under certain situations turn into relationship conflict. However it is still an open question as to what the mechanisms of this transformation are (Jehn, 1997). According to a balance theory of intra-group conflict, task conflict can occur in two ways: regulated () and non-regulated (). Only the non-regulated form of conflict is harmful to performance relevant outcomes. The same is true for relationship conflict. Each topical conflict type can express itself in two ways, and previously the regulated and non-regulated form have mostly been confounded into the respective topical categories. Distinguishing topic from affect, it is also possible to explain why under certain circumstances task conflict "transforms" into relationship conflict. This transformation can be explained through a shift towards negative interpersonal attributions, led by a continued engagement in non-regulated conflict interactions. The transformations from task to relationship conflict marks a point at which the state of relationship conflict emerges out of continuing non-regulated conflict interactions.

Many researchers have argued that the intensity of emotions during conflict plays an important role. Consequently they advised managers to avoid heated conflict interactions. According to a balance theory of conflict, however, the arousal level is not important. Even researchers of marital interactions initially thought that emotional intensity matters. Gottman (1994) however found that the emotional intensity during conflict is irrelevant. What matters is the balance between positivity and negativity during the interaction. He was able to find three types of functional conflict interactions that vary in their level of intensity but not in their balance between positivity and negativity: Volatile, validating, and conflict-avoiding. The volatile couples are the highest in emotional expressivity, the validating couples intermediate, and the conflict-avoiding couple are the lowest.

The balance theory of intra-group conflict was developed with small highly interdependent teams in mind. It is motivated by my experiences in a small four-person design team and by the findings form the three studies in this dissertation. The theory also focuses on teams that are less regulated through formal processes and more dependent on direct interpersonal interactions. For example consider a small production-focused team with an assigned project manager that follows a clearly laid out project plan. In such a team the affective quality of a teams' interactions might have less impact on the team outcomes because who does what is clearly laid out and whether tasks get done is more dependent on the relationship with the manager than with other team members. Emotions are regulated externally through a strong superimposed procedural structure and hierarchy and this structure takes control precedence. Even if two team members don't like each other they still have clearly assigned tasks and responsibilities. This is different in small, innovation focused, start-up like teams, that often do not have an explicitly assigned project manager nor a precisely laid out project plan that structures their work. In these teams, emotions are more likely to take control precedence over what gets done. Those teams are highly dependent on frequent interactions and information exchange and therefore the affective quality of each interaction directly influences the direction and outcomes of a project.

The recommendations that can be derived from current theories are often difficult to put into practice. A common recommendation to managers and team members is to carefully manage conflict in ways that one can benefit from the helpful effects of task conflict while avoiding the detrimental effects of relationship conflict (Jehn & Mannix, 2001; Amason, 1996; Jehn, 1997). How managers, engineers, and team members can actually do this, remains almost always unclear. Specific recommendations are often limited to avoiding overt expressions of high negative emotions, for example by providing spaces in which employees can "let of steam" (Jehn & Bendersky, 2003). Researchers of marital conflict used to hold the belief that both the avoidance of conflict by suppressing all negative emotions as well as highly vola-

tile and loud engagements in conflict are detrimental for relationships. This changed when Gottman (1994) found that it is not the volatility of conflict that matters but the balance of positivity and negativity that does. He showed that there can be three types of conflict interactions that vary in their level of volatility but that are al functional. A balance theory of intra-group conflict suggests that the same is true for organizational teams. What matters is not the "what" of conflict, whether people disagree about task, relationship, or process related issues, but "how" they actually do it matters a great deal. Eisenhardt and others (Eisenhardt, Kahwajy, & Bourgeois, 1998) when quoting a manager "We scream a lot, then laugh, and then resolve the issues" support this idea that there can be highly volatile disagreements that are not harmful for the overall progress of a project. The balance theory of intra-group conflict introduced here derives its core distinctions from observable behavior. These behaviors can be acted on directly and conflict interactions can be steered from negatively to positively balanced ones.

Engineering team performance is still seen as a determinant of rational thought and technical engineering skills. These skills however can only come to bear in the context of emotionally functional interactions. The ability to build those interactions by regulating ones own emotions and the emotions of fellow team members is therefore critical for building a team interaction context that is conducive for high engineering team performance. It it time that emotions are considered as crucial in understanding engineering team work. Emotions are not ornament, they are the drivers of high engineering team performance.

... that’s it ...

... yup.

BIBLIOGRAPHY

Alloy, L. B., Abramson, L. Y., & Viscusi, D. (1981). Induced mood and the illusion of control. [doi:10.1037/0022-3514.41.6.1129]. *Journal of Personality and Social Psychology, 41*(6), 1129-1140.

