Affective Process in Multi-User Interaction and Trust in Shared Technology

By

Jie Xu

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Industrial and Systems Engineering)

at the

UNIVERSITY OF WISCONSIN – MADISON

2016

Date of final oral examination: 1/25/2016

The dissertation is approved by the following members of the Final Oral Committee:
Enid Montague, Associate Professor, Computing and Digital Media
John D. Lee, Professor, Industrial and Systems Engineering
Paula Niedenthal, Professor, Psychology
Lyn Van Swol, Professor, Communication Arts
Douglas A. Wiegmann, Associate Professor, Industrial and Systems Engineering

Dedication

This dissertation is dedicated to my parents and grandparents who raised me up.

Acknowledgements

I want to thank many individuals that supported me in my journey of pursuing my graduate education and completing this dissertation work. Without these individuals I would not have been reached my goals in this journey.

First of all, I would like to thank my advisor, Dr. Enid Montague. She has been a very supportive mentor throughout my 6 years of graduate school process. I have learned a lot from her in the discipline of human factors and ergonomics and human computer interaction. She also provided a lot of research opportunities that broadened my experiences and built the foundation of my future career.

I also would like to thank my other dissertation committee members, Dr. John D. Lee, Dr. Paula Niedenthal, Dr. Lyn Van Swol, and Dr. Douglas Wiegmann. They provided valuable guidance during the writing of this dissertation.

I must thank colleagues and students who provided support in my work in this dissertation: Ping-yu Chen, Dr. Onur Asan, Dr. Erin Chiou, Dhruv Mahta, Kim Le, and many others who worked in the HCI lab. They helped me in the data collection and data analysis; provided ideas, and most importantly, friendship.

Last but not least I must also thank my entire family and my many close friends who supported me in my study and life in the United States. I would not have been here today without them.

Abstract

The purpose of this dissertation research was to understand affect in team interactions and to investigate the effect of affective process on the individuals' trust in technology. Specifically, this research studied two-person teams that consisted of an active user and a passive user of a shared technology in a psychomotor multi-tasking environment.

Efforts have been made in previous research to understand affect in teams and the effect of affect in technology use. For example, research has shown that affect influences trust in technology thus could potentially impact the use of technology and performance of individual operators. However, little research was conducted with focus on team use of technology and differentiation of the effects of incidental affect and integral affect. This distinction is particularly important in teams since interactions among the team members and the technology can shape the integral affect of the team. Integral affect may have a stronger link to the team members' experiences with and attitudes toward the technology being used than incidental affect does.

This dissertation research consisted of two studies. The first study aimed to understand dynamics of integral affect in teams and the second study aimed to understand the effects of incidental affect and integral affect on trust in technology. In both studies, participants worked as two-person teams and multi-tasked under varied technology reliability and task difficulty levels. Participants' integral affect was measured by facial expression recognition analysis of videotaped data. In the first study, expertise of the passive user was manipulated by the level of training prior to the tasks. In the second study, incidental affect was manipulated by having the participants to view images of positive, neutral, or negative valence prior to the tasks; trust in technology was measured by surveys administered after the task.

In the first study, it was found that individuals worked in a same team had similar level of positive affect. Analysis performed using actor-partner interdependence model indicated that the two individuals worked in a same team influenced each other from a moment to moment basis. Task/technological conditions and level of training of the passive user influenced how the two users influenced each other. In the second study, it was found that integral affect mediated the effects of incidental affect and task/technological conditions on trust in technology. However, the main effect of incidental affect on trust in technology was not significant.

These findings provided insights into the affective process of team interactions with shared technology, revealed the role of affect in the mechanism of trust formation and calibration process, and had important implications for system design. In addition, the methods used in this study for modeling integral affect could be useful for future efforts in investigating emotional contagion and affect dynamics in teams.

Table of Contents

List of Figures

List of Tables

1. Chapter I: Introduction

1.1. Purpose of the Research

The purpose of this dissertation research was to understand the affective process in multi-user interaction with shared technology and how affect influence trust in technology. Systems involving active users and passive users of a shared technology (Montague & Xu, 2012) were investigated.

Traditional research in human technology interaction focuses primarily on the cognitive process due to behavioral and information processing traditions (J. D. Lee, 2006). As the dramatic progress in affect research in psychology in recent years (Frijda, 2008; J. J. Gross, 1999), there has been an increasing interest in affective process in the field of human factors and human computer interaction (Brave & Nass, 2003). For example, a series of studies conducted by Reeves and Nass (1996) showed that humans interact with technologies as if they were humans. Subsequent studies have shown that humans respond to computers affectively and their affective states can be significantly influenced by emotionally expressive computers (Brave, Nass, & Hutchinson, 2005; Klein, Moon, & Picard, 2002). Affective computing emerged as the study of technologies that can recognize, interpret, and express affect (Picard, 1999, 2000). In addition, Hancock, Pepe, & Murphy (2005) introduced the term of "Hedonomics" as the study of the promotion of pleasure, while ergonomics is the study of prevention of pain. However, there is a general lack of research in multi-user interactions and shared technology (Inbar & Tractinsky, 2009; Montague & Xu, 2012; Xu & Montague, 2012), and fewer studies of the affective process in such interactions.

This research aimed to contribute to the ergonomics side of affective human factors studies and applications. For example, in primary care, the use of electronic health records (EHRs) could affect clinician patient interaction thus influencing patient outcomes such as compliance to treatment (Ong, De Haes, Hoos, & Lammes, 1995) and satisfaction (Bensing, 1991). There may be undesirable consequences to patient care if a clinician gets frustrated with the EHR and transfers his/her negative affect towards the EHR to the patient during the clinical encounter. Design guidelines should be proposed to enable the EHRs to facilitate the positive interaction between the clinician and the patient to improve patient care. Since affect is an important part of interaction process in the clinic (Roter, Frankel, Hall, & Sluyter, 2006), one should understand how the physician and the patient's affective states influence each other during the visit, and how these affective states may influence the outcome of the visit. This research could provide insights to these types of questions. There also has been research on the group level that showed group affect influence group performance. For example, Cole, Walter, & Bruch (2008) found that negative mood is correlated with low performance in automotive manufacturing work teams. In another study, Gibson (2003) found that group affect is correlated with group efficacy in nurse teams. This research could also contribute to the understandings of affective process on the group/team level.

Another reason why affective human factors design is paid much attention is that "The hedonomics of economics is the economics of hedonomics". As Norman (2004) discussed, sometimes people choose attractive, emotional designs over efficient ones. In organizational design, research has been conducted to investigate how the "emotional labor" (such as service workers and salesperson) improves performance of an organization (Totterdell & Holman, 2003). This research could potentially contribute to the design of customer service encounters where the service worker and the customer are active user and passive user of a shared technology, correspondently (Inbar & Tractinsky, 2009). The affective process in the interaction is critical for customer trust and satisfaction.

1.2. Overview of the Background

1.2.1. Active user, passive user, and shared technology

In a system involves multiple individuals and shared technologies, the group of users of the shared technologies can have different roles or perspectives of the system. One way of characterizing different users is to introduce the concept of active and passive users.

A passive user is an individual who has limited direct control of technologies and artifacts in a work system (Montague & Xu, 2012). The opposite concept is the active user who actively controls the technological environment. Examples of systems that consist of both passive and active users include: customers (passive user) and service providers (active user) in face-to-face service encounters (Inbar & Tractinsky, 2010); Pilots not flying (passive user) and pilots flying (active user) in commercial airplane cockpits; students (passive user) and teachers (active user) in a classroom; and patients (passive user) and physicians (active user) in a clinical encounter (Montague, Winchester III, & Kleiner, 2010). A general conceptual model of active user and passive user with a shared technology is showed in Figure 1.

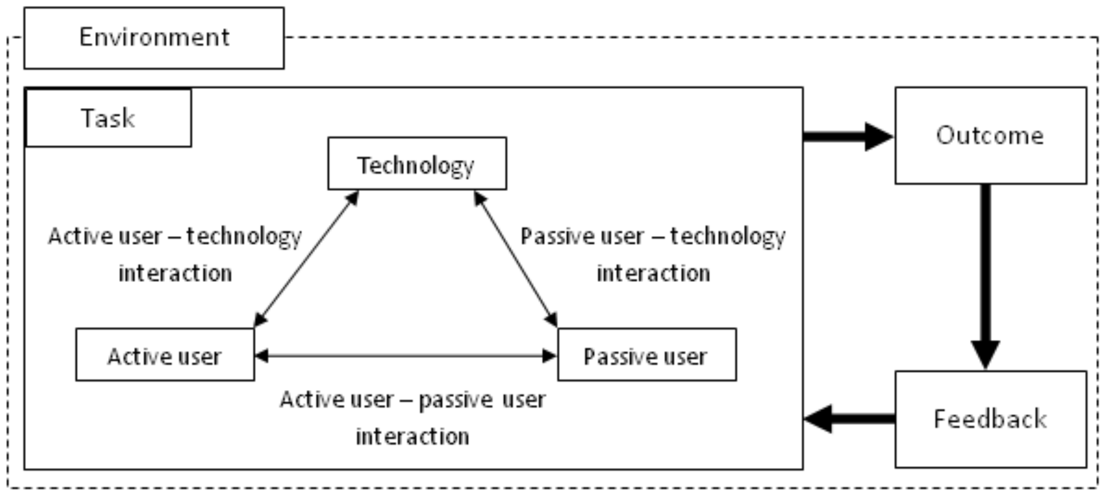


Figure 1. A general conceptual model of a system involving a passive user. Adapted from “Understanding active and passive users: The effects of an active user using normal, hard and unreliable technologies on user assessment of trust in technology and co-user,” by E. Montague and J. Xu, 2012, *Applied Ergonomics, 43*(4), 702-712.

In the cases mentioned above, passive users are at least as important as active users. For this reason, passive users should be treated as important stakeholders in the design process.

Incorporating design features for passive users, including sharing information with the passive user and increasing the passive user's degree of control over the technology, can optimize trust and effectiveness, thus improving satisfaction of the customers (passive users) in face-to-face service encounters (Inbar & Tractinsky, 2012). However, technologies are typically only designed for active users, such as operators, administrators, or technicians, rather than for passive users (Inbar & Tractinsky, 2009). When designing technologies to be used by active and passive users, some design requirements for active users can also be applied to passive users. For example, good data visualization can help the passive user understand the current state of the technology. However, situations involving three-way interactions that include the technology, the active user, and the passive user, are complicated and more research is needed to inform design requirements.

1.2.2. Trust in shared technology

Trust in technology is "the attitude that a technology will help achieve the user's goal in a uncertain and vulnerable situation" (J. D. Lee & See, 2004). Previous research has found appropriate use, misuse, and disuse of technology can happen depending on if the technology is trusted appropriately (Parasuraman & Riley, 1997). An individual's trust in technology may also influence his/her trust in other elements of the system, such as interpersonal trust and institutional trust (Muir, 1994). This issue is critical for industries such as healthcare (Montague & Lee, 2012) and e-commerce (M. K. O. Lee & Turban, 2001) where interpersonal trust and institutional trust are important. When multiple people or a group of users use a technology, trust in technology could be a factor that influences how the technology is used and the collaboration between group members. For example, whether or not the team trusts the technology appropriately can affect overall team performance (Bowers, Oser, Salas, & Cannon-Bowers, 1996). Trust in technology on the group level is especially relevant for multi-user shared technologies, such as health technologies shared by physicians, nurses, and patients, or interactive interface shared by a customer service representative and a customer, or robots used by a military team.

1.2.3. Conceptual model of the research

Many of the multi-user systems involving active users and passive users are affect critical systems, where affective process has an important impact on the outcome of those systems. Examples include customer service encounters such as primary care encounters. To guide this dissertation research, a framework bases on an input-process-outcome structure (Wittenbaum et al., 2004) was utilized as a general conceptual model. This framework is shown in Figure 2.

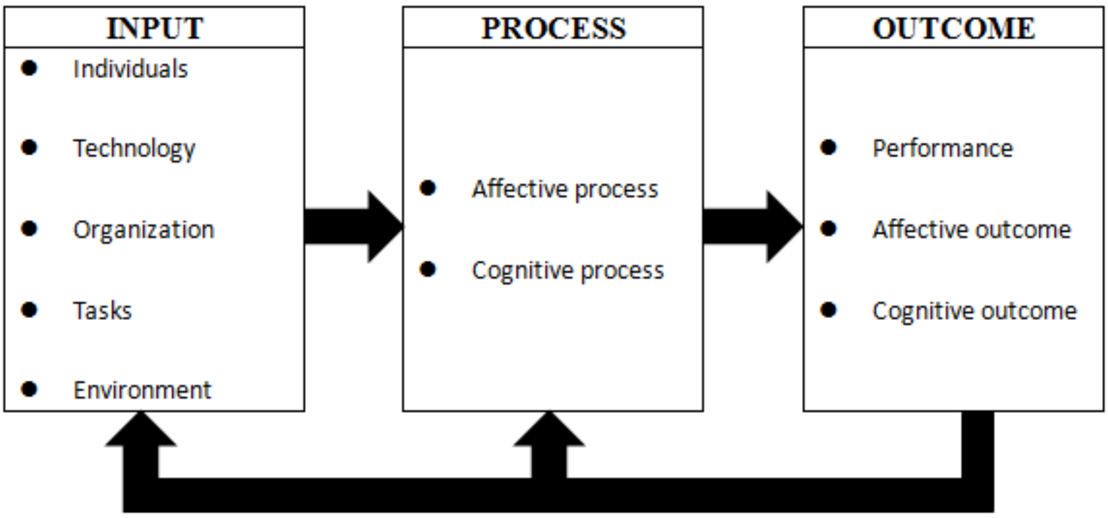


Figure 2. A general conceptual model guiding this dissertation research.

On the input side, it is the designed multi-user work system that could be viewed as consisting of five basic elements: individuals, technology, organization, tasks, and environment (Smith & Sainfort, 1989). These five elements interact and influence the process of interaction and further the outcome of the interaction (Carayon et al., 2006). For example, individual differences will lead to differed intensity in response to the same emotional stimuli (R. J. Larsen & Diener, 1987). Individuals' differing levels of expertise in the task or technology will lead to different level of perceived uncertainty of the situation, thus further lead to different emotional response to the situation (Brave & Nass, 2003; Ellsworth, 1994). Technology problems may make the users feel frustrated (Klein et al., 2002). The organization may require the employees to express certain emotion even if the employees do not feel that emotion (Barsade & Gibson, 2007). Poorly designed tasks may lead to employee burnout (Maslach, Schaufeli, & Leiter, 2001). Characteristics of the physical environment, such as color (Levy, 1984) and sound (Gregory, Worrall, & Sarge, 1996), can also influence individuals' mood.

During the interaction process, the affective process and cognitive process interact to affect the outcome of the interaction. On the individual level, there have been increasing amount of documentations of affect – cognition interaction. Affect can influence cognitive process such as attention, memory, decision making, problem solving (Brave & Nass, 2003; Izard, 2009; J. D. Lee, 2006; Lottridge, Chignell, & Jovicic, 2011; Niedenthal, Krauth-Gruber, & Ric, 2006). On the group level, it is known that distributed cognition (Hollan, Hutchins, & Kirsh, 2000; Nardi, 1996) and group affect (Barsade & Gibson, 1998; Kelly & Barsade, 2001) have important implications for group/team outcomes such as performance and satisfaction; however, there have been little research conducted to investigate the interaction between these two constructs.

1.3. Research Scope and Questions

The scope of this research was first to investigate the dynamics of affective process in a multi-user system involving active user and passive user. In this step, the focus was integral affect (Kugler, Connolly, & Ordóñez, 2012), which is the affective state of the users during the process of interacting with the shared technology and with each other. An experimental study aimed to answer the following research question:

Research Question 1: How the affective states of the individual users are influenced by each other during the interaction process?

In the second step of the research, how integral affect mediate the relationship between incidental affect and trust in technology was investigated. Incidental affect is the affective state of an individual before the interaction (Kugler et al., 2012), thus can be consider the input in the input-process-outcome model. Trust in technology was measured at the end of the interaction thus it was considered to be an outcome of the interaction. A second experimental study aimed to answer the following research question:

Research Question 2: How an individual's trust in technology is influenced by the affective process?

Since little research has been conducted to investigate the affective process in multi-user interaction with shared technologies, this dissertation research made an exploratory attempt in this field. Figure 3 shows the scope of the two studies in the proposed research.

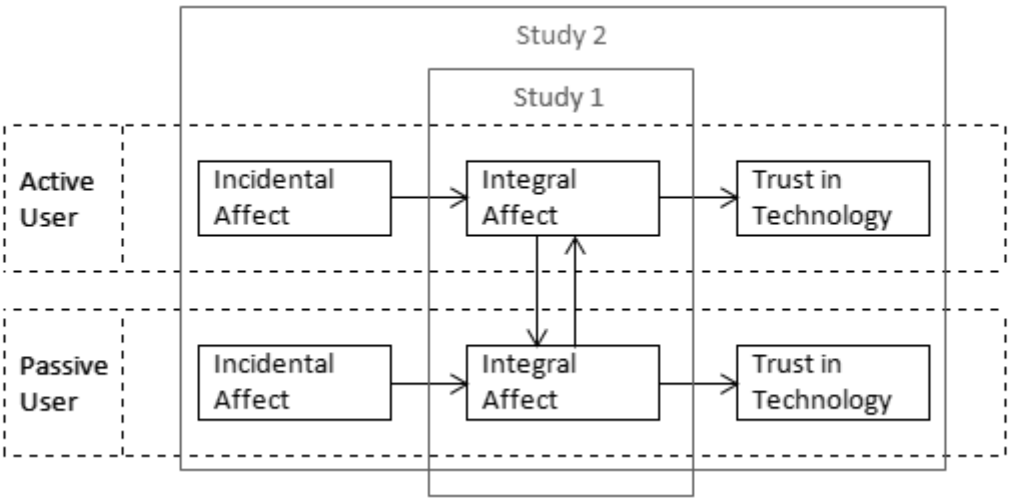


Figure 3. The scope of this research.