Altman, D. G. (1991). Practical statistics for medical research: Chapman & Hall.

Amabile, T. M., Barsade, S. G., Mueller, J. S., & Staw, B. M. (2005). Affect and Creativity at Work. [Article]. *Administrative Science Quarterly, 50*(3), 367-403.

Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of Management Journal, 39*(1), 123-123.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. [doi:10.1037/0033-2909.111.2.256]. *Psychological Bulletin, 111*(2), 256-274.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. [doi:10.1037/0022-3514.64.3.431]. *Journal of Personality and Social Psychology, 64*(3), 431-441.

Aquino, K., Douglas, S., & Martinko, M. J. (2004). Overt Anger in Response to Victimization: Attributional Style and Organizational Norms as Moderators. [doi:10.1037/1076-8998.9.2.152]. *Journal of Occupational Health Psychology, 9*(2), 152-164.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. [doi:10.1037/0022-3514.63.4.596]. *Journal of Personality and Social Psychology, 63*(4), 596-612.

Bakeman, R., & Gottman, J. (1997). Observing interaction: An introduction to sequential analysis. Cambridge, UK: Cambridge University Press.

Bakeman, R., & Quera, V. (1995). Analyzing interaction: Sequential analysis with SDIS & GSEQ. Cambridge, UK: Cambridge University Press.

Bales, R. F. (1950). A set of categories for the analysis of small group interaction. [doi:10.2307/2086790]. *American Sociological Review, 15*, 257-263.

Barki, H., & Hartwick, J. (2004). Conceptualizing The Construct of Interpersonal Conflict. [doi:10.1108/eb022913]. *International Journal of Conflict Management, 15*(3), 216-244.

Baron, R. A. (1990). Environmentally induced positive affect: Its impact on self-efficacy, task performance, negotiation, and conflict. [doi:10.1111/j.1559-1816.1990.tb00417.x]. *Journal of Applied Social Psychology, 20*(5, Pt 2), 368-384.

Baron, R. A., & Neuman, J. H. (1996). Workplace violence and workplace aggression: Evidence on their relative frequency and potential causes.

[doi:10.1002/(SICI)1098-2337(1996)22:3<161::AID-AB1>3.0.CO;2-Q]. *Aggressive Behavior, 22*(3), 161-173.

Barsade, S. G. (2002). The ripple effects: Emotional contagion and its influence on group behavior. [doi:10.2307/3094912]. *Administrative Science Quarterly, 47*(4), 644-675.

Barsade, S. G., & Gibson, D. E. (2007). Why does affect matter in organizations? *The Academy of Management Perspectives, 21*(1), 36-59.

Beck, K., & Andres, C. (2004). Extreme Programming Explained: Embrace Change (2nd Edition): Addison-Wesley Professional.

Begley, T. M. (1994). Expressed and suppressed anger as predictors of health complaints. [doi:10.1002/job.4030150603]. *Journal of Organizational Behavior, 15*(6), 503-516.

Bucciarelli, L. L. Between thought and object in engineering design. [doi:10.1016/S0142-694X(01)00035-7]. *Design Studies, 23*(3), 219-231.

Bucciarelli, L. L. (1988). An ethnographic perspective on engineering design. [doi: 10.1016/0142-694X(88)90045-2]. *Design Studies, 9*(3), 159-168.

Carleton, T., & Leifer, L. (2009, 30-31 March 2009). *Stanford's ME310 Course as an Evolution of Engineering Design.* Paper presented at the 19th CIRP Design Conference – Competitive Design, Cranfield University, UK.

Carnevale, P. J., & Isen, A. M. (1986). The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation. [doi:10.1016/0749-5978(86)90041-5]. *Organizational Behavior and Human Decision Processes, 37*(1), 1-13.

Carrere, S., & Gottman, J. M. (1999). Predicting divorce among newlyweds from the first three minutes of a marital conflict discussion. [doi:10.1111/j.1545-5300.1999.00293.x]. *Family Process, 38*(3), 293-301.

Carrizosa, K., Eris, Ö., Milne, A., & Mabogunje, A. (2002). *Building the Design Observatory: a Core Instrument for Design Research.* Paper presented at the 7th International Design Conference - DESIGN 2002.

Chong, J., & Hurlbutt, T. (2007). *The Social Dynamics of Pair Programming*. Paper presented at the Proceedings of the 29th international conference on Software Engineering.