2. Chapter II: Background and Literature Review

2.1. Multi-User Interaction and Shared Technology

2.1.1. Shared technology: active user and passive user

In many human-computer interaction models an individual user interacts with a technology. In some of these models, users are divided into “active process operators” and “passive process operators” (Johansson, 1989; Persson, Wanek, & Johansson, 2001). According to Persson et al.’s (2001) definition, an active process operator’s work is distinguished from the passive process operator’s work by the predominance of monitoring tasks. If most of the tasks are monitoring tasks rather than tasks such as prediction, planning, control, etc., then the work could be considered a passive process operator’s work. This dichotomization of users has a continuous underlying dimension of “activeness.” The “activeness” of a user is a result of task allocation between the user and the technology. The level of control changes as different levels of automated technology are introduced into the work environment, from purely manual control to a fully automated process (Endsley & Kiris, 1995; Sheridan, 2002).

At the group level, the role of individual users can be further differentiated. In work system models, the system is comprised of people, technology, tasks, the organization, and the environment (Smith & Carayon-Sainfort, 1989). Individuals in the environment can have different roles or perspectives of the system. The “activeness” dimension in individual user – technology interaction can be expanded and applied to complex sociotechnical systems. A similar idea in characterizing different users is the concept of “incidental user” in customer service settings (Inbar & Tractinsky, 2009). This concept relates to the “activeness” dimension of technology usage. An incidental user is interested in the information output presented by the technology, but has little or no control over the technology. The user’s communication with the technology is mediated by another user referred to as an “active user” who has control over the technology (Inbar & Tractinsky, 2009). The term “active user” and “passive user” are defined as opposite cases on a continuous dimension of “activeness” in cooperative work. Passive users do not have full control of the technology, while active users have enough control to operate the technology. A passive user will be similar to a passive process operator in his/her relationship with the technology being used. However, the passive user may be faced with less monotony than the passive process operator, since sometimes the passive user can “control” the technology indirectly by communicating with the active user. In sum, the role of passive process operator in automated systems is characterized by indirect control of task process through automated systems; in group use of shared technology, a passive user of a shared technology is characterized by indirect control of the technology through the active user.

2.1.2. Examples of active and passive user interactions

A thorough discussion of passive users in face-to-face customer service encounters is discussed in Inbar and Tractinsky (2010). Some other examples of areas where passive users are important include aviation, education, and healthcare. In the cockpits of commercial airplanes, the captain and the first officer will alternate roles between Pilot Flying (PF) and the Pilot Not Flying (PNF). The PNF could be considered a passive user who takes a supportive role and attends to duties such as communicating with the air traffic controller and monitoring instruments (Hutchins, 1995). Monitoring instruments is an important duty for the PNF, and is associated with a high percentage of aviation safety problems (Sumwalt, Thomas, & Dismukes, 2002). Research indicates that the PNF is less likely to lose situational awareness than the PF. Thus, the captain of the aircraft should consider taking the role of the PNF in an emergency situation (Jentsch, Barnett, Bowers, & Sales, 1999).

In educational environments, students can be passive users if teachers are sole operators of the technology. For example, research shows that although interactive whiteboards are versatile and effective tools for teaching, student access is limited (Hall & Higgins, 2005). In other cases, technologies such as socialized computers, designed to facilitate collaborative learning, have advantages over individual learning techniques (Mori et al., 2009). Some students may take the role of a passive user when activities require taking turns to interact with the technology (Inkpen, Mcgrenere, Booth, & Klawe, 1997) or access to the technology is limited (Pal, Pawar, Brewer, & Toyama, 2006).

In healthcare, the patient is often a passive user. They receive information or treatment from health technologies operated by the clinician, who is the active user. Previous research indicates that doctor-patient interaction is important in influencing health outcomes (Pearson & Raeke, 2000). In healthcare, technologies and computer use can function as moderators in the doctor-patient interaction (Kaplan, Greenfield, & Ware, 1989; Woolley, Kane, Hughes, & Wright, 1978). In primary care offices, researchers have investigated the effects of computer use on patient outcomes. The results indicate that the physician's computer skill can significantly affect patients' ratings of satisfaction with the care they receive (Garrison, Bernard, & Rasmussen, 2002). To improve interactions, researchers have proposed that patient perspectives should be considered when designing system improvements (Delbanco, 1992) and for risk management (Itoh, Andersen, Madsen, Østergaard, & Ikeno, 2006).

2.1.3. Social requirements and shared technology

When the design and implementation of technology is focused on the group level, one needs to put more emphasis on social requirements. Ackerman (2000) discussed three important social requirements of designing technologies in computer supported collaborative work (CSCW) systems. First, social interactions are nuance and un-formalized. This highly flexible nature of social interaction requires the use of a flexible system which is not available most of the time. In consequence, the users often "redesign" the intention of the designer and use the system in unexpected ways (Orlikowski, 1992). For example, Asan and

Montague (2013) found that in primary care encounters, there are multiple ways that the information on the computer is shared between the physician and patient: a physician may actively turn the computer monitor to the patient and direct the patient to information that he/she wants to share; a physician may not actively share the information on the screen but the patient may actively lean in and search the screen for information. The design of the interface of the computer and the physical layout of the room may facilitate or inhibit such interaction styles.

Second, social interactions are dynamic and contextual. For example, people interact with each other in different modes in different contexts and they switch between modes often (Grudin, 1994). According to the time/space classification of collaboration (Bafoutsou & Mentzas, 2002; Desanctis & Gallupe, 1987; Grudin, 1994), the design of the technology for a multi-user system could be classified into four categories: same location and same time, same location and different time, different location and same time, and different location and different time. For example, traditionally healthcare service is provided in a face-to-face fashion; as the advances in information technology progress, distance collaboration is also possible. telemedicine technology was used for distance medical consultation (Perednia & Allen, 1995); telepresence technology enable the physician to use a robot to “walk around” intensive care units to monitor and interact with patients (Vespa, 2005); surgeries can also be performed by the collaboration of multiple surgeons in multiple different locations through robotic technology (Anvari, McKinley, & Stein, 2005; Cadiere, Himpens, Vertruyen, & Favretti, 1999). These different modes of collaboration will lead to different social requirements (O’Kane & Hargie, 2007). In addition, changes in personnel sub-system factors, such as degree of professionalism, demographic characteristics, or value system (Hendrick, 2002), could also change the way people interact with each other in the system. The system should be able to adapt to the dynamics of the changes in social interactions.

Third, different users often have different preferences and goals. For example, physicians and nurses have different concerns on how a medical error reporting system should work, in terms of voluntariness to report, what to report, whether to use an electronic system, etc. (Escoto, Karsh, & Beasley, 2006) Also, individuals have their own preferences for what information to be shared with others and how (Ackerman, 2000). The system should be able to adapt to the requirements of different stakeholders.

Inbar and Tractinsky's work about incidental users (Inbar & Tractinsky, 2009, 2010, 2012) expanded the third point and discussed the social requirements in a multi-user system involving active users and incidental users. Active users and incidental users have different degree of control, information need, and level of expertise (Inbar & Tractinsky, 2012). The system designer should calibrate the degree of control for the passive users and decide what information is shared and how it is shared, in order to optimize the collaboration between the passive users and active users.

Clegg's work (Clegg, 1993, 2000) further revealed social requirements in an organizational level. In team/group work situation, the design of the system should support task identity and task significance, fulfill the need for autonomy and provide clear feedback. According to the job characteristics model (Hackman & Oldham, 1975, 1976), these designs relate to work motivation, satisfaction and effectiveness. In addition, although senior managers and investors may not be the direct user of the system, the system should address their needs.

In a socio-technical system, the performance of the system is determined by how different sub-systems, such as personnel sub-system, technology sub-system, and organization sub-system, is designed to work together (Kleiner, 2006; Smith & Carayon, 2001). Social requirements emerge from the personnel sub-system and organizational sub-system, and need to be joint optimized. Designing for the performance of the system could also be understood in a hierarchical fashion: from physical requirements, information requirements, personal requirements, to social requirements; and these requirements should be addressed in hardware system, software system, human-computer interaction system, and social system, correspondingly (Whitworth, 2009). The performance of the system is often defined at the social system level (Whitworth, 2009) – that is, how the social requirements are addressed. The link between social requirements and system outcomes may be stronger in systems involving groups or teams since social interaction is taking a more important role.

Human factors and ergonomics mainly concern about performance, safety and satisfaction as system outcomes (Wickens, Lee, Liu, & Gordon, 2004). Social requirements relate to all three of these outcomes. For example, implementation of groupware may encounter the critical mass problem, that there should be enough users using the technology to motivate other users in the group to use the it (Ackerman, 2000; Markus, 1987). Otherwise, the anticipated performance level won't be achieved due to disuse of the system. Team members in distributed teams develop trust in each other slower than face-to-face teams due to the difficulty in acquiring social information to evaluate trust (Wilson, Straus, & McEvily, 2006). With low trust level, the team members will allocate resources to monitor each other thus hinders performance of the team (McAllister, 1995). Low trust also relates to low satisfaction and low willingness to stay in the team (Golembiewski & McConkie, 1975). In aviation, a large amount of accidents could be attributed to failures of crew resource management (Wiegmann & Shappell, 2001), which is related to the coordination of the team (E. Salas, Burke, Bowers, & Wilson, 2001; Wiegmann & Shappell, 2003). These failures may be related to that the social requirements of the interaction are not being met. It is also found that aviation system design that works in individualism culture (e.g. US) does not work in collectivism culture (e.g. China), due to a mismatch between the organizational design and social requirements raised by cultural values (Li, Harris, Li, & Wang, 2009).

Traditionally social requirements are addressed in macroergonomics research on the organization level and mesoergonomics research on the group level (Karsh, Waterson, & Holden, 2014). In an effort to better integrate microergonomics and macroergonomics, Karsh (2006) redefined mesoergonomics as "an open systems approach to ergonomic theory and research whereby the relationship between variables in at least two different levels or echelons is studied, where the dependent variables are human factors and ergonomic constructs". The studies of multi-user system with shared technology will need to be mesoergonomics research that the social requirements of such systems should be investigated on both the individual level and the group level. A series of previous experimental studies on active/passive users and shared technology investigated trust in technology and co-user (Xu & Montague, 2013b) and physiological response patterns (Xu & Montague, 2012) for the passive user on the individual level, and antecedents of trust in technology and physiological compliance (Montague, Xu, & Chiou, 2014; Xu, Le, Deitermann, & Montague, 2014) crossing the individual and group level. This research continued this direction and investigates the effect of affective process on trust in technology crossing the two levels.

2.2. Trust in Technology

2.2.1. Trust calibration

Human factors researchers have investigated the means for trust calibration (J. D. Lee & See, 2004) or how users develop an appropriate level of trust toward the technology. In order to develop effective means for trust calibration, one needs to first identify antecedents of trust in technology. Researchers have identified a wide variety of factors that influence an individual user's level of trust in technology, including reliability of the technology (Bisantz & Seong, 2001; J. D. Lee & Moray, 1992; Madhavan, Wiegmann, & Lacson, 2006), usability (Corritore, Kracher, & Wiedenbeck, 2003; Koufaris & Hampton-Sosa, 2002; Yuviler-Gavish & Gopher, 2011), visual design (Fogg et al., 2003; J. Kim & Moon, 1998; Weinstock, Oron-Gilad, & Parmet, 2012), and propensity to trust technology (Merritt & Ilgen, 2008), etc. Previous research also investigated social influence on trust in technology. For example, Xu and Montague (2013a) found that group polarization (Stoner, 1961; Sunstein, 2002; Van Swol, 2009) happened after group discussion of the trustworthiness of the technology. In a meta-analysis, Hancock et al. (2011) found that factors relate to team collaboration, such as culture, communication, and shared mental model, influence trust in robots of individuals in a team. Moods, which are heavily influenced by social interactions, are also found to be significantly related to trust in automation (Merritt, 2011; Stokes et al., 2010). However, more research is needed to investigate the effect of affective states on trust in technology in multi-user systems.

2.2.2. General and specific trust

The difference between general attitude and specific attitude is important for predicting behaviors (Ajzen & Fishbein, 1977). According to the compatibility principle proposed by Ajzen and Fishbein (1977), a specific behavior can be predicted by a specific attitude but not a general attitude; general attitude predicts aggregated behaviors across situations. Lee and See (2004) discussed functional and temporal specificity of trust in technology that describe the extent that an individual's trust maps to specific aspect of the technology and specific period of time. Here general trust in technology is defined as a general attitude towards the trustworthiness of a specific technology across functions and times.

General trust may influence an individual's perception of trust calibration information. According to social judgment theory (C. W. Sherif & Sherif, 1967; M. Sherif & Hovland, 1981), an individual evaluate a message not only by the merits of the information but also by their initial attitude. Specifically, an individual assimilates and contrasts a piece of information with his/her initial attitude as an anchor point (C. W. Sherif & Sherif, 1967; M. Sherif & Hovland, 1981). For example, if an individual has high trust towards a device, he/she may perceive a piece of information that indicates the device is moderately trustworthy as it indicates the device is more trustworthy than “moderately trustworthy”. This is the assimilation effect: the individual perceives the information to be more similar with his/her attitude than it actually is. On the other hand, contrast effect happens when an individual perceives the information to be more different with his/her attitude than it actually is. For example, if the individual in the previous example receives a piece of information indicating the device is moderately untrustworthy, he/she may perceive it as an indication of highly untrustworthy. As a result of assimilation and contrast effect, the development of specific trust is closely related to general trust.

According to confluence theory, elements of an individual’s mental system tend to flow together (Trafimow, 2009). For example, cognitive dissonance theory (Festinger, 1957) states that behavior changes attitude, while theory of reasoned action (Fishbein & Ajzen, 1975) states that attitude changes behavior. Rice, Trafimow, Keller, and Bean (2012) argued that general trust and specific trust always influence each other with a tendency to converge.

In this research, measurement of general trust in technology were used. The instrument for the measure was the scale developed by Jian, Bisantz, and Drury (2000).

2.3. Affective Process

2.3.1. Individual affect

Affect is a general term for the affective process and experience of human mind. Several terms are used for describing affect at more specific levels. These terms include (but not exclusively) emotion, mood, and sentiment (Brave & Nass, 2003; Kelly & Barsade, 2001). Emotion is a focused and intense affective state toward a specific stimulus in a relatively short time (Frijda, 1994). In comparison, mood has lower intensity and persists for a longer time than emotion, and mood is not directed to a specific stimulus (Frijda, 1994). Finally, sentiment could be defined as a property of an object that an individual is assigned to, as a result of appraisal (Frijda, 1994); for example, a person may like or dislike a specific object. In this research, "affect" was used as a general term combining emotion and mood.

Some researchers believe that emotions present in the mind in a relatively discrete form, consisting of basic emotions such as surprise, anger, sad, happy, disgust, and fear (Izard, 2009). Other researchers suggested that a higher order, two dimensional model (valence and arousal) is usually sufficient to describe affect (Barrett & Russell, 1999; Russell, 2009). In this research, the dimensional view of affect was adopted. Figure 4 shows a simplified model representing how the affective state of an individual changes over time.

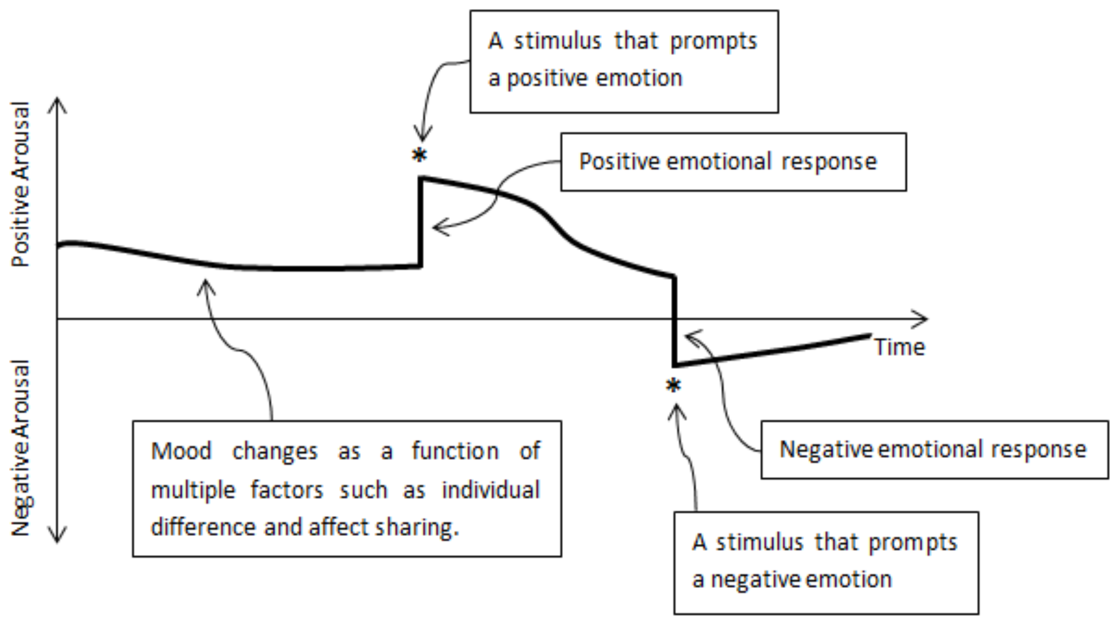


Figure 4. A simplified model depicting the affective process of an individual. Affect is represented as a value of the two dimensional map with the time dimension and the positive arousal - negative arousal dimension. Note that collapsing valence and arousal into a single positive arousal - negative arousal dimension to represent affect might be problematic (Watson & Vaidya, 2003).