Chong, J., Plummer, R., Leifer, L., Klemmer, S. R., Eris, O., & Toye, G. *Pair programming: When and why it works*. Paper presented at the 17th Workshop of the Psychology of Programming Interest Group PPIG17.

Coan, J. A., & Gottman, J. M. (2007). The Specific Affect Coding System (SPAFF). In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment.* (pp. 267-285): New York, NY, US: Oxford University Press.

Cockburn, A., & Williams, L. (2000). *The Costs and Benefits of Pair Programming.* Paper presented at the First International Conference on Extreme Programming and Flexible Processes in Software Engineering (XP2000).

Cohen, J. (1960). A coefficient of agreement for nominal scales. [doi:10.1177/001316446002000104]. *Educational and Psychological Measurement, 20*, 37-46.

Coser, L. A., & Merton, R. K. (1977). ´The significance of Simmelís Work *Masters of Sociological Thought: Ideas in Historical and Social Context*. New York, NY: Harcourt Brace Jovanovich.

Curhan, J., Elfenbein, H., & Eisenkraft, N. (2010). The Objective Value of Subjective Value: A Multi-Round Negotiation Study. [doi:10.1111/j.1559-1816.2010.00593.x]. *Journal of Applied Social Psychology, 40*(3), 690-709.

Curhan, J. R., Elfenbein, H. A., & Kilduff, G. J. (2009). Getting off on the right foot: Subjective value versus economic value in predicting longitudinal job outcomes from job offer negotiations. [doi:10.1037/a0013746]. *Journal of Applied Psychology, 94*(2), 524-534.

Curhan, J. R., Elfenbein, H. A., & Xu, H. (2006). What do people value when they negotiate? Mapping the domain of subjective value in negotiation. [doi:10.1037/0022-3514.91.3.493]. *Journal of Personality and Social Psychology, 91*(3), 493-512.

Curhan, J. R., & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. [doi:10.1037/0021-9010.92.3.802]. *Journal of Applied Psychology, 92*(3), 802-811.

De Dreu, C. K. W., & Van Vianen, A. E. M. (2001). Managing relationship conflict and the effectiveness of organizational teams. [doi:10.1002/job.71]. *Journal of Organizational Behavior, 22*(3), 309-328.

De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. [doi:10.1037/0021-9010.88.4.741]. *Journal of Applied Psychology, 88*(4), 741-749.

Desmet, P. (2003). A Multilayered Model of Product Emotions. *The Design Journal, 6*(2), 4-13.

Desmet, P. (2005). Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Products. In M. Blythe, K. Overbeeke, A. Monk & P. Wright (Eds.), *Funology* (Vol. 3, pp. 111-123): Springer Netherlands.

Desmet, P., & Hekkert, P. (2007). Framework of Product Experience. *International Journal of Design, 1*(1), 57-66.

Dick, A. J., & Zarnett, B. Paired programming and personality traits. Paper presented at the XP2002.

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. [doi:10.1037/0003-066X.55.1.34]. *American Psychologist, 55*(1), 34-43.

Diener, E., & Diener, C. (1996). Most people are happy. [doi:10.1111/j.1467-9280.1996.tb00354.x]. *Psychological Science, 7*(3), 181-185.

Diener, E., & Suh, E. (1997). Measuring quality of life: Economic, social, and subjective indicators. [doi:10.1023/A:1006859511756]. *Social Indicators Research, 40*(1-2), 189-216.

Dong, A. (2006). *How am I doing? The language of appraisal in design*. Paper presented at the Design Computing and Cognition'06.

Dong, A., Kleinsmann, M., & Valkenburg, R. (2009). Affect-in-cognition through the language of appraisals. [doi: 10.1016/j.destud.2008.12.003]. *Design Studies, 30*(2), 138-153.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. [doi:10.2307/2666999]. *Administrative Science Quarterly, 44*(2), 350-383.

Eisenhardt, K. M., & Kahwajy, J. L. (1997). Conflict and strategic choice: how top management teams disagree. *California management review, 39*(2), 42-42-62.

Ekman, P., & Friesen, W. V. (1978). Facial action coding system: A technique for the measurement of facial movement: Consulting Psychologists Press, Palo Alto, CA.

Eris, Ö. (2004). Effective Inquiry for Innovative Engineering Design: From Basic Principles to Applications. Norwell, Massachusetts, USA: Kluwer Academic Publishers.

Estrada, C. A., Isen, A. M., & Young, M. J. (1997). Positive affect facilitates integration of information and decreases anchoring in reasoning among physicians. [doi:10.1006/obhd.1997.2734]. *Organizational Behavior and Human Decision Processes, 72*(1), 117-135.

Felps, W., Mitchell, T. R., & Byington, E. (2006). How, When, and Why Bad Apples Spoil the Barrel: Negative Group Members and Dysfunctional Groups. [doi: 10.1016/S0191-3085(06)27005-9]. *Research in Organizational Behavior, 27*, 175-222.

Finger, S., & Dixon, J. R. (1989). A review of research in mechanical engineering design. Part I: Descriptive, prescriptive, and computer-based models of design processes. *Research in Engineering Design, 1*(1), 51-67.

Finger, S., & Dixon, J. R. (1989). A review of research in mechanical engineering design. Part II: Representations, analysis, and design for the life cycle. *Research in Engineering Design, 1*(2), 121-137.

Fisher, R., Ury, W., & Patton, B. (1991). Getting to Yes: Negotiating agreement without giving in. New York, NY: Penguin Books.

Forgatch, M. S. (1989). Patterns and outcomes in family problem solving: The disrupting effect of negative emotion. [doi:10.2307/352373]. *Journal of Marriage & the Family, 51*(1), 115-124.

Frankenberger, E., & Auer, P. (1997). Standardized observation of team-work in design. *Research in Engineering Design, 9*(1), 1-9.

Fredrickson, B. L. (1998). What good are positive emotions? [doi:10.1037/1089-2680.2.3.300]. *Review of General Psychology, 2*(3), 300-319.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. [doi:10.1037/0003-066X.56.3.218]. *American Psychologist, 56*(3), 218-226.

Fredrickson, B. L., & Losada, M. F. (2005). Positive Affect and the Complex Dynamics of Human Flourishing. [doi:10.1037/0003-066X.60.7.678]. *American Psychologist, 60*(7), 678-686.

Garcia-Prieto, P., Bellard, E., & Schneider, S. C. (2003). Experiencing Diversity, Conflict, and Emotions in Teams. [doi:10.1111/1464-0597.00142]. *Applied Psychology An International Review, 52*(3), 413-440.

George, J. M. (1991). State or trait: Effects of positive mood on prosocial behaviors at work. [doi:10.1037/0021-9010.76.2.299]. *Journal of Applied Psychology, 76*(2), 299-307.

George, J. M. (1995). Leader positive mood and group performance: The case of customer service. [doi:10.1111/j.1559-1816.1995.tb01775.x]. *Journal of Applied Social Psychology, 25*(9), 778-794.

Giese-Davis, J. (1995). *SPAFF for Group Therapy Sessions*.Unpublished manuscript, Stanford.

GIESE-DAVIS, J., PIEMME, K. A., DILLON, C., & TWIRBUTT, S. (2005). Macrovariables in affective expression in women with breast cancer participating in support groups. In J. A. Harrigan, R. Rosenthal & K. R. Scherer (Eds.), *The New Handbook of Methods In Nonverbal Behavior Research* (pp. 397-446). Oxford, UK.: Oxford University Press.

Gottman, J., Markman, H., & Notarius, C. (1977). The topography of marital conflict: A sequential analysis of verbal and nonverbal behavior. [doi:10.2307/350902]. *Journal of Marriage & the Family, 39*(3), 461-477.

Gottman, J. M. (1979). Marital interaction: Experimental investigations. New York, NY: Academic Press.

Gottman, J. M. (1993). A theory of marital dissolution and stability. [doi:10.1037/0893-3200.7.1.57]. *Journal of Family Psychology, 7*(1), 57-75.

Gottman, J. M. (1994). Nonlinear differential equation models of marital interaction. In S. Johnson & L. Greenberg (Eds.), *The heart of the matter: Perspectives on emotion in marital therapy* (pp. 256-293). New York, NY: Brunner/Mazel, Inc.

Gottman, J. M. (1994). What predicts divorce? The relationship between marital processes and marital outcomes: Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Gottman, J. M. (1999). The marriage clinic: A scientifically based marital therapy: New York, NY, US: W W Norton & Co.

Gottman, J. M. (2001). What the study of relationships has to say about emotion research. *Social Science Information, 40*(1), 79-94.

Gottman, J. M. ((1996)). *What predicts divorce: The measures*. Hillsdale NJ, USA: : Lawrence Erlbaum Associates.

Gottman, J. M., Coan, J., Carrere, S., & Swanson, C. (1998). Predicting marital happiness and stability from newlywed interactions. [doi:10.2307/353438]. *Journal of Marriage & the Family, 60*(1), 5-22.