2.3.2. Group affect

Barsade and Gibson (1998) proposed the construct of group affect which is a combination of affect of the individuals who form the group. Group affect is defined as "the affective state arising from a combination of the group's top-down components (i.e., the affective context) and its bottom-up components (i.e., the affective composition of the group) as transferred and created through explicit and implicit affective transfer processes" (Barsade & Gibson, 2012). One may argue that at a specific point of time, group affect is simply the aggregation of the affective states of the individuals and is described by measures such as mean and variance. However, if we view group affect as a dynamic process (Butler, 2011), we have to consider the change of affective state of each individual and how the affective states of one individual influences the affective states of other individuals in the group (affective composition; the bottom-up process). At the same time, we also need to consider the norms that govern how affect is expressed in a group (affective context; the top-down process). This norm may be induced by the norms and culture of an organization or the previous interaction history of the group (Kelly & Barsade, 2001). The concept of group affect is depicted in Figure 5.

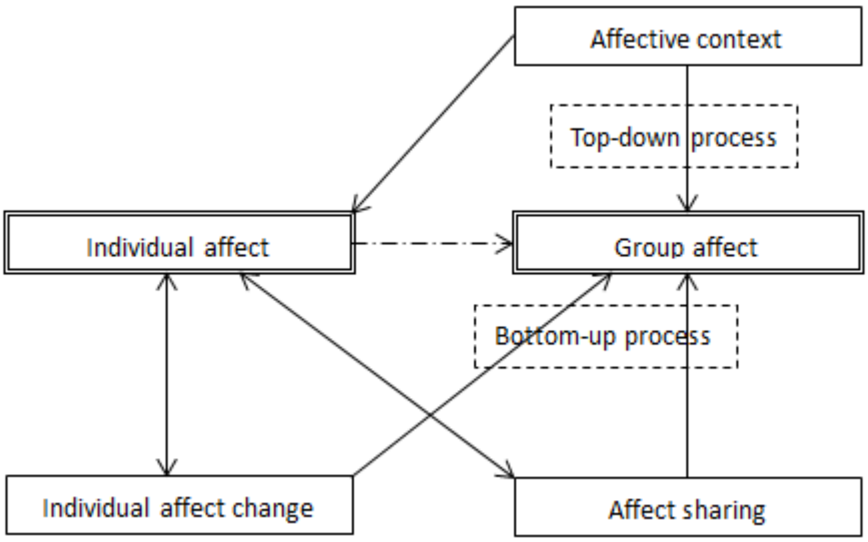


Figure 5. The concept of group affect.

In this research, the dynamics of group affect as a bottom-up process was investigated. Specifically, the individuals' affective states may change during the process of interacting with the shared technology; and the affective state changes in one individual may influence another individual. This affective process of the group may influence the individuals' perception about the system and their interaction with the system.

2.3.3. Measurement of affect

2.3.3.1. Different approaches to measure affect

The main methods for measuring affect include self-report measures, physiological measures, and observational measures (Mauss & Robinson, 2009; Wiles & Cornwell, 1991).

Rating scales of affect is sensitive to the dimensional framework of affect (Mauss & Robinson, 2009). Two of the most widely used scales are the Positive and Negative Affect Schedule (PANAS) (Tellegen, Watson, & Clark, 1988) and the Brief Mood Introspection Scale (BMIS) (Mayer & Gaschke, 1988). Lang (1995) also developed an popular pictorial based rating scale to measure affect on valence and arousal dimensions. Although these rating scales are easy to implement, they have disadvantages that they rely on conscious experiences of affect and the individual's memory (Brave & Nass, 2003). Continuous self-report measure is also obtainable through affect-rating-dials (Gottman & Levenson, 1985). When applied to current research, self-report measures could be good measurements for incidental affect but not for integral affect as they are intrusive to the task process.

Autonomic nerve system (ANS) measures, such as electrodermal activity and cardiovascular activity, are known to be sensitive to arousal that is related to emotion. It is reported that affect on arousal and valence dimensions by combining multiple ANS measures (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000), even basic emotions to some extent (J. T. Larsen, Berntson, Poehlmann, Ito, & Cacioppo, 2008). Recent developments in affect computing made progression in emotion recognition using ANS measures. In a study, Picard, Vyzas, & Healey (2001) achieved 81% correct recognition rate of basic emotions in a subject with their algorithm. 60% - 70% success rate was achieved when the approach was extended to multiple subjects (K. H. Kim, Bang, & Kim, 2004). These physiological measures provide continuous and un-intrusive measurements of affect, however, the results might be confounded by the individual's cognitive process such as changes in mental workload (Boucsein & Backs, 2000).

Affect is also measured using observable behaviors as indicators (Mauss & Robinson, 2009), these behaviors include: vocal tone (Planalp, 1998), facial expression (Ekman, Friesen, & Hager, 2002), and bodily posture (Van den Stock, Righart, & de Gelder, 2007). Bartel & Saavedra (2000) developed an instrument for observer rating for group affect on video-taped group interactions. This instrument showed high correlation with group members' self-report of affect (Bartel & Saavedra, 2000). Following the development in computer vision and related analysis techniques, automated continuous facial expression recognition is possible (Gunes & Schuller, 2012). For example, the computer expression recognition toolbox (CERT) developed by Littlewort et al. (Bartlett et al., 2006; G. C. Littlewort et al., 2011) was able to perform well even under the condition that was different with the condition in which the algorithm of program was originally trained (G. C. Littlewort et al., 2003). Given the setting of current study, using facial expression recognition approach to measure integral affect was most appropriate.

2.3.3.2.Computer expression recognition toolbox (CERT)

2.3.3.2.1. From human face detection to action units (AUs) recognition

CERT use video recording as input and the data is analyzed frame-by-frame. In each frame, human face is detected using the Viola-Jones approach (Viola & Jones, 2004), which is considered to be one of the most reliable algorithm for face detection in visible light spectrums (Reese, Zheng, & Elmaghraby, 2012). Then the facial features, including inner and outer eye corners, eye centers, tip of the nose, inner and outer mouth corners, and center of the mouth, will be detected using the approach proposed by Eckhardt, Fasel, & Movellan (2009). The facial features are then used as inputs to a support vector machine (SVM) to estimate the intensities of 19 facial action units (AUs) as defined by the Facial Action Coding System (FACS) (Ekman & Friesen, 1978), which is one of the most widely used facial expression coding system in behavior science.

CERT showed reliable performance in AU recognition in validation studies (Gwen C Littlewort, Bartlett, Salamanca, & Reilly, 2011; G. C. Littlewort et al., 2011) and it has been used for a number of studies in different fields. For example, Rossi, Fasel, & Sanfey (2011) used analyzed the facial expressions of a responder's facial expressions when he/she was given a proposal in an ultimatum game economic experiment. The out-of-sample-classification accuracy of 0.78 was achieved when SVM was used for training the AU values captured from CERT to predict the responses to fair or unfair proposals. Bosch et al. (Bosch, Chen, & D'Mello, 2014; Bosch et al., 2015; Y. Chen, Bosch, & D'Mello, 2015) conducted a series of studies to use CERT AU outputs to classify learning-centered affect, such as confusion and frustration, in computerized learning environments. Lalot, Delplanque, & Sander (2014) examined the effectiveness of different emotion regulation techniques using self-report of affect as well as facial expressions as captured by CERT. They found that facial expression analyses effectively captured that expression suppression (an individual consciously controls facial expression to be neutral regardless of the actual emotional state) significantly reduced the activation of the AUs.

2.3.3.3.2. Basic emotions recognition

CERT feeds the AU values to a multivariate logistic regression (MLR) classifier to perform emotion recognition and produces intensity values for six basic emotions (surprise, anger, sad, joy, disgust, and fear) and neutral expression.

CERT's ability to automate the analysis of video data make it possible to collect large amount of affect data effectively and un-intrusively. RUBI-5 (Malmir, Forster, Youngstrom, Morrison, & Movellan, 2013; Movellan, Malmir, & Forester, 2012), a prototype social robot being developed for research and enrichment of early childhood education environments integrated CERT facial expression analysis to identify which kind of activities the children liked the most. The researchers compared the results from CERT and ratings from 3 human judges. They found that the correlation between the CERT output and the average human ratings was significant ($r = 0.73$). The average agreement between CERT output and human ratings was 0.67, which is similar to the average agreement between the 3 human judges (0.68). The authors concluded that RUBI-5 could be a useful autonomous digital ethnographer in this aspect.

The basic emotions recognition feature of CERT was widely used in different research areas. Biel et al. (Biel, Teijeiro-Mosquera, & Gatica-Perez, 2012; Biel, Tsiminaki, Dines, & Gatica-Perez, 2013; Teijeiro-Mosquera, Biel, Alba-Castro, & Gatica-Perez, 2014) used facial expressions and verbal contents to predict Big Five personality ratings from vloggers in their YouTube videos. Schaafsma et al. (2014) used CERT to investigate individuals' emotional response when they were exclude from a three person communication. Donkor et al. (2014) attempted to measure "emotional validity" of driving simulators using CERT in combination with questionnaires, postural movements, and physiological measures, such as heart rate and respiration rate. In this preliminary investigation, they found that study participants' emotional responses were similar to those reported in previous studies which were done on the road.

CERT also has been a useful tool in clinical assessment and intervention. Deriso et al. (Deriso, Susskind, Krieger, & Bartlett, 2012; Deriso, Susskind, Tanaka, et al., 2012) designed an intervention system called Emotion Mirror to help autism children improve facial expression perception and production. The intervention system is a game that a child will be rewarded when he/she is able to mimic the expression of a cartoon character. CERT was used as a critical component in this system to recognize the expressions of the player to judge the mimicry. In an effort to develop automatic audiovisual behavior descriptors for depression assessment, Scherer et al. (S. Scherer, Stratou, & Morency, 2013; Yu et al., 2013) used CERT outputs in combination with acoustic features as inputs for a classifier. CERT outputs used included neutral expression, emotion variability (the variation of the values of neutral and all six basic emotions), and head rotation. These outputs were considered as visual cues depression. The authors found that using visual or acoustic cue alone, the classification accuracies was 64.10% and 51.28%, respectively; combining both cues, the accuracy reached 89.74%.

2.3.3.4.Measuring affect in teams

There are two different classes of group affect measurement as a result of the top-down process and bottom-up process of team affect; they can be labeled as group-as-a-whole approach and sum-of-its-parts approach (Barsade & Gibson, 1998). The group-as-a-whole approach involves measuring group affect directly without the need to aggregate affective values from individual group members. For example, Bartel and Saavedra (2000) developed an instrument for outside observers to rate group mood of work groups. This instrument allowed observers to record facial, vocal, and postural affective cues exhibited by any of the work group members. The results obtained from this instrument were consistent with aggregated self-reports from the work group members. The same instrument was used by Barsade (2000) in videotaped team interactions and demonstrated high inter-rater reliability. The direct measurement of group affect could be considered a major advantage of the group-as-a-whole approach. However, some sort of aggregation still takes place in the current available methods. For example, the instrument developed by Bartel and Saavedra (2000) required observers to aggregate behavioral cues across group members during the observation. There is little evidence showed that this method has advantage over observing individual behaviors and aggregate afterwards. A second limitation of the group-as-a-whole approach is that current available methods are labor intensive that they require intensive manual coding of observed behaviors.

The sum-of-its-parts approach involves the measurement of affective states of the individual group members and the aggregation of the individual values to represent the group affect. Three major methods could be found in the literature: mean aggregation, variance aggregation, and extreme values (Barsade & Gibson, 1998). For example, Cole, Walter, and Bruch (2008) measure work team negative affective tone by median aggregation of the individual team member ratings and found that collective emotion played a significant role in team performance. A review of literature (Barsade & Knight, 2015; Collins, Lawrence, Troth, & Jordan, 2013) showed two limitations in previous research utilizing sum-of-its-parts approach: first, more guidelines are need to decide when and how to aggregate individual scores; second, few research measured affect at single time points so dynamic temporal aspect of affect was not addressed. Thus there is a need to: first, develop affect measurement instruments that are able to capture affect dynamics and are unobtrusive to group interaction; second, develop valid and reliable aggregation methods to generate group-level measures of affect.

2.4. Affective Process and Trust

2.4.1. The affect infusion model

The affect infusion model proposed by Forgas (1995) is a comprehensive model that describes the affect-cognition interaction on judgments and this model have received substantial empirical support (Dunn & Schweitzer, 2005; Forgas, 2001). The affect infusion model proposed that there are four types of processing strategies one may potentially use depending on the characteristics of the target of judgment, the individual, and the situation. If the judgment target is familiar and the importance of the judgment is low, one will use direct access strategy which prior stored judgment is retrieved and used. If there is a specific way of gathering and processing information for a specific goal, an individual will use motivated strategy. These two strategies are labeled low infusion strategies which affect has little effect during the process. Under these two conditions, trust evaluation is directly retrieved or formed under a very specific way (for example, an individual may be instructed to compare the performance data of two devices and to trust the device which has higher average performance), thus the effect of affective state is minimal.

The two high infusion strategies are influenced by affective states and are particularly relevant to trust formation and calibration. If the target of judgment is unfamiliar and complex, the judgment task is important, and there is cognitive resource available for an individual, the individual may adopt substantive processing strategy that the individual acquire, analyze, and interpret information and select a decision. If the target of judgment is less complex, or there is insufficient cognitive resource available, an individual may adopt heuristic processing strategy that judgment is made with a small amount of information and simple rules. These two processing strategies roughly resemble system 2 and system 1 of the dual process theory (Kahneman, 2011). One important finding about the selection between these two processing strategies is that affect will influence which strategy is to be used (Forgas, 1995). In general, positive affect promotes the use of heuristic process strategy and negative affect promotes the use of substantive process strategy (Fiedler, Asbeck, & Nickel, 1991; Schwarz, 2000).

A large amount of empirical studies have been documented showing that affect influences information processing during substantive processing of information, such as attention, perception, memory retrieval, and decision selection (J. D. Lee, 2007; Niedenthal et al., 2006). This effect is partly due to the activation of an affective state also activate memories associated with this affective state in the associative network, resulting an affect priming effect (Bower, 1981; Niedenthal, Setterlund, & Jones, 1994). For example, in perception, a study found that participants in happy state make lexical decision on happy words faster than participants in sad state; while participants in sad state make lexical decision on sad words faster than those who are in happy state (Niedenthal & Setterlund, 1994). Eich and Metcalfe (1989) demonstrated the mood state dependent memory effect that retrieval of memories was enhanced if the mood state of the participants during the encoding process matched with the mood state during the retrieval process. As a result of the affect priming effect, trust evaluation using the substantive process strategy may be systematically biased.

Trust that is formulated or calibrated with heuristic process strategy is also influenced by affective state according to the affect-as-information model (Clore et al., 2001; Schwarz, 1990). This model proposes that when judgment is made through heuristic processing route, one's affective state may be used as a piece of information that is relevant to the judgment task. For example, an individual may evaluate his/her life satisfaction using his/her current mood which is influenced by current weather (Schwarz & Clore, 1983).

2.4.2. The effect of affective process on interpersonal trust

Literature suggests that there are similarities between interpersonal trust and technology trust (P. A. Hancock et al., 2011; J. D. Lee & See, 2004; Madhavan & Wiegmann, 2007). Thus it is valuable to review the research that examined the relationship between affect and interpersonal trust.

According to the affect infusion model, as long as a judgment is made under a high infusion strategy (substantive processing strategy and heuristic processing strategy), there is a main effect of affect on the judgment: positive affect will lead to a positive judgment; negative affect will lead to a negative judgment. This prediction is supported in the studies conducted by Dunn and Schweitzer (2005) that positive incidental affect increased trust in another person and negative incidental affect decreased trust. They also found that if the participants aware that their affective state is unrelated to the judgment task, then their affective states do not influence trust. The main effect predication is also supported in another study (Gino & Schweitzer, 2008) that the researchers found that positive incidental affect made people more receptive to the advice provided by others while negative incidental affect made people less receptive. In yet another study (Ferrin, Bligh, & Kohles, 2007), the researchers found that participants in negative incidental affect rated a set of human faces as less trustworthy comparing to participants who were in positive or neutral incidental affect.