Gottman, J. M., Katz, L. F., & Hooven, C. (1996). Parental meta-emotion philosophy and the emotional life of families: Theoretical models and preliminary data. [doi:10.1037/0893-3200.10.3.243]. *Journal of Family Psychology, 10*(3), 243-268.

Gottman, J. M., & Levenson, R. W. (1985). A valid procedure for obtaining self-report of affect in marital interaction. [doi:10.1037/0022-006X.53.2.151]. *Journal of Consulting and Clinical Psychology, 53*(2), 151-160.

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. [doi:10.1037/0022-3514.63.2.221]. *Journal of Personality and Social Psychology, 63*(2), 221-233.

Gottman, J. M., & Levenson, R. W. (2000). The timing of divorce: Predicting when a couple will divorce over a 14-year period. [doi:10.1111/j.1741-3737.2000.00737.x]. *Journal of Marriage & the Family, 62*(3), 737-745.

Gottman, J. M., & Silver, N. (1995). Why marriages succeed or fail: And how you can make yours last. New York, NY: Simon and Schuster Paperbacks.

Guetzkow, H., & Gyr, J. (1954). An analysis of conflict in decision-making groups. [doi:10.1177/001872675400700307]. *Human Relations, 7*, 367-382.

Hackman, J. R. (1987). The design of work teams. In R. T. Golembiewski (Ed.), *Handbook of organizational behavior* (pp. 315-342). Englewood Cliffs, NJ: Prentice Hall.

Hagedorn, J., Hailpern, J., & Karahalios, K. G. (2008). *VCode and VData: illustrating a new framework for supporting the video annotation workflow*. Paper presented at the Proceedings of the working conference on Advanced visual interfaces.

Hanks, B., McDowell, C., Draper, D., & Krnjajic, M. (2004). Program quality with pair programming in CS1. *SIGCSE Bull., 36*(3), 176-180.

Hanks, B. F. (2006). *Empirical studies of distributed pair programming.* Unpublished Dissertation, University of California at Santa Cruz, Santa Cruz, CA.

Hannay, J., Dyba, T., Arisholm, E., & Sjoberg, D. I. K. (2009). The effectiveness of pair programming: A meta-analysis. [doi: 10.1016/j.infsof.2009.02.001]. *Information and Software Technology, 51*(7), 1110-1122.

Hannay, J. E., Arisholm, E., Engvik, H., & Sjoberg, D. I. K. (2010). Effects of Personality on Pair Programming. *IEEE Trans. Softw. Eng., 36*(1), 61-80.

Harvey, J. B. (1974). The Abilene paradox: The management of agreement. [doi:10.1016/0090-2616(74)90005-9]. *Organizational Dynamics, 3*(1), 63-80.

Hekkert, P. (2006). Design aesthetics: Principles of pleasure in design. *Psychology Science, 48*(2), 157.

Hinds, P. J., & Bailey, D. E. (2003). Out of Sight, Out of Sync: Understanding Conflict in Distributed Teams. [doi:10.1287/orsc.14.6.615.24872]. *Organization Science, 14*(6), 615-632.

Hobman, E. V., Bordia, P., Irmer, B., & Chang, A. (2002). The expression of conflict in computer-mediated and face-to-face groups. *33*(4).

Isen, A. M. (1970). Success, failure, attention, and reaction to others: The warm glow of success. [doi:10.1037/h0029610]. *Journal of Personality and Social Psychology, 15*(4), 294-301.

Isen, A. M., & Baron, R. A. (1991). Positive affect as a factor in organizational behavior. *Research in Organizational Behavior, 13*, 1-53.

Isen, A. M., Clark, M., & Schwartz, M. F. (1976). Duration of the effect of good mood on helping: "Footprints on the sands of time.". [doi:10.1037/0022-3514.34.3.385]. *Journal of Personality and Social Psychology, 34*(3), 385-393.

Isen, A. M., & Daubman, K. A. (1984). The influence of affect on categorization. [doi:10.1037/0022-3514.47.6.1206]. *Journal of Personality and Social Psychology, 47*(6), 1206-1217.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1987). Positive affect facilitates creative problem solving. [doi:10.1037/0022-3514.52.6.1122]. *Journal of Personality and Social Psychology, 52*(6), 1122-1131.

Isen, A. M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. [doi:10.1037/h0032317]. *Journal of Personality and Social Psychology, 21*(3), 384-388.

Isen, A. M., Shalker, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? [doi:10.1037/0022-3514.36.1.1]. *Journal of Personality and Social Psychology, 36*(1), 1-12.

Jacobson, N. S., Gottman, J. M., Gortner, E., Berns, S., & Shortt, J. W. (1996). Psychological factors in the longitudinal course of battering: When do the couples split up? When does the abuse decrease? *Violence and Victims, 11*(4), 371-392.

Janis, I. L. (1982). Groupthink: psychological studies of policy decisions and fiascoes: Houghton Mifflin.

Janssen, O., Van de Vliert, E., & Veenstra, C. (1999). How task and person conflict shape the role of positive interdependence in management teams. [doi:10.1016/S0149-2063(99)80006-3]. *Journal of Management, 25*(2), 117-141.

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. [doi:10.2307/2393638]. *Administrative Science Quarterly, 40*(2), 256-282.

Jehn, K. A. (1997). A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative Science Quarterly, 42*(3), 530-530-557.

Jehn, K. A., & Bendersky, C. (2003). Intragroup conflict in organizations: A contingency perspective on the conflict-outcome relationship. In B. M. Staw & R. M. Kramer (Eds.), *Research in Organizational Behavior* (Vol. 25, pp. 187-242). Kidlington, Oxford, UK: Elsevier Ltd.

Jehn, K. A., & Mannix, E. A. (2001). The Dynamic Nature of Conflict: A Longitudinal Study of Intragroup Conflict and Group Performance. [doi:10.2307/3069453]. *Academy of Management Journal, 44*(2), 238-251.

Johnson, E. J., & Tversky, A. (1983). Affect, generalization, and the perception of risk. [doi:10.1037/0022-3514.45.1.20]. *Journal of Personality and Social Psychology, 45*(1), 20-31.

Jung, M., Chong, J., & Leifer, L. (2010). Pair Programming Performance: An emotional dynamics point of view from marital pair counseling. *Electronic Colloquium on Design Thinking Research (ECDTR)*,

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology.* (pp. 3-25): New York, NY, US: Russell Sage Foundation.

Kahneman, D., & Krueger, A. B. (2006). Developments in the Measurement of Subjective Well-Being. [doi:10.1257/089533006776526030]. *The Journal of Economic Perspectives, 20*(1), 3-24.

Kelly, J. R., & Barsade, S. G. (2001). Mood and emotions in small groups and work teams. [doi:10.1006/obhd.2001.2974]. *Organizational Behavior and Human Decision Processes, 86*(1), 99-130.

Kemper, T. D. (1984). Power, status, and emotions: A sociological contribution to a psychophysiological domain *Approaches to emotion* (Vol. 369, pp. 384).

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. [doi:10.1007/BF02281954]. *Journal of Nonverbal Behavior, 20*(3), 165-182.

Krokoff, L. J., Gottman, J. M., & Hass, S. D. (1989). Validation of a global rapid couples interaction scoring system. *Behavioral Assessment, 11*(1), 65-79.

Landis, J. R., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. [doi:10.1037/0003-066X.46.8.819]. *American Psychologist, 46*(8), 819-834.

Leifer, L., Plummer, R., Chong, J., & Toye, G. (2008). Pair Programming: Myths Dispelled by Evidence. Unpublished Paper. Stanford University.Unpublished manuscript, Stanford, CA.

Levenson, R. W. (1999). The intrapersonal functions of emotion. [doi:10.1080/026999399379159]. *Cognition and Emotion, 13*(5), 481-504.

Levenson, R. W., Carstensen, L. L., & Gottman, J. M. (1994). Influence of age and gender on affect, physiology, and their interrelations: A study of long-term marriages. [doi:10.1037/0022-3514.67.1.56]. *Journal of Personality and Social Psychology, 67*(1), 56-68.

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: Physiological linkage and affective exchange. [doi:10.1037/0022-3514.45.3.587]. *Journal of Personality and Social Psychology, 45*(3), 587-597.

Levenson, R. W., & Gottman, J. M. (1985). Physiological and affective predictors of change in relationship satisfaction. [doi:10.1037/0022-3514.49.1.85]. *Journal of Personality and Social Psychology, 49*(1), 85-94.

Lewis, K. M. (2000). When leaders display emotion: How followers respond to negative emotional expression of male and female leaders.

[doi:10.1002/(SICI)1099-1379(200003)21:2<221::AID-JOB36>3.0.CO;2-0]. *Journal of Organizational Behavior, 21*(Spec Issue), 221-234.

Losada, M. (1999). The complex dynamics of high performance teams. [doi: 10.1016/S0895-7177(99)00189-2]. *Mathematical and Computer Modelling, 30*(9-10), 179-192.