Another prediction made by the affect infusion model is that an individual is more likely to use heuristic processing strategy when he/she is in a positive affective state and use substantive processing strategy when in a negative affective state. A series of experiments conducted by Lount (2010) showed that when the participants were in a positive mood, they were more likely to evaluate trust according to schemas and stereotypes. For example, they tended to send less money to an out-group player in a trust game. The researcher also generated faces with stereotypical trustworthy features (e.g., large round eyes) or untrustworthy features (e.g. narrow eyes) and had the participants to rate their trustworthiness. The results indicated that the participants in a positive mood state had higher reliance on the facial features to judge trust comparing to participant in a neutral mood state. Interestingly, these results actually contradict with the predictions of the affect-as-information model. It is possible that in this particular study the stereotypical cues were made salient so that the participants were more likely to use the stereotypical cues instead of affective cues.

Studies also had conflicting findings about the influence of emotion with different appraisals on trust. Dunn and Schweitzer (2005) found that other-person control emotions (anger and gratitude) influence trust more than personal control emotions (pride and guilt) or situational control emotions (sadness). However, Myers and Tingley (2011) did not find anger (other-person control) have an effect on trusting behavior in trust games while anxiety (situational control) decreased trust.

2.4.3. Affective process and trust in technology

There are some recent studies investigated the effect of affective states on individual operator's trust in technology. In the study conducted by Stokes et al. (2010), the relationship between induced mood states and an individual's level of trust in an automated decision aid was tested. 76 participants sampled from an Air Force base participated in this study. Before performing tasks on the Convoy Leader platform (Lyons, Stokes, Garcia, Adams, & Ames, 2008) with an automated decision aid, the participants were induced with positive or negative mood state using the International Affect Picture System (Lang, Bradley, & Cuthbert, 2008). The results indicated that positive mood state resulted in a higher level of trust in technology than negative mood state; but this difference was statistically significant only in the first experimental session, not in the second and third sessions. This finding is interesting that even for participants that had high expertise in relevant fields, their mood states still had an effect on trust in technology. However, there were limitations in this study. First, a baseline mood state condition was lacking so it was not able to determine if the positive mood state increased trust, or the negative mood state decreased trust, or both. Second, the mood states were induced in the training session of the experiment so it was unknown that if participants' mood states changed during the course of the subsequent experimental sessions and this might be the cause of the non-significance in trust levels.

In the study conducted by Merritt (2011), 130 students participants performed X-ray screening task where they were asked to identify weapons among other items in a luggage with the assistance of an automated decision aid. Before the experimental sessions, mood inducing videos with contents that were irrelevant to the experimental task were shown to the participants. The results indicated that mood states overall did not correlate with participants' level of trust in technology. However, “happy” as a specific emotion state correlated with higher level of trust in technology comparing to other emotion states. Although this study suffered similar limitations as the Strokes et al.’s study, it offered additional insights into the effect of affective states on trust in technology. It demonstrated that affective state could affect trust in technology not only in the initial stage of the interaction but also for the subsequent interactions, though the length of the experiment was not particularly long (a total of 60 task trials). This study also implied that there may be some differences in using the discrete emotion approach and the dimensional emotion approach to understand affective process in trust in technology.

Positive mood was found to potentially benefit trust calibration in a recent study by Niederée et al. (2012). In this study, 24 student participants were asked to monitor an aircraft automation which was design for maintaining the aircraft’s pitch angle. The participants needed to react to abnormal pitch angles. In this task, handling errors (repeated unnecessary reactions) to an abnormal event implied a potential breakdown in human automation cooperation and further a potential implication for inappropriately calibrated trust. The results of this study showed that positive mood correlated with lower number of handling errors. This implied that positive mood might be potentially linked to appropriately calibrated trust in the context of this study.

2.4.4. Directions in the research about affective process and trust in technology

This brief review of studies about the relationship between affective process and trust indicated some possible directions for expanding the understandings related to this topic. In the reviewed studies, little insight was provided for the effect of integral affect on trust. This is the case for both the studies in interpersonal trust and trust in technology. However, a complete understanding of the effect of affective process on trust in technology could not be achieved without investigating integral affect. The process of trust formation and calibration involve iterative task performing and interactions with the technology. During the interaction process, technological condition and task condition can influence an individual's affective state. For example, frustration and anxiety often result from an unpleasant interaction with a computer (Klein et al., 2002). Task performance can also alter mood state that there is a positive correlation between task performance and post-task mood (Rozell & Gardner III, 2000). If incidental affect is found to influence trust in technology, the integral affect which is a result of technology interaction and task performance may also influence trust in technology. With the unobtrusive measurements of affect (for example, autonomic nerve system measures or facial expression recognition techniques), it is possible for researchers to measure integral affect and investigate its relationship with trust in technology. This proposed research was an attempt in this direction to investigate the "incidental affect – integral affect – trust in technology" relationship in an input-process-outcome framework.

3. Chapter III: Hypotheses

3.1. Study One

Literature suggests that affect can be contagious that one individual's affective state can influence another individual's affective state. This process is known as emotional contagion (Hatfield, Cacioppo, & Rapson, 1992). Emotional contagion is the process that an individual converge his/her affective state with another individual by mimicking and synchronizing affective expressions (Hatfield et al., 1992; Hatfield, Cacioppo, & Rapson, 1993). As a result, the individuals converge affectively. Previous studies found this affect convergence effect in romantic couples (Anderson, Keltner, & John, 2003) as well as work teams (Totterdell, Kellett, Teuchmann, & Briner, 1998). The mimicry, synchronization, or entrainment of affective expressions of others is found in facial, postural, and verbal behaviors (Kelly, 2001; Kelly & Spoor, 2006). For example, a research found that mirror neuron triggered the activation in the limbic system which makes an observing individual feels the emotion facially expressed by the observed individual (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003). Based on these results, two hypotheses were proposed. The first hypothesis stated the tendency for the individuals to converge affectively in general:

Hypothesis 1.1: the affective states of the individuals worked in a same team are more similar than that of the individuals did not work in a same team.

The second hypothesis stated that the affective states of the individuals influence each other during the interaction process:

Hypothesis 1.2: an individual's current affective state is positively related to another individual's previous affective state.

In other words, if affective state was measured on two time points (t1 and t2), then an individual's affective state on t2 would be influenced by the other individual's affective state on t1. This described process of the affective convergence of two individuals.

Group composition factors may influence group integral affective process. This study investigated two of these factors that are particularly relevant to multi-user systems: roles (active/passive user) and expertise of the team members. These are two of the three critical attributes (degree of control, familiarity of the technology, and level of interest in the task) of a user in multi-user shared technology systems discussed by Inbar and Tractinsky (2012). It was hypothesized in this study that:

Hypothesis 1.3: in terms of affective states, the active user's influence over the passive user is stronger than the passive user's influence over the active user.

A previous study found that communicating with the active user is one of the major ways for trust in technology calibration for passive user while communication doesn't relate to active user's trust calibration (Montague & Xu, 2012). This result indicated that the active user has strong influence over the passive user on this affective outcome. It may further imply that the active user has strong influence over the passive user in terms of affective state. As the passive user observes the interaction between the active user and the technology, this observer role may facilitate the passive user's mimicking and synchronizing the expressions of the active user. In a study of eye gaze patterns in clinical encounters (Asan, Montague, & Xu, 2012), the researchers found that the patient's gaze at the computer follows the physician's gaze at the computer, while physician's gaze does not follow the patient's gaze. This behavior pattern may partially due to the fact that the physician is the active user and the patient is the passive user of the computer. However, the expertise differences between the physician and the patient may also lead to this "following" behavior of the patient. If a passive user had low expertise, he/she may have higher uncertainty about the situation and may be more likely to observe the active user to evaluate the situation. Thus in this case, the passive user is more likely to be influenced by the active user affectively. So in this study, it was hypothesized that:

Hypothesis 1.4: in terms of affective states, the active user's influence over the passive user will be stronger when the passive user has lower expertise than when the passive user has higher expertise.

3.2. Study Two

Lee and See (2004) proposed that affective process has a dominant influence on trust in technology. Some recent studies have reported that an operator's level of trust in technology is related to incidental affect of the operator (see section 2.4 of this proposal). However, Stokes et al. (2010) only found the significant relationship between incidental affect and trust in technology in the initial task session; in the subsequent task sessions, this relationship becomes insignificant. It may be partially due to the mediation effect of the integral affect on the relationship between incidental effect and trust in technology. In the context of multi-user systems with shared technology, this mediation effect may be more significant. This is because integral affect may be influenced by social interaction between the users in addition of interacting with the technology individually (study one). Thus, in this study, it was hypothesized that:

Hypothesis 2.1: initial mood state (incidental affect) influences an individual's affective state during the interaction process (integral affect).

In this study, it was also hypothesized that integral affect partially mediates the relationship between incidental affect and trust in technology (see section 2.4 of this proposal). Thus:

Hypothesis 2.2: initial mood state (incidental affect) affects trust in technology.

Hypothesis 2.3: affective state of an individual during the task (integral affect) mediates the effect of initial mood state (incidental affect) on trust in technology.

4. Chapter IV: Study One – Integral Affect of the Team

4.1. Methods

4.1.1. Participants

A Total number of 72 participants participated in this study. The sample was recruited from the University of Wisconsin – Madison. The participants received information about the study either through an introductory human factors engineering class or an online posting on the Student Job Center (https://jobcenter.wisc.edu/) website. The participants were randomly assigned to two-person teams (n = 36) for the experiment. All the participants received $10 compensation for their participation. In addition, monetary incentive was offered for team performance. The four teams who achieved the highest performance were awarded $20 per team member.

The participants ranged in age from 18 to 28 (mean = 21.60, SD = 1.96). Thirty-one of the participants were female (43.06%). There were 10 female-female teams (27.78%), 15 male-male teams (41.67%), and 11 mixed gender teams (30.56%). Fifty-one participants (70.83%) reported that they are Caucasians, and 13 participants (18.06%) reported that they are Asian. Other reported ethnicity groups included African American, Native American, and multiracial. The majority of the participants (53, 73.61%) were engineering students, majored in disciplines such as industrial engineering, mechanical engineering, biomedical engineering, and civil engineering. Other participants majored in a variety of disciplines, such as psychology, communication arts, economics, and anthropology. The majority of the participants were in their third (25, 34.72%) or fourth year (23, 31.94%) of college, and there were 12 participants (16.67%) who were attending graduate school. Finally, participants in 19 teams (52.78%) reported that they did not know their teammate prior to the experiment.

4.1.2. Task and setting

4.1.2.1. Task

The experiment required the participants to multi-task on a shared computerized platform as teams. The participants performed the tasks in a updated and revised version of the Multi Attribute Task Battery (MATB) program (Comstock & Arnegard, 1992) which runs on a PC Windows XP operating system. The MATB contains three tasks: a monitoring task, tracking task, and a resource management task. The interface of MATB is showed in Figure 6. The monitoring task located on the upper left panel and required participants to respond as quickly as possible to the two lights and the fluctuating four dials via keystrokes on the keyboard. The tracking task located on the upper center panel and required participants to maintain the position of a randomly moving target (a green circular object) in the center of the task area using a joystick. The resource management task located in the lower center panel and required participants to control several pumps to maintain optimum liquid level in two tanks. The panel on the lower right showed the flow rates of the pumps. The participants had to perform all three tasks simultaneously.

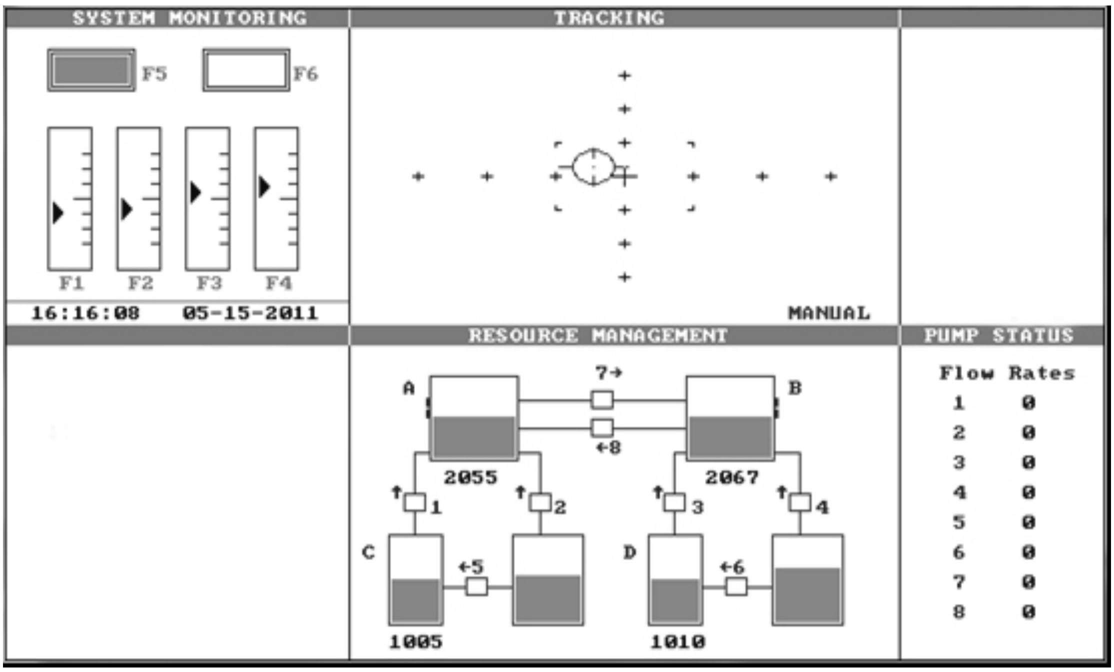


Figure 6. The interface of the updated and revised Multi Attribute Task Battery (MATB) program.

4.1.2.2. Apparatus and laboratory setting

There were a total number of two computer work stations set up for the experiment. The computer work stations were equipped with 23-inch monitors and keyboards, mice, and joysticks as control devices. During the individual training session, each participant occupied one computer work station to perform the tasks individually. During the experimental trials, the participants worked together in one shared computer work station. This is illustrated in Figure 7. The active user was seated on the left side of the work station with control devices in the front. The passive user was seated on the right side of the work station. Two webcams were placed on top of the computer screen to capture the face of the participants. A Biopac MP150 Data Acquisition System was placed behind the seats of the participants to collect the electrocardiogram (ECG) data from the participants. However, the data collected from this device was not included in this study.

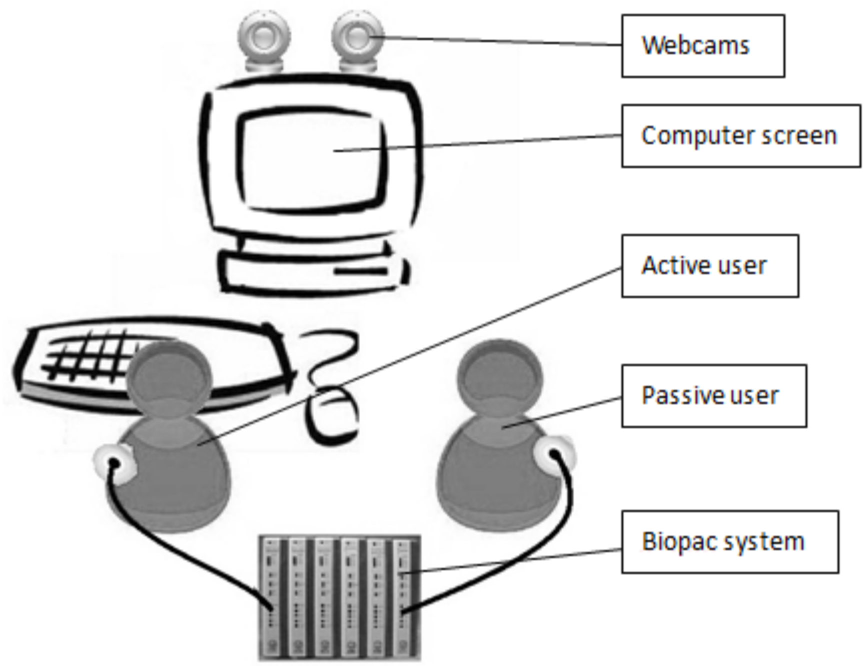


Figure 7. The setting for the two participants during the experimental trials.

MATB programs were installed on the computers in the work stations. Time stamps of the MATB program were generated according to the system time. The system times of the two computers were synchronized with internet time and the synchronization was performed right before each experiment. The webcams and the Biopac system were connected to another computer located in a separate room of the laboratory. The two video streams from the two webcams were recorded and synchronized using the Observer XT 9.0 program (Noldus Information Technology, 2009). The Observer program used system time as time stamps and the system time was synchronized with internet time right before each experiment.

4.1.3. Design

This study was a mixed design with two between-subject variables and one within subject variable. The first between-subject variable was the role of participants. As the participants arrive as a two-person teams, one of the participants was randomly assigned as the active user, and the other participant was assigned as the passive user. The active user had full access to the control devices of the computer, including the keyboard and the joystick. The passive user did not have access to any of the control devices but he/she could monitor the tasks and communicate with the active user to assist with the tasks. The participants were instructed that all the responsibility of the task, such as decision making and planning, except physical control, would be shared equally between the two group members.