Losada, M., & Heaphy, E. (2004). The role of positivity and connectivity in the performance of business teams: A nonlinear dynamics model. [doi:10.1177/0002764203260208]. *American Behavioral Scientist, 47*(6), 740-765.

March, J. G., & Simon, H. A. (1958). *Organizations*: Oxford, England: Wiley.

Marks, M., Mathieu, J., & Zaccaro, S. (2001). A temporally based framework and taxonomy of team processes. *The Academy of Management Review, 26*(3), 356-376.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The Tie That Binds? Coherence Among Emotion Experience, Behavior, and Physiology. [doi:10.1037/1528-3542.5.2.175]. *Emotion, 5*(2), 175-190.

McDowell, C., Werner, L., Bullock, H. E., & Fernald, J. (2003). *The impact of pair programming on student performance, perception and persistence*. Paper presented at the Proceedings of the 25th International Conference on Software Engineering.

Mills, T. M. (1953). Power relations in three-person groups. [doi:10.2307/2087546]. *American Sociological Review, 18*, 351-357.

Miner, A. G., Glomb, T. M., & Hulin, C. (2005). Experience sampling mood and its correlates at work. [doi:10.1348/096317905X40105]. *Journal of Occupational and Organizational Psychology, 78*(2), 171-193.

Minneman, S. L. (1991). The social construction of a technical reality: Empirical studies of group engineering design practice. Unpublished Dissertation, Stanford University, Stanford.

Moore, C. C., Romney, A. K., Hsia, T.-L., & Rusch, C. D. (1999). The Universality of the Semantic Structure of Emotion Terms: Methods for the Study of Inter- and Intra-Cultural Variability. *American Anthropologist, 101*(3), 529-546.

Morris, M. W., & Keltner, D. (2000). How emotions work: The social functions of emotional expression in negotiations. [doi: 10.1016/S0191-3085(00)22002-9]. *Research in Organizational Behavior, 22*, 1-50.

Nawrocki, J., & Wojciechowski, A. (2001). *Experimental evaluation of pair programming*. Paper presented at the European Software Control and Metrics (Escom 2001).

Nicolescu, R., & Plummer, R. (2003). A Pair Programming Experiment in a Large Computer Course. *Romanian Journal of Information Science and Technology, 6*(1), 199-216.

Norman, D. A. (2004). Emotional design: Why we love (or hate) everyday things: New York, NY, US: Basic Books.

Nosek, J. T. (1998). The case for collaborative programming. *Commun. ACM, 41*(3), 105-108.

Patton, B. (1982). Eazy's Garage - Two-party (lawyers only) version (Dispute Resolution Case). Cambridge, MA, USA: Program on Negotiation, Harvard Law School.

Pelled, L. H. (1996). Demographic diversity, conflict, and work group outcomes: An intervening process theory. [doi:10.1287/orsc.7.6.615]. *Organization Science, 7*(6), 615-631.

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. [doi:10.2307/2667029]. *Administrative Science Quarterly, 44*(1), 1-28.

Priem, R. L., & Price, K. H. (1991). Process and outcome expectations for the dialectical inquiry, devil's advocacy, and consensus techniques of strategic decision making. [doi:10.1177/105960119101600207]. *Group & Organization Studies, 16*(2), 206-225.

Purdy, L. R. (1967). Organizational conflict: Concepts and models. [doi:10.2307/2391553]. *Administrative Science Quarterly, 12*(2), 296-320.

Roberts, N. A., Tsai, J. L., & Coan, J. A. (2007). Emotion elicitation using dyadic interaction tasks. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment.* (pp. 106-123): New York, NY, US: Oxford University Press.

Ruef, A. M., & Levenson, R. W. (2007). Continuous measurement of emotion: The affect rating dial. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment.* (pp. 286-297): New York, NY, US: Oxford University Press.

Russell, J. A. (1980). A circumplex model of affect. [doi:10.1037/h0077714]. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. [doi:10.1037/0022-3514.57.3.493]. *Journal of Personality and Social Psychology, 57*(3), 493-502.

Salleh, N., Mendes, E., Grundy, J., & Burch, G. S. J. (2009). *An empirical study of the effects of personality in pair programming using the five-factor model*. Paper presented at the Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement.

Sfetsos, P., Stamelos, I., Angelis, L., & Deligiannis, I. (2009). An experimental investigation of personality types impact on pair effectiveness in pair programming. *Empirical Softw. Engg., 14*(2), 187-226.

Simmel, G. (1908). Die quantitative Bestimmtheit der Gruppe. In D. Humblot (Ed.), *Soziologie. Untersuchungen über die Formen der Vergesellschaftung.* (pp. 32-100). Berlin.

Simons, T. L., & Peterson, R. S. (2000). Task conflict and relationship conflict in top management teams: The pivotal role of intragroup trust. [doi:10.1037/0021-9010.85.1.102]. *Journal of Applied Psychology, 85*(1), 102-111.

Solomon, R. C. The philosophy of emotions *Handbook of emotions* (pp. 3-15). New York, NY: Guilford Press.

Stapel, K., Knauss, E., Schneider, K., & Becker, M. (2010). Towards Understanding Communication Structure in Pair Programming. In A. Sillitti, A. Martin, X. Wang &

E. Whitworth (Eds.), *Agile Processes in Software Engineering and Extreme Programming* (Vol. 48, pp. 117-131): Springer Berlin Heidelberg.

Staw, B. M. (1975). Attribution of the "causes" of performance: A general alternative interpretation of cross-sectional research on organizations. [doi:10.1016/0030-5073(75)90060-4]. *Organizational Behavior & Human Performance, 13*(3), 414-432.

Staw, B. M., & Barsade, S. G. (1993). Affect and managerial performance: A test of the sadder-but-wiser vs. happier-and-smarter hypotheses. [doi:10.2307/2393415]. *Administrative Science Quarterly, 38*(2), 304-331.

Staw, B. M., Sutton, R. I., & Pelled, L. H. (1994). Employee positive emotion and favorable outcomes at the workplace. [doi:10.1287/orsc.5.1.51]. *Organization Science, 5*(1), 51-71.

Stefik, M., & Stefik, B. (2004). Breakthrough: stories and strategies of radical innovation. Cambridge, MS, USA: MIT Press.

Tabachnik, N., Crocker, J., & Alloy, L. B. (1983). Depression, social comparison, and the false-consensus effect. [doi:10.1037/0022-3514.45.3.688]. *Journal of Personality and Social Psychology, 45*(3), 688-699.

Tang, J. C. (1989). Gesturing in Design: A Study of the Use of Shared Workspaces by Design Teams. Stanford University, Stanford.

Thatcher, S. M. B., Jehn, K. A., & Zanutto, E. (2003). Cracks in Diversity Research: The Effects of Diversity Faultlines on Conflict and Performance. [doi:10.1023/A:1023325406946]. *Group Decision and Negotiation, 12*(3), 217-241.

Tran, V. (2004). The influence of emotions on decision-making processes in management teams. University of Geneva, Geneva, Switzerland.

Van Kleef, G. A. Emotion in Conflict and Negotiation: Introducing the Emotions as Social Information (EASI) Model (2007). Paper presented at the IACM 2007 Meetings.

van Kleef, G. A., De Dreu, C. K. W., & Manstead, A. S. R. (2004). The Interpersonal Effects of Anger and Happiness in Negotiations. [doi:10.1037/0022-3514.86.1.57]. *Journal of Personality and Social Psychology, 86*(1), 57-76.

Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team Diagnostic Survey. *The Journal of Applied Behavioral Science, 41*(4), 373-398.

Walle, T., & Hannay, J. E. (2009). *Personality and the nature of collaboration in pair programming*. Paper presented at the Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. [doi:10.1037/0022-3514.54.6.1063]. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. [doi:10.1037/0033-2909.98.2.219]. *Psychological Bulletin, 98*(2), 219-235.

Weingart, L. R., Olekalns, M., & Smith, P. L. (2004). Quantitative Coding of Negotiation Behavior. [doi:10.1163/1571806053498805]. *International Negotiation, 9*(3), 441-456.

Wiggins, J. S., Trapnell, P., & Phillips, N. (1988). Psychometric and geometric characteristics of the Revised Interpersonal Adjective Scales (IAS,ÄîR). [doi:10.1207/s15327906mbr2304_8]. *Multivariate Behavioral Research, 23*(4), 517-530.

Williams, L., & Kessler, R. (2003). *Pair Programming Illuminated*: Addison-Wesley.

Williams, L., Kessler, R. R., Cunningham, W., & Jeffries, R. (2000). Strengthening the Case for Pair Programming. *IEEE Softw., 17*(4), 19-25.

Yang, J., & Mossholder, K. W. (2004). Decoupling task and relationship conflict: The role of intragroup emotional processing. [doi:10.1002/job.258]. *Journal of Organizational Behavior, 25*(5), 589-605.

Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition.* (pp. 31-58): New York, NY, US: Cambridge University Press.