The second between-subject variable was the expertise of the passive users. This variable was manipulated on the team level. In the high expertise condition, both the active user and the passive user in a team received training of the tasks before the experimental trials. Thus passive user had hands-on experience with the tasks in this condition. In the low expertise condition, only the active user received training before the experimental trials. Thus the passive user did not have any experience of controlling the computer throughout the experiment.

The within-subject variable of this study was named task/technological conditions, which consisted of three levels: normal condition, hard condition, and low reliability condition. The participants went through these three levels in three task trials. In the normal condition, the demand level in the monitoring task and difficulty level in the tracking task was low, and the reliability of the pumps in the resource management task was high. In the hard condition, the demand level in the monitoring task was high, in that more frequent responses from the user were needed. The tracking task was more difficult, in that the speed and movement amplitude of the target were increased. All the parameters of the resource management task were same as in the normal condition. In the low reliability condition, the difficulty level in the monitoring task and the tracking task was the same as in the normal condition. However, the pumps in the resource management task became unreliable, that certain pumps would “fail” at times. When a pump fails, its flow rate becomes zero and turns red in color. The participant was not able to control the flow rate of the pump when it failed. Thus it was difficult for the participants to maintain the desired liquid level in this condition. In sum, the normal condition could be a baseline for group performance, the hard condition could alter the group’s performance in the monitoring task and the tracking task, and the low reliability condition could alter the group’s performance in the resource management task. The manipulation of this variable is summarized in Table 1. Previous studies utilizing the same experimental treatment (Xu, Le, Deitermann, & Montague, 2013) showed that the team performance differences in the three conditions were consistent with what was expected; in addition, low reliability condition lad to significantly lower level of trust in technology than the other two conditions. Thus, no manipulation check was performed.

Table 1: Manipulation of the within-subject variable – task/technological conditions.

MATB tasks	Levels of task/technological conditions		
	Normal condition	Hard condition	Low reliability condition
Monitoring task	Standard demand	High Demand	Standard Demand
Tracking task	Standard difficulty	High difficulty	Standard Difficulty
Resource management task	Reliable pumps	Reliable pumps	Unreliable pumps

The sequence of presenting the three task/technological conditions were counter-balanced across the participated teams. The overall design of the study is shown in Figure 8.

Overall study:

Expertise		Task/technological conditions (counterbalancing sequence)	Number of groups
High	n = 36 (18 groups)	N – H – L	3
		N – L – H	3
		H – N – L	3
		H – L – N	3
		L – N – H	3
		L – H – N	3
Low	n = 36 (18 groups)	N – H – L	3
		N – L – H	3
		H – N – L	3
		H – L – N	3
		L – N – H	3
		L – H – N	3

Within each group:

Role	Active user	Passive user
Number of participants	1	1

Figure 8. Design of the experiment of study one.

4.1.4. Procedure

Before the experiment, all the participants filled out an online pre-experiment survey which measured the participants' individual characteristics such as propensity to trust technology (Singh, Molloy, & Parasuraman, 1993) and the Big Five personality (John & Srivastava, 1999).

Upon arrival, the following general introduction of the experiment was given to the participants:

> *"Thank you for participating in this experiment. I am the research assistant who is responsible for running this experiment. The purpose of this experiment is to investigate the interaction process of two persons with a shared computer. Thus this experiment will require the two of you to work as a group to accomplish some tasks on a computer. You two will have different roles. One of you will be the active user and the other of you will be the passive user. During the experiment, the active user will have access to the keyboard and joystick of the computer and is responsible for performing the inputs. The passive user will not have the access to the control devices. But the passive user can assist the active user with the task through communication. In sum, the active user and the passive user share all the responsibilities except the control inputs. According to the random assignment, XXX (name of the participant), you will be the active user. And XXX (name of the participant), you will be the passive user."*
>
> *"You need to perform three tasks simultaneously with a computer program called MATB. One of the tasks will require you to respond to changes in lights and dials as quickly as possible. Another task will require you to use a joystick to control the position of a moving target. The other task*

will require you to control pumps to make sure the liquid levels in two tanks are at their optimal level. More detailed instructions about tasks will be given later. You will also need to go through a training session individually before the experiment trials."

"The experiment will be divided into three trials. You need to interact with different versions of MATB. After each trial, you need to fill out a survey individually to evaluate the process. Physiological data including palm sweating and heart activity will be gathered from you. So I need to put a few sensors on the surface to your skin. In addition, the experiment trials will be videotaped."

"The experiment will last about 90 minutes. You will receive $10 compensation for participating in this experiment. We also have a performance award for the groups with best performance. The award will be $20 for each of the group member."

Informed consents were then obtained from the two participants individually. Next, the participants took turns to go to a private space to put on the sensors for the physiological measurements. During the sensor placement, the following information was given to the participants:

"I am going to put 3 electrodes on your skin. They will be put right below your clavicle and ribs, for measuring the heart activity measurement. I will also wrap this device on your left palm to measure your palm sweating."

"In order to ensure appropriate measurements for the heart activity, you will need to minimize your movements of your upper body throughout the study. This is because the sensors are sensitive to movements which could affect the data collection. It is fine when you need to answer surveys and type,

since we are not recording during that time. But be careful during the tasks."

After the sensor placement, the participants were seated in separate computer workstations. A 6-minute baseline physiology recording was collected. Before this baseline data collection, the participants were instructed not to interact with each other. The instruction was as follows:

> *"To help with the physiological data analysis, a baseline recording needs to be done when you are not doing the computer tasks. So in the next six minutes, a baseline recording will be performed. In this period, please sit quietly and don't interact with each other. I will come back when the recording is finished."*

After the baseline recording, written instructions of the MATB were given to the participants to read. Copies of the written instructions are showed in Appendix A. The participants were given 6 minutes to read the instructions. After the reading, the experimenter came back and showed the interface of MATB on the computer screen. Oral instructions of the tasks were given to the participants. The content of the oral instructions was similar with that of the written instructions:

> *"In the monitoring task, you need to respond to the two lights and the four dials. The green light is normally on and the red light is normally off. If the green light turned off or the red light turned on, you need to press F5 or F6. For the four dials, the pointers will move up and down around the center of the scale within plus or minus one unit from the center. Once the pointer goes out of the plus and minus one range, you need to report to the system by pressing the corresponding keys. You just need to press the key once when you spot abnormalities. If you press the key multiple times, the presses other than the first one will be considered false alarms and those will reduce you*

performance score. Missing any signal will also reduce your performance score. And you need to respond as quickly as possible.”

“In the tracking task, this green target will move around randomly by itself. The task for you is to keep the target in the center of the window. You should move the joystick around to make sure the target moves within the dotted lines of the rectangle.”

“In the resource management task, you need to switch these pumps on and off. The goal is to make sure the fuel levels in Tank A and Tank B are as close to 2500 as possible. So you don't want the levels to be higher or lower than that. Tank C and Tank D are backup tanks so we don't concern about their fuel levels. To turn on the pumps, you just need to press the corresponding keys. Once a pump is turned on, it will keep working until you press the key again to shut it down.”

The active user then went through a 6-minute training session on a computer. The passive user did or did not go through the training depending on the expertise condition the team received. For the teams in high expertise condition, the passive users went through the same training as the active users. The participants were instructed not to interact with each other during the training. For the teams in low expertise condition, the passive users were instructed to move to a separate room and wait until the active user finish the training. After the training section, the experimenter answered the participants’ questions about the control and the goal of the tasks and received verbal confirmation that the participants understood the tasks.

The participants then worked together on one computer work station with one shared display for the experimental trials. There were three trials total, lasting a fixed length of 6 minutes each. The participants receive one task/technological condition in each trial.

Participants filled out surveys about their subjective experience after each task trial using separate computers. Scales used included the NASA task load index (NASA-TLX) (Hart & Staveland, 1988) and trust in technology scale (Jian et al., 2000). The participants moved to the next task trial once both of them finished the survey. Physiological data and video recording were collected during all the task trials. All the participants filled out a demographic information survey at the end of the experiment. There was also a question asking the participants to rate their familiarity with the other participant in the group prior to the experiment in a 1-5 scale. Lastly, the participants were debriefed about the study. The procedure is summarized in Figure 9.

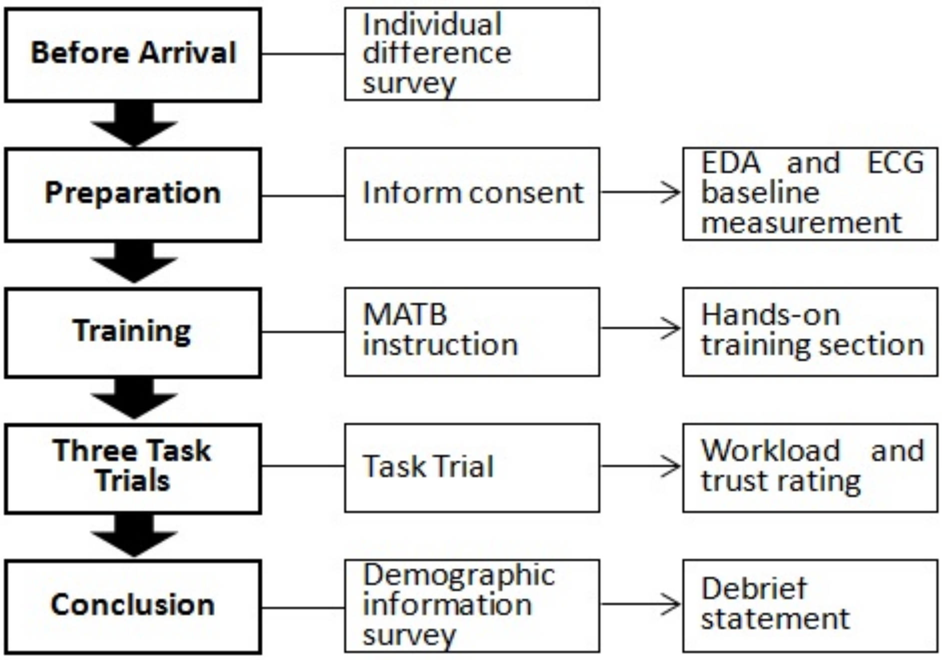


Figure 9. Procedure of the experiment of study one.

4.1.5. Measurements

4.2.5.1. Affective states measure derived from the videotaped data

The videos were captured at 30 frames per second with the resolution of 640*480. To reduce the video processing time, frame rate of the videos was reduced to 12 frame per second. The Computer Expression Recognition Toolbox (CERT) (G. C. Littlewort et al.,

2011) was used to analyze the facial features of the participants frame by frame. The resultant data that were included in the analysis were seven 12Hz time series indicating the intensity of six basic emotions (surprise, anger, sad, joy, disgust, and fear) and neutral expression. The range of the intensity output is 0 and 1, which represents the intensity of an expression as if it was estimated by expert human coders (iMotions, 2013). An example of the output of the analysis is shown in Figure 10.

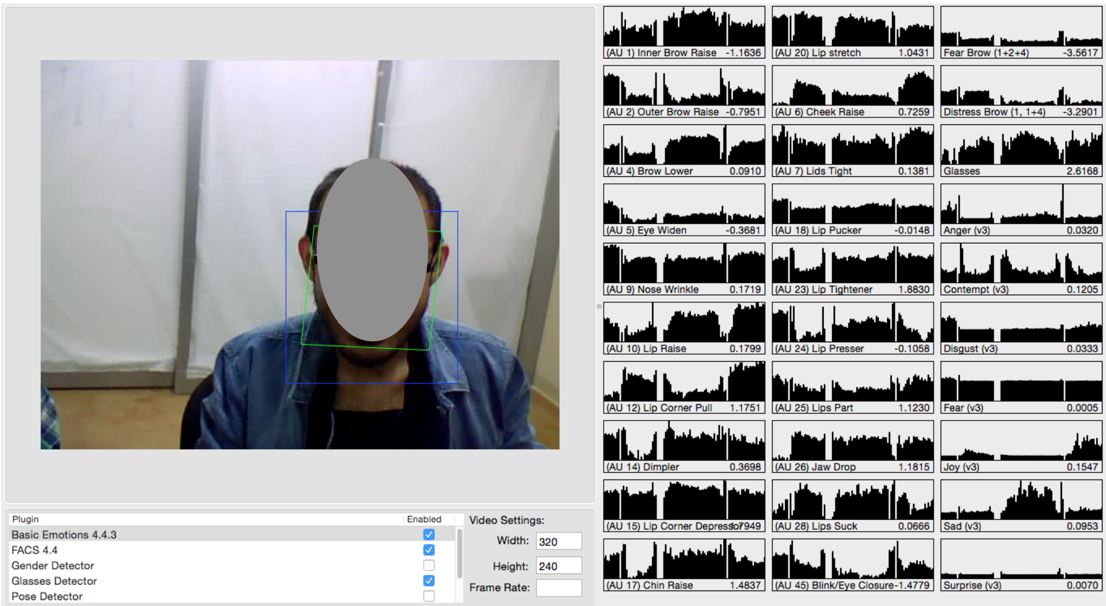


Figure 10. A screenshot of the interface of CERT. The plots are intensity estimations of the AUs and basic emotions. Participant's face was covered for privacy protection purpose.

The time series data was resampled for statistical analysis. Previous research used different time units for analysis. For example, research in marital conflict with the Specific Affect Coding System (SPAFF) used 1-second as time unit (Coan & Gottman, 2007). Levenson and Gottman (1983) averaged physiological signals into 10-second segments to investigate affective exchange in couples. However, unconscious facial mimicry can happen within 500 milliseconds (Dimberg, Thunberg, & Elmehed, 2000; Mancini, Ferrari, & Palagi, 2013). In this study, data was resampled at 0.1 Hz, 1Hz, and 2Hz (corresponding time units are 10s, 1s, and 0.5s) to examine if different time units can alter the results of the analyses significantly.

4.1.6. Data analysis

4.1.6.1.Missing data

4.1.6.1.1. Missing data overview

As there were 72 participants and each participant went through 3 experimental trials, there was a total number of 216 video recordings. The analyses from CERT resulted 216 time series from each emotion category (surprise, anger, sad, joy, disgust, and fear). Each time series contained 4,320 data points (360 seconds in length and 12 frame per second). Failing to detect a face in a frame would lead to missing values in all the emotion categories in that frame. As a result, the patterns of missing data were identical for all the emotion categories.

The percentage of missing values in each emotion category was 6.19% (57,750 out of 933,120 data points). The average percentage of missing values in each video recording is 6.19%, with a standard deviation of 9.04%. 17 out of 216 recordings had missing value percentages higher than 20%. The main cause of missing values was face detection failure when the participants looked down on the keyboard or looked at their teammate.

4.1.6.1.2. Multiple imputation

R (R Core Team, 2014) package Amelia II (J. Honaker, King, & Blackwell, 2011) was used to perform multiple imputation to create 5 imputed complete datasets. Methodologists considered that multiple imputation is the preferred technique as it provides advantages in accuracy and statistical power (van Buuren, 2013). Amelia II imputes a data point with the information both from the estimation of linear combination of other variables and polynomials of the time series (James Honaker & King, 2010). By using polynomials of time, the imputation algorithm corresponds to the assumption that time-series variables' value changes relatively smoothly over time. The imputations in this study were performed using the linear combination of the following variables: participant ID, team ID, experimental trial

ID, expertise (high/low), task/technological condition (normal/hard/low reliability), role (active user/passive user), and sequence of experimental trial (1/2/3). Second-order polynomials of time was also fitted to each time series in the imputations, as aforementioned.

The number of imputed data sets (m) to create is an important attribute for multiple imputation procedure. According to the rule of thumb proposed by Von Hippel (2009), the m value in multiple imputation should be similar to the percentage of missing values. The default number of m is 5 in Amelia as this number is considered adequate unless the percentage of missing is very high (J. Honaker et al., 2011; Rubin, 1987; Schafer, 1997). Van Buuren (2013) also noted that “substantive conclusions are unlikely to change as a result of raising m beyond m=5”. Thus m=5 was used in this study. The subsequent analyses were applied to each of the 5 imputed datasets.

To test the statistical significance, first, estimates and standard errors were combined using Rubin’s rules (Rubin, 1987). Second, the degrees of freedom estimation of the original Rubin’s Rules was replaced by an adjustment proposed by Barnard and Rubin (1999), as the original rules tend to overestimate the degrees of freedom (van Ginkel & Kroonenberg, 2014).

4.1.6.1.3. Imputation diagnostics

Distribution density. A good imputation is likely to have a distribution similar to the observed data (van Buuren, 2013). Kernel density estimates were created to compare the observed data with the imputed data. The distribution was inspected for discrepancies.

Autocorrelation. The autocorrelations of the observed data and imputed data were also compared to investigate if there were significant discrepancies. The autocorrelations were calculated for each time series for each variable both for the observed data and the 5 imputed datasets. The autocorrelations computed from each imputed dataset were compared against the observed data in pairwise t-tests.

4.1.6.2. Hypothesis 1.1: the affective states of the individuals worked in a same team are more similar than that of the individuals did not work in a same team

4.1.6.2.1. Similarity indexes and dynamic time warping (DTW)

To measure the similarity (or dissimilarity) of the time series, Euclidean distance was used. Euclidean distance was defined as the square root of the averaged sum of squared differences between two variables (Kenny, Kashy, & Cook, 2006).

In the calculation of the similarity in the emotion expressions between two people, however, direct application of Euclidean distance is not always appropriate due to the contagious nature of emotion. For example, if individual A's joy influences individual B, there would be a time difference (or time lag) between the joy expressions of the two individuals. Furthermore, in certain time period, there may be times when individual A influences individual B, and there may be times when individual B influences individual A. The result of emotional contagion is that the time lag of the two time series is not a constant. This requires aligning the time series flexibly. Traditional alignment methods aligns time series with a constant time lag thus they are not appropriate in this study. A technique known as dynamic time warping (DTW) was used for the alignment.

The DTW technique was originally used to compare speech patterns in automatic speech recognition (Müller, 2007), where the matching time series are similar in general shapes but may have different lengths or require flexible alignments. As this technique is flexible that one can specify constraints on how time series are aligned, it was then widely used in other fields such as electro-cardiogram analysis, biometrics, and process monitoring (Berndt & Clifford, 1994; Giorgino, 2009). Figure 11 shows an example of how two time series data were aligned. The alignment can also be illustrated in a warping path plot (see Figure 12). In DTW, the warping path for time series should satisfy three conditions (Giorgino, 2009;

Müller, 2007): (1) the path begins at point $P_1 = (1, 1)$ and ends at point $P_L = (N, M)$, given the length of the time series is N and M (in this study, N=M); (2) the path won't go backwards that $n_1 \leq n_2 \leq \ldots \leq n_L$ and $m_1 \leq m_2 \leq \ldots \leq m_L$; (3) the size of each step is limited to $p_{k+1} - p_k \in (1, 0), (0,1), (1,1)\}$. In this study, in addition to these three conditions, a global constraint known as Sakoe-Chiba band (Sakoe & Chiba, 1978) with width 1 was applied to the warping path (see Figure 13). With this constraint, the warping path would only go through the diagonal region where the time difference of the two aligned data points are smaller or equal to 1 unit. In other words, individual A's expression value at time t may be aligned with individual B's expression value at time t-1, t, or t+1, but not anything else.

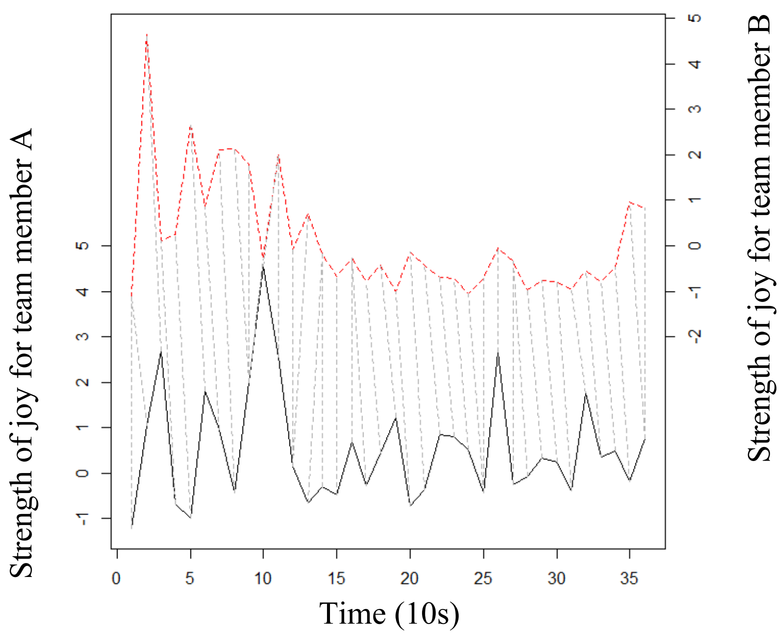


Figure 11. An example of time alignment of two time series data. The grey dashed lines indicate how the two time series were aligned to calculate Euclidean distance.

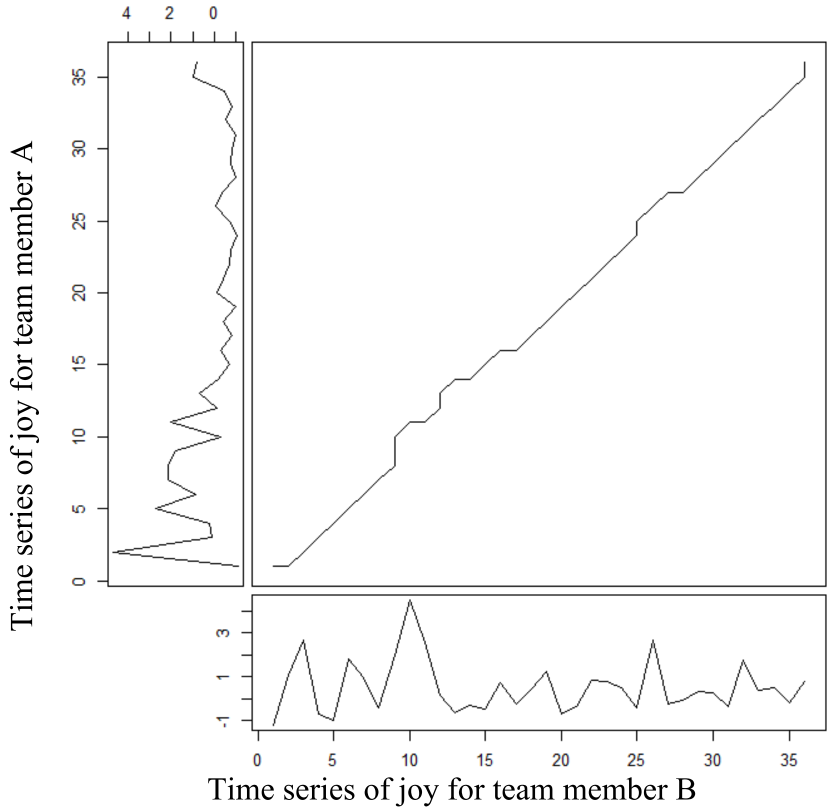


Figure 12. The warping path of the time series used in Figure 11.

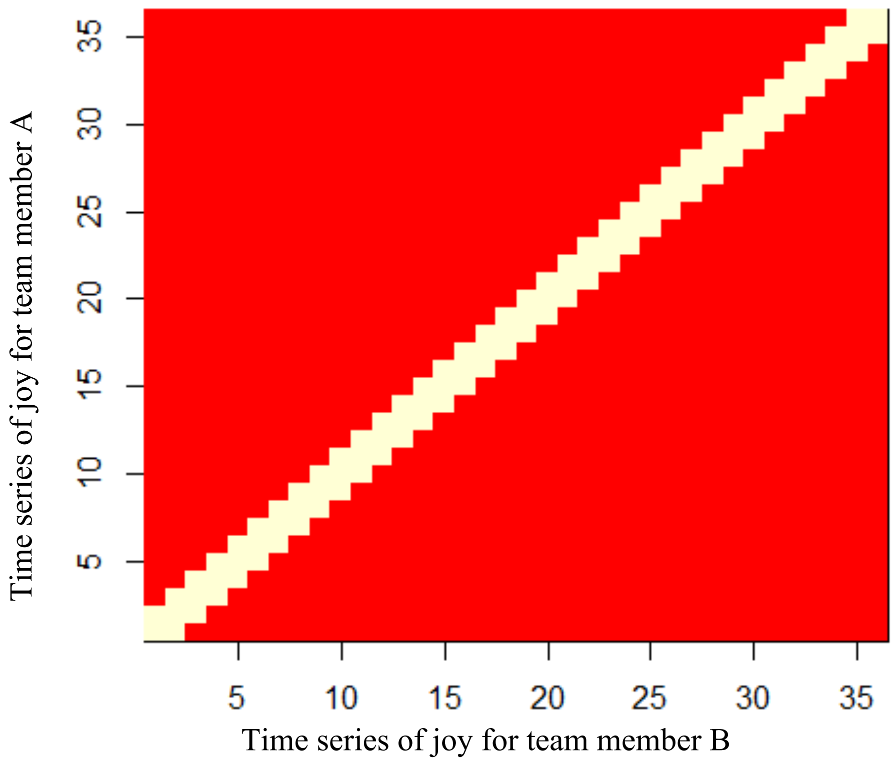


Figure 13. Global constraint (Sakoe-Chiba band with width 1) for the warping path for the time series illustrated in Figure 11 and Figure 12. The yellow area is the area that is admissible for the warping path, and the red area is the area that is not admissible for the warping path.

Setting the 1-unit global constraint on the warping path reflected the assumption that there is a time boundary to emotional contagion. For 0.1 Hz sample rate (1 data point per 10 seconds), the 1-unit global constraint meant the effect of emotional contagion only applies to adjacent data points which are 10 seconds apart. Affective values that were further than 10 seconds apart in time would not be considered in the calculation of the similarity index. The application of 1-unit global constraint to the sample rate of 1Hz and 2Hz may be too strict. However, this made the algorithm consistent across different sample rates so the results could be comparable.

The computations related to DTW were performed in R (R Core Team, 2014) with dtw package (Giorgino, 2009).

4.1.6.2.2. Pseudo-couple analysis

To compare the individuals who worked in the same team with the individuals who did not, a technique known as "pseudo-couple analysis" was used in the data analysis (Corsini, 1956; Kenny & Acitelli, 1994). First, similarity indexes of affective states were calculated for the participants who worked in a same team. Second, pseudo-teams were generated by pairing active users and passive users who did not worked in a same team. To make sure the actual teams and the pseudo-teams were comparable, the pseudo teams were formed with couples of participants who received the exact same experimental treatments. Third, the similarity indexes of affective states were calculated for the pseudo-teams. Finally, the similarity indexes from the real teams and the pseudo teams were compared. Figure 14 depicts the idea of pseudo-couple analysis with two real teams and two pseudo teams.

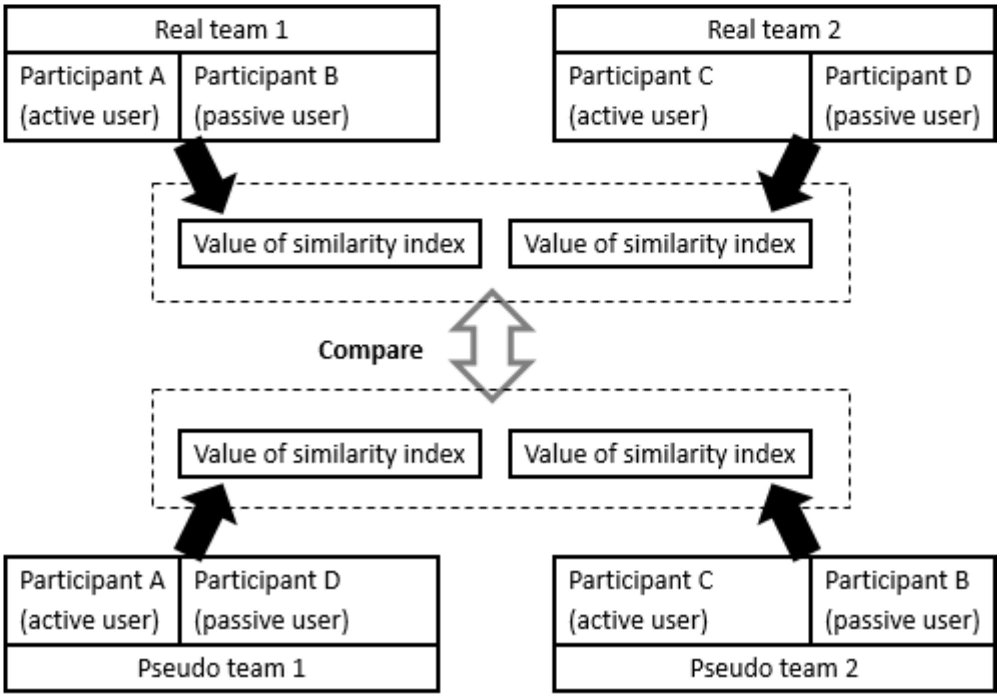


Figure 14. The general idea of the pseudo-couple analysis.

To test if there was higher similarity in the real teams than that in the pseudo teams, the following method was used: first, real/pseudo team was coded as a dummy variable in the data. Pseudo team was coded as 0 and real group was coded as 1. Second, a linear mixed effects (LME) model was fitted to the data using values of similarity index as dependent variable; real/pseudo team, task/technological conditions (normal/hard/low reliability), expertise (high/low), and their interactions as independent variables; and teams as random intercepts. The model specification is shown in Table 2. Third, if the coefficient of real/pseudo team was negative and significant, then the hypothesis that real teams have higher values of similarity indexes would be supported.

Table 2: LME model specification for testing hypothesis 1.1 in study one.

Dependent variable	Fixed effect variables	Random effect variables
1. Similarity index for positive affect or negative affect	1. Real/pseudo team 2. Task/technological conditions 3. Expertise 4. Interaction terms (all possible two-way interactions)	1. Teams

The LME model fittings were conducted using R (R Core Team, 2014) with the lme4 package (Bates, Maechler, & Bolker, 2013). However, the lme4 package omits the output of p values associated with sequential ANOVA decompositions of fixed effects (Bates, Mächler, Bolker, & Walker, 2014). This is because the null distributions of the parameter estimates in complex designs are not t distributed for finite size samples. The authors did propose recommendations for varies methods to approximate p values, however, when this computation is necessary (Bates et al., 2014). In this study, Satterthwaite approximation for degrees of freedom was used for obtaining p values when a single imputed dataset was tested, using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2012). In the results, the combined results from all 5 imputed datasets using Rubin's rules (Rubin, 1987; see section 4.1.6.1.2) were reported.

4.1.6.3. Hypothesis 1.2: an individual's current affective state is positively related to another individual's previous affective state

To test hypothesis 1.2, an analysis framework known as the actor-partner interdependence model (APIM) (Cook & Kenny, 2005; Kenny, 1996; Kenny et al., 2006) was used. A visualization of the model is shown in Figure 15.

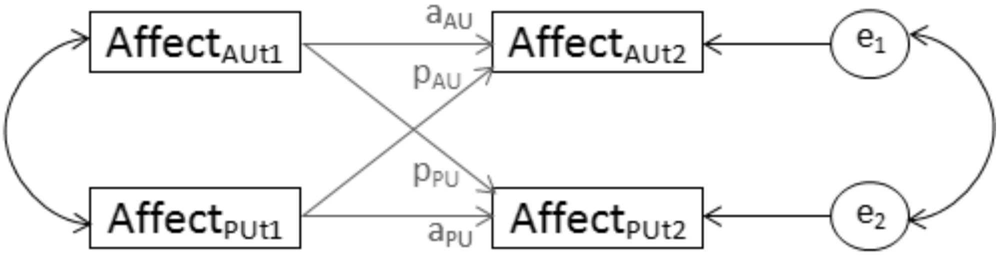


Figure 15. The APIM framework. $Affect_{AUt1}$ is the integral affect measure of active user in time 1. $Affect_{AUt2}$ is the integral affect measure of active user in time 2. $Affect_{PUt1}$ is the integral affect measure of passive user in time 1. $Affect_{PUt2}$ is the integral affect measure of passive user in time 2. a_{AU} and a_{PU} is the actor effect for active user and passive user. p_{AU} and p_{PU} is the partner effect for active user and passive user.

In the model, $Affect_{AUt1}$ and $Affect_{AUt2}$ are the integral affect measures of active user in time 1 and time 2. $Affect_{PUt1}$ and $Affect_{PUt2}$ are the integral affect measures of passive user in time 1 and time 2. The effect of $Affect_{AUt1}$ on $Affect_{AUt2}$ and the effect of $Affect_{PUt1}$ on $Affect_{PUt2}$ (labeled a_{AU} and a_{PU} in the figure) are called the actor effects. The actor effect describes how an individual's own affective state changes over time. The effect of $Affect_{AUt1}$ on $Affect_{PUt2}$ and the effect of $Affect_{AUt1}$ on $Affect_{PUt2}$ (labeled p_{AU} and p_{PU} in the figure) are called the partner effects. The partner effect describes how an individual's affective influence another individual's affective state. Thus, to test hypothesis 1.2 was to test the partner effects.

The APIM framework were modeled using LME model. The dependent variable was the positive or negative affective state value for either the active user or the passive user at time t in a task trial. The value of t should be t>1, thus the first data point for each task trial was omitted in the dependent variable. The independent variables included: affective value for the "actor" (whose data was used in the dependent variable) in time t-1, affective value for the "partner" (the other individual in the same team) in t-1, task/technological conditions (normal/hard/low reliability), expertise (high/low), role of the participants (active/passive user), and the interaction terms. The random variables included intercepts of task trials, individuals, and teams. The model specification is presented in Table 3.

Table 3: LME model specification for testing hypothesis 1.2, 1.3 and 1.4 in study one.

Dependent variable	Fixed effect variables	Random effect variables
1. Positive or negative affective state value for the actor in time t	1. Affective value for the actor in time t-1 2. Affective value for the partner in t-1 3. Task/technological conditions 4. Expertise 5. Role 6. Interaction terms (all possible three-way interactions)	1. Task trials 2. Individuals 3. Teams

The coefficient of the affective value for the actor (partner) in time t-1 represented the average actor (partner) effect. So the partner effect could be directly tested in this model.

4.1.6.4.Hypothesis 1.3: in terms of affective states, the active user's influence over the passive user is stronger than the passive user's influence over the active user

This hypothesis stated that the partner effect is different for the active user and passive user. To test this hypothesis, the same model described in Table 3 was fitted to the data. The coefficient for the affective value for the partner in t-1 X role interaction was used for testing this hypothesis.

4.1.6.5. Hypothesis 1.4: in terms of affective states, the active user's influence over the passive user will be stronger under the low expertise condition than the high expertise condition

The same model described in Table 3 were used for testing this hypothesis. This hypothesis stated that the effect of the partner effect (when the actor is the passive user) is different in different levels of expertise. Thus, this hypothesis could be tested by the simple effect of partner in t-1 X expertise interaction when the role of the actor was held as passive user.

4.2. Results

4.2.1. Data processing and multiple imputation

The emotional states analyses using CERT were summarized in Table 4. As fear and surprise consisted of very low percentages of total emotional state values expressed by the participants (less than 4%), they were not included in subsequent analyses. The sum of anger, disgust, and sad was used as the measure of negative affective state, and joy was used as the measure of positive affective state.

Table 4: Averages, standard deviations, and percentages out of the sum of emotional states values for anger, disgust, fear, joy, sad, and surprise.

Emotional states	**Average**	**Standard deviation**	**Percentage out of the sum of emotional states values**
Anger	0.055	0.100	28.803%
Disgust	0.030	0.064	15.107%
Fear	0.003	0.013	1.485%
Joy	0.035	0.124	15.113%
Sad	0.069	0.116	35.669%
Surprise	0.005	0.017	3.822%

The multiple imputation procedure was performed for positive affective state values and negative state values. Since the distribution of the values of the positive affective state was highly skewed, log transformation was performed before the imputation. After the imputation, the log values were transformed back to original units. Figure 16 shows the comparisons of density plots for the positive affective state and negative state for the original dataset and the imputed values for the 5 imputations.

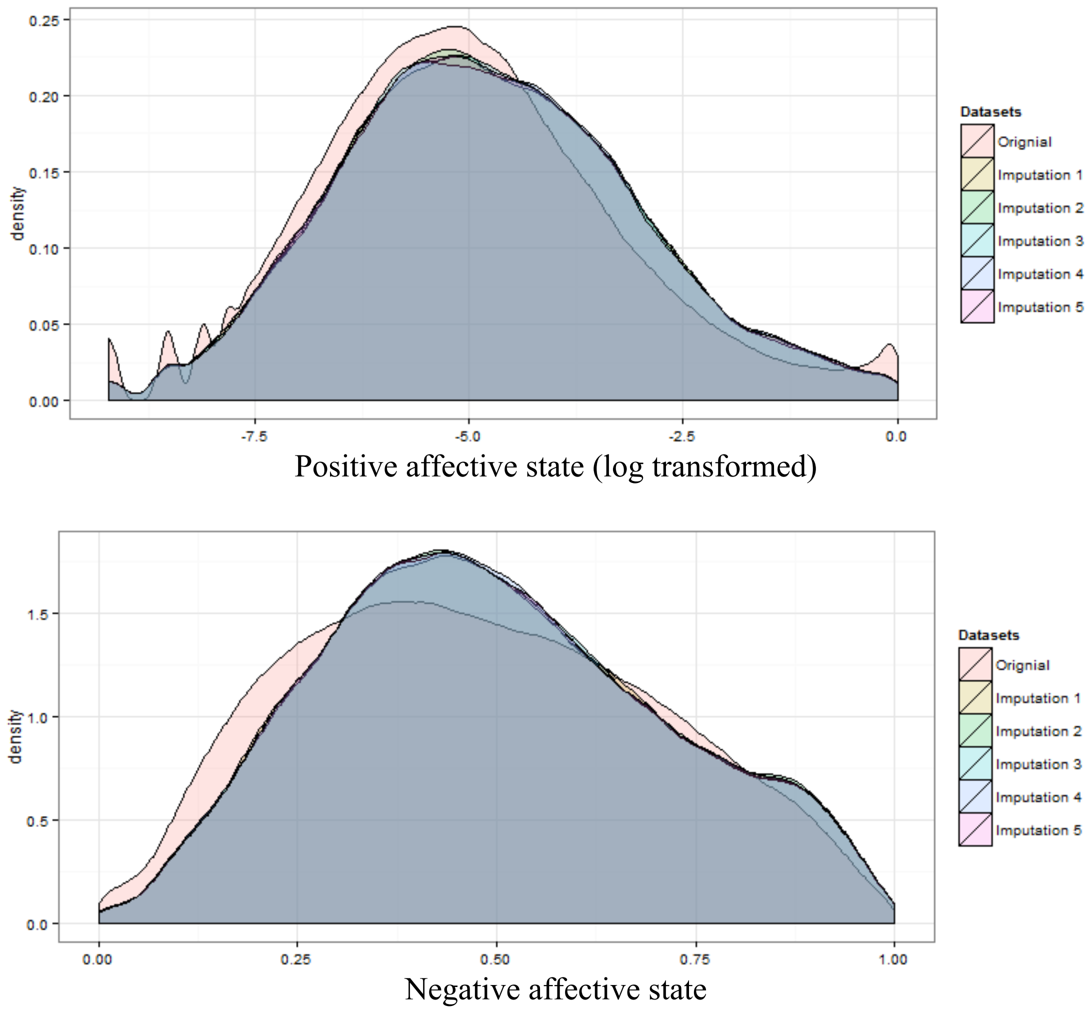
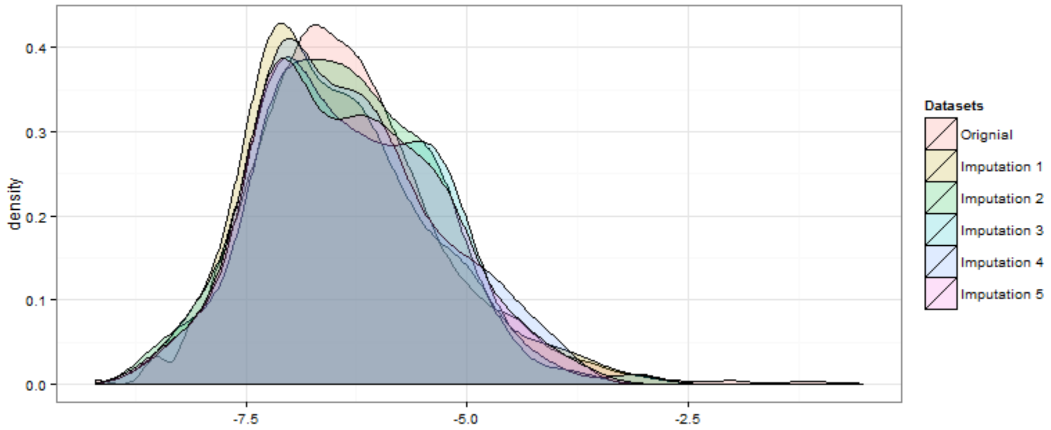
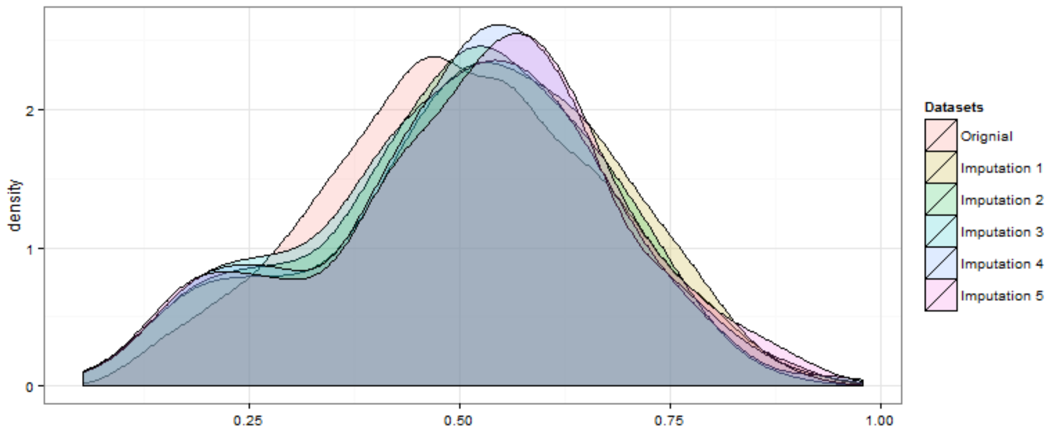


Figure 16. Kernel density estimates for the marginal distributions of the observed data and the m = 5 densities per variable calculated from the imputed data.

Figure 17 showed an example of the plots for one of the selected trials of one of the participants. Similar plots were also created for 10 randomly selected experimental trials. The patterns of similarity/discrepancy of the densities were similar.



Positive affective state (log transformed)



Negative affective state

Figure 17. Kernel density estimates for the marginal distributions of the observed data (blue) and the m = 5 densities per variable calculated from the imputed data (thin red lines) for a randomly selected experimental trial.

Lag 1 through lag 3 autocorrelations were calculated for the observed dataset and the imputed datasets (see Table 5) to exam if imputations significantly influenced the autocorrelations of the time series data. LME models were fitted to the date to test if the differences in the means of the autocorrelations were significant. In the model, observed/imputed data was included as a fixed effect, and individuals and teams were included as random effects. The results indicated that the imputed datasets had significantly lower lag 1 autocorrelations for negative affective state. No significant difference was observed for the other values.

Table 5: Average lag 1 to lag 3 autocorrelations for the observed data and the 5 imputed datasets for each variable imputed. The t values were the results LME models to test if the imputed data is significantly different from the observed data.

Variable	Observed	Imputed 1	Imputed 2	Imputed 3	Imputed 4	Imputed 5
Positive affect (lag 1)	0.83	0.82 (t=-1.11)	0.82 (t=-0.98)	0.82 (t=-0.96)	0.82 (t=-0.90)	0.82 (t=-1.03)
Positive affect (lag 2)	0.75	0.75 (t=-0.05)	0.75 (t=0.04)	0.75 (t=0.01)	0.75 (t=0.06)	0.75 (t=0.00)
Positive affect (lag 3)	0.70	0.70 (t=-0.03)	0.70 (t=-0.02)	0.70 (t=0.04)	0.70 (t=0.02)	0.70 (t=0.00)
Negative affect (lag 1)	0.79	0.77 (t=-2.44*)	0.77 (t=-2.42*)	0.77 (-2.31*)	0.78 (t=-2.28*)	0.77 (t-2.35*)
Negative affect (lag 2)	0.69	0.69 (t=-0.11)	0.69 (t=-0.06)	0.69 (t=-0.01)	0.69 (t=-0.02)	0.69 (t=0.00)
Negative affect (lag 3)	0.64	0.63 (t=-0.12)	0.63 (t=-0.14)	0.63 (t=-0.11)	0.64 (t=0.00)	0.64 (t=0.00)

After the imputation, the time series in each imputed dataset were down-sampled from 12Hz (corresponding time units was 0.08s) to 0.1Hz, 1Hz, and 2Hz (corresponding time units were 10s, 1s, and 0.5s). The down-sampling was performed using the signal package (Ligges, 2015) in R. After the down-sampling, the autocorrelations for lag 1 were tested again (see Table 6). No significant difference between the original dataset and the imputed datasets was observed.

Table 6: Average lag 1 autocorrelations for the observed data and the 5 imputed datasets for each down-sampled frequencies. The t values were the results LME models to test if the imputed data is significantly different from the observed data.

Variable	Observed	Imputed 1	Imputed 2	Imputed 3	Imputed 4	Imputed 5
Positive affect (0.1Hz)	0.09	0.09 (t=0.24)	0.09 (t=0.45)	0.09 (t=0.12)	0.09 (t=-0.19)	0.08 (t=0.00)
Positive affect (1Hz)	0.58	0.57 (t=-0.37)	0.57 (t=-0.48)	0.57 (t=-0.39)	0.57 (t=-0.44)	0.57 (t=-0.51)
Positive affect (2Hz)	0.75	0.74 (t=-0.57)	0.74 (t=-0.50)	0.74 (t=-0.49)	0.74 (t=-0.53)	0.74 (t=-0.60)
Negative affect (0.1Hz)	0.13	0.13 (t=0.21)	0.13 (t=0.30)	0.13 (t=0.19)	0.13 (t=0.10)	0.13 (t=0.00)
Negative affect (1Hz)	0.56	0.55 (t=-0.53)	0.55 (t=-0.54)	0.55 (t=-0.57)	0.55 (t=-0.53)	0.55 (t=-0.59)
Negative affect (2Hz)	0.73	0.72 (t=-1.26)	0.72 (t=-1.21)	0.72 (t=-1.25)	0.72 (t=-1.25)	0.72 (t=-1.25)

4.2.2. Pseudo-couple analysis

Similarity indexes were calculated for both actual teams and pseudo teams. LME model were fitted to the five imputed datasets according the specification in Table 2. In the model, POC coding was used for task/technological conditions and expertise level. So the test for actual/pseudo team can be interpreted as the differences between the actual teams and pseudo teams when the other variables were held at their mean.

4.2.2.1.Hypothesis 1.1: the affective states of the individuals worked in a same team are more similar than that of the individuals did not work in a same team

The results were summarized in Table 7. The average distance of the positive affect between the two individuals working in actual teams were lower than those of the pseudo teams, indicated that individuals worked in a same team had similar positive affect expressions. The pooled significance tests using Rubin's rules showed the differences were significant. For the negative affect, although the average distances for the actual teams were lower than the pseudo teams, none of the tests showed that the differences were statistically significant.

Table 7: The DTW distances for actual teams and pseudo teams for positive affective state and negative state for 0.1Hz, 1Hz, and 2 Hz sample rate. These values were computed from the first imputation dataset.

	Time series sample rate	**DTW distance for actual teams**	**DTW distance for pseudo teams**	**Multiple imputation test**
Positive affect	0.1 Hz	20.10 (SD=10.58)	23.90 (SD=9.83)	b=0.38, t(318)=3.18, p<0.05
	1 Hz	188.39 (SD=83.30)	223.33 (SD=84.24)	b=0.42, t(320)=3.35, p<0.05
	2 Hz	375.81 (SD=167.25)	446.31 (SD=168.84)	b=0.41, t(320)=3.30, p<0.05
Negative affect	0.1 Hz	30.41 (SD=6.24)	30.50 (SD=7.15)	b=0.01, t(320)=0.06, p=0.95
	1 Hz	354.05 (SD=63.18)	359.27 (SD=70.21)	b=0.08, t(320)=0.63, p=0.53
	2 Hz	731.09 (SD=124.49)	741.21 (SD=135.93)	b=0.08, t(320)=0.62, p=0.62

Note. b values were mean standardized coefficients across 5 imputed datasets. p values for the multiple imputation were calculated using Rubin's rule. Grey cell indicates that the statistical test was significant at α=0.05 level.

In summary, the hypothesis was supported for positive affective state but not for negative affective state.

4.2.3. Actor-partner interdependence model analysis

The imputed datasets were transformed and analyzed using the APIM model described in section 4.1.6.3. LME models as specified in Table 3 were fitted to the data and statistical tests were performed and results were pooled across imputed datasets. The partner effect X role X expertise three-way interaction effect was tested first. The results suggested that the interaction effects were not significant for negative affect. For positive affect, the effect was significant if the sample rate was 1Hz and 2Hz. The results were summarized in

Table 8. Figure 18 and Figure 19 are visualizations of the three-way interaction effects for positive affect on 1Hz and 2Hz sample rates respectively. Note that in the figures, the steeper the line is, the stronger the influence of actor on the partner is. This indicated that, for positive affect (when sample rate was 1Hz and 2Hz), active user's influence over the passive user were stronger than the other way around under the high expertise condition.

Table 8: The partner effect X role X expertise interaction for 0.1Hz, 1Hz, and 2 Hz sample rate.

	Sample rate	**Multiple imputation test**
Positive affect	0.1 Hz	$b=0.08$, $t(1361)=1.52$, $p=0.13$
	1 Hz	$b=0.05$, $t(10263)=3.75$, $p<0.05$
	2 Hz	$b=0.02$, $t(79092)=3.05$, $p<0.05$
Negative affect	0.1 Hz	$b=0.04$, $t(4845)=0.77$, $p=0.44$
	1 Hz	$b=0.002$, $t(23572)=0.17$, $p=0.87$
	2 Hz	$b=0.001$, $t(10017)=0.08$, $p=0.93$

Note. b values were mean standardized coefficients across 5 imputed datasets. p values for the multiple imputation were calculated using Rubin's rule. Grey cell indicates that the statistical test was significant at $\alpha=0.05$ level.

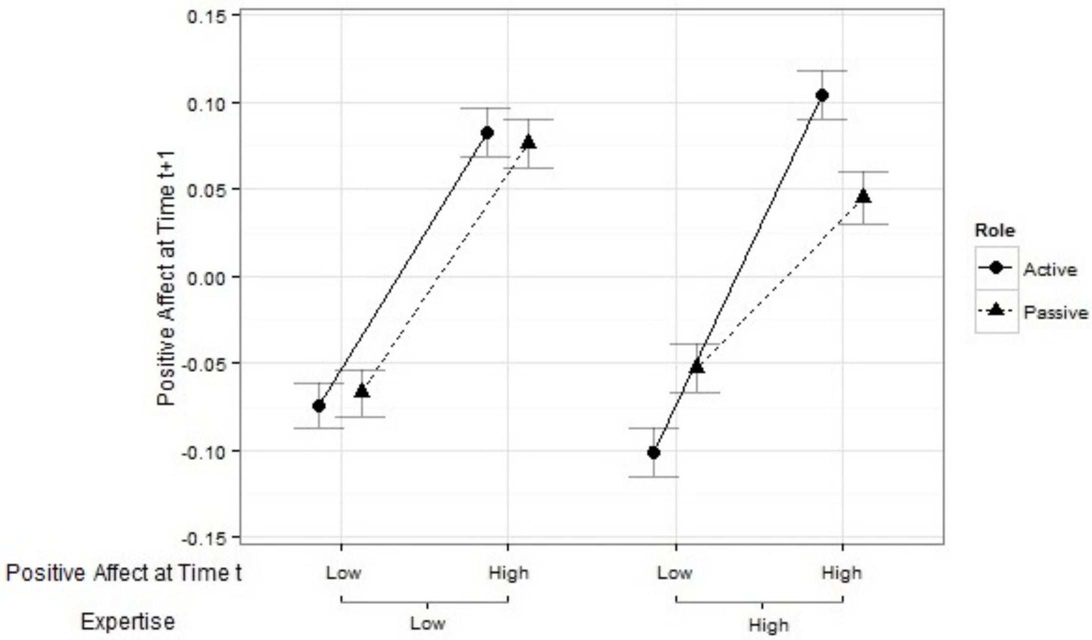


Figure 18. The predicted means and standard errors of actor's positive affect at time t+1 given partner's positive affect levels at time t (high was defined as the value of +1 SD and low was defined as the value of -1 SD), partner's role, and expertise level. The sample rate is 1Hz. All the values were computed when the values of all the other variables held at their means.

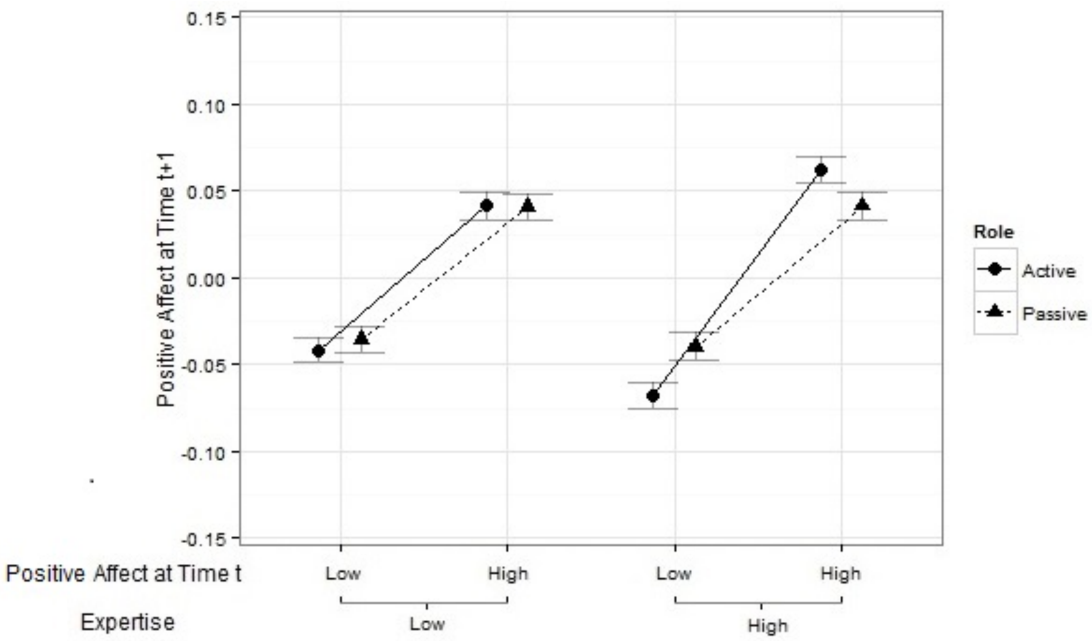


Figure 19. The predicted means and standard errors of actor's positive affect at time t+1 given partner's positive affect levels at time t (high was defined as the value of +1 SD and low was defined as the value of -1 SD), partner's role, and expertise level. The sample rate is 2Hz. All the values were computed when the values of all the other variables held at their means.

4.2.3.1.Hypothesis 1.3: in terms of affective states, the active user's influence over the passive user is stronger than the passive user's influence over the active user

Hypothesis 1.3 and 1.4 were tested before hypothesis 1.2 as they involved the testing of two-way interactions.

For positive affect sampled at 1Hz and 2Hz, since the partner effect X role X expertise three-way interaction effect was significant, two methods were used to examine the effect of the partner effect X role interaction. First, this effect was tested as the simple effect when all the other variables held at their mean values. Second, simple effects were tested given different expertise levels (low and high).

When all other variables were held at their mean, the results showed that the interaction effect was significant (for 1Hz sample rate, b=0.03, t(2207)=3.87, p<0.05; for 2Hz sample rate, b=0.02, t(17007)=3.21, p<0.05); see Figure 20 for visualizations. This indicated that on average, active user's influence over the passive user on positive affect was stronger than the passive user's influence over the active user. The test results for the simple effects under varying expertise level were summarized in Table 9. These results showed that the interaction affect was not significant if the passive user was not trained.

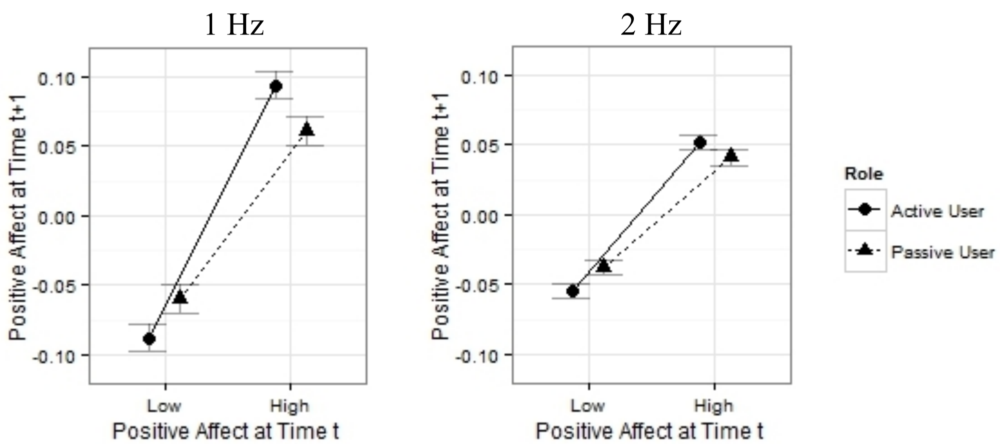


Figure 20. The predicted means and standard errors of actor's positive affect at time t+1 given partner's positive affect levels at time t (high was defined as the value of +1 SD and low was defined as the value of -1 SD) and partner's role. All the values were computed when the values of all the other variables held at their means.

Table 9: The simple effects of the partner X role interaction effect for 1Hz, and 2 Hz sample rate given different expertise levels (low and high).

	Simple effect	Sample rate	Multiple imputation test
Positive affect	Low expertise	1 Hz	b=0.01, t(3686)=0.82, p=0.42
		2 Hz	b=0.004, t(10755)=0.70, p=0.48
	High expertise	1 Hz	b=0.06, t(3273)=5.19, p<0.05
		2 Hz	b=0.03, t(96720)=4.31, p<0.05

Note. b values were mean standardized coefficients across 5 imputed datasets. p values for the multiple imputation were calculated using Rubin's rule. Grey cell indicates that the statistical test was significant at $\alpha=0.05$ level.

For other cases where the three-way interaction was not significant, including positive affect sampled at 0.1Hz and negative affect sampled at all sample rates, the effect of the partner effect X role interaction was tested when all the other variables held at their mean. No significant result was found in these tests (for positive affect at 0.1Hz sample rate, b=-0.04, t(1271)=-1.18, p=0.24; for negative affect at 0.1Hz sample rate, b=0.04, t(6693)=1.69, p=0.09; for negative affect at 1Hz sample rate, b=-0.01, t(36174)=-0.96, p=0.34; for negative affect at 2Hz sample rate, b=-0.01, t(17419)=-1.57, p=0.12).

In summary, hypothesis 1.3 was not supported for negative affect. It was partially supported for positive affect under the sample rate of 1Hz and 2Hz. The tests of simple effect suggested if the passive user had high expertise, active user's influence over passive user was stronger than passive user's influence over active user; but this was not true if the passive user had low expertise.

4.2.3.2. Hypothesis 1.4: in terms of affective states, the active user's influence over the passive user will be stronger under the low expertise condition than the high expertise condition

To test this hypothesis, the simple effect of the partner effect X expertise interaction was tested when the partner's role was held as passive user.

For positive affect, the results showed that the simple interaction effects were significant for 1Hz and 2Hz sample rates (for 0.1Hz sample rate, b=0.01, t(2942)=0.20, p=0.84; for 1Hz sample rate, b=0.02, t(3985)=2.60, p<0.05; for 2Hz sample rate, b=0.02, t(11002)=4.38, p<0.05); see Figure 21 for visualizations.

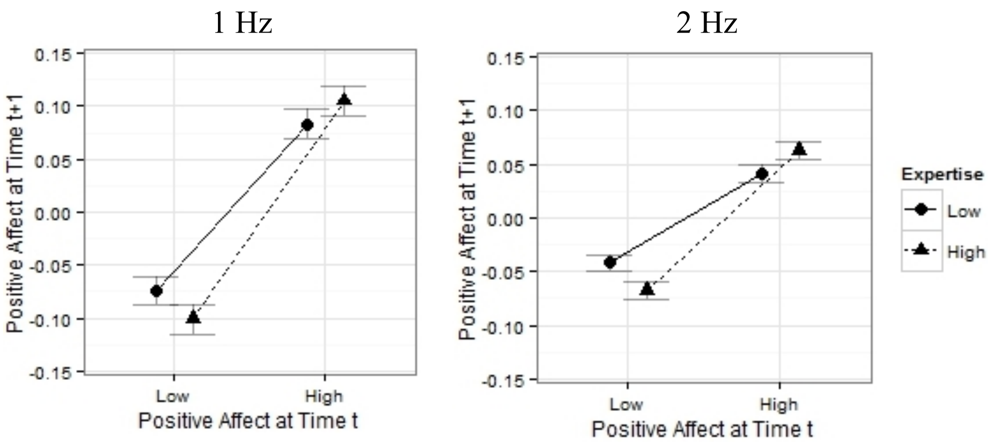


Figure 21. The predicted means and standard errors of actor's positive affect at time t+1 given partner's positive affect levels at time t (high was defined as the value of +1 SD and low was defined as the value of -1 SD) and expertise conditions, given that the actor's role was passive user. All the values were computed when the values of all the other variables held at their means.

For negative effect, the simple interaction effects were not significant across all the sample rates (for 0.1Hz sample rate, b=-0.01, t(2541)=-0.52, p=0.6; for 1Hz sample rate, b=-0.01, t(66852)=-0.74, p=0.45; for 2Hz sample rate, b=-0.001, t(33893)=-0.15, p=0.87).

In summary, hypothesis 1.4 was not supported for negative affect. It was supported for positive affect under the sample rate of 1Hz and 2Hz.

4.2.3.3.Hypothesis 1.2: an individual's current affective state is positively related to another individual's previous affective state.

For positive affect sampled at 1Hz and 2Hz, since both the partner effect X role X expertise three-way interaction effect the partner effect X role two-way interaction effect were significant, two methods were used to examine the main effect of the partner effect. First, this effect was tested as the simple effect when all the other variables held at their mean values. Second, simple effects were tested given all possible combinations of partner's role (active user and passive user) and expertise levels (low and high).

When all other variables were held at their mean, the results showed that the partner effect was significant (for 1Hz sample rate, $b=0.08$, $t(13866)=18.13$, $p<0.05$; for 2Hz sample rate, $b=0.05$, $t(49508)=19.91$, $p<0.05$). The test results for the simple effects under varying partner's role and expertise level were summarized in Table 10. This results indicated that the partner effect was significant across all conditions for positive affect sampled at 1Hz and 2Hz.

Table 10: The simple effects of the partner effect for 1Hz, and 2 Hz sample rate given different combinations of partner's role (active user and passive user) and expertise (low and high).

	Simple effect	**Sample rate**	**Multiple imputation test**
Positive affect	Active user, Low expertise	1 Hz	b=0.07, t(21111)=10.4, p<0.05
		2 Hz	b=0.04, t(84545)=9.87, p<0.05
	Active user, High expertise	1 Hz	b=0.1, t(2946)=14.35, p<0.05
		2 Hz	b=0.06, t(6782)=16.09, p<0.05
	Passive user, Low expertise	1 Hz	b=0.06, t(3131)=9.25, p<0.05
		2 Hz	b=0.03, t(4741)=8.74, p<0.05
	Passive user, High expertise	1 Hz	b=0.04, t(15614)=5.67, p<0.05
		2 Hz	b=0.06, t(6782)=16.09, p<0.05

Note. b values were mean standardized coefficients across 5 imputed datasets. p values for the multiple imputation were calculated using Rubin's rule. Grey cell indicates that the statistical test was significant at $\alpha=0.05$ level.

For other cases where the interaction effects were not significant, including positive affect sampled at 0.1Hz and negative affect sampled at all sample rates, the effect of the partner effect was tested when all the other variables held at their mean.

The results showed that the partner effect was not significant if the sample rate was 0.1Hz (for positive affect, $b=0.02$, $t(1855)=1.41$, $p=0.16$; for negative affect, $b=-0.02$, $t(4568)=-1.45$, $p=0.15$). For negative affect, the partner effect was significant for both 1Hz and 2Hz sample rates (for 1Hz sample rate, $b=-0.01$, $t(1892)=-4.53$, $p<0.05$; for 2Hz sample rate, $b=-0.01$, $t(2860)=-4.03$, $p<0.05$). Interestingly, the coefficient of the partner effect was negative for negative affect.

In summary, hypothesis 1.2 was supported for positive affect under the sample rate of

1Hz and 2Hz. However, partner's negative affect at time t-1 was negatively related to actor's negative affect at time t under the sample rate of 1Hz and 2Hz, which was to the contrary of the hypothesis. In addition, the hypothesis was not supported under the sample rate of 0.1Hz.

4.2.4. Additional analysis

4.2.4.1. Integral affect in different experimental conditions

On average, the active users expressed lower positive affect (0.037, SD=0.13) than the passive users (0.041, SD=0.11); and they also expressed higher negative affect (0.50, SD=0.22) than the passive users (0.45, SD=0.22). Table 11 and Table 12 show the mean and standard deviation of the active and passive users' positive and negative affect under different experimental conditions. Task/technological conditions did not significantly alter the participants' affect, however, expertise conditions did. Participants worked in high expertise condition expressed lower positive affect and higher negative affect than those who worked in the low expertise condition (see Table 11 and Table 12).

Table 11: Mean (standard deviation) of positive affect values of the active users and the passive users under different experimental conditions.

		Normal	**Hard**	**Low reliability**	*Total*
High expertise	AU	0.03 (0.10)	0.02 (0.09)	0.04 (0.13)	*0.03 (0.14)*
	PU	0.02 (0.06)	0.03 (0.08)	0.03 (0.10)	
Low expertise	AU	0.04 (0.14)	0.04 (0.14)	0.05 (0.16)	*0.05 (0.10)*
	PU	0.05 (0.13)	0.05 (0.14)	0.06 (0.14)	
Total		*0.04 (0.11)*	*0.04 (0.12)*	*0.04 (0.13)*	

Note. AU, active user. PU, passive user.

Table 12: Mean (standard deviation) of negative affect values of the active users and the passive users under different experimental conditions.

		Normal	**Hard**	**Low reliability**	*Total*
High expertise	AU	0.53 (0.22)	0.56 (0.20)	0.53 (0.23)	*0.49 (0.24)*
	PU	0.46 (0.23)	0.44 (0.26)	0.44 (0.23)	
Low expertise	AU	0.49 (0.22)	0.48 (0.22)	0.45 (0.22)	*0.46 (0.21)*
	PU	0.45 (0.22)	0.45 (0.18)	0.45 (0.21)	
Total		*0.48 (0.22)*	*0.48 (0.22)*	*0.47 (0.22)*	

Note. AU, active user. PU, passive user.

4.2.4.2. Task/technological conditions, expertise, team member prior relationship, and integral affect similarity

Additional analyses were preformed to explore whether task/technological conditions, expertise, and team member prior relationship (whether the team members know each other before the experiment) could predict the similarity of integral affect within a team. LME models were fitted to the data using positive/negative affect similarity (as measured using the DTW technique described in section 4.1.6.2.1) as dependent variable. Task/technological conditions, team member prior relationship, expertise, and all the possible two-way interactions were used as fixed effects variables. Teams were used as random intercepts.

The results indicated that task/technological conditions had significant effects on affect similarities. For positive affect, mean similarity was higher in hard and low reliability conditions than that in the normal condition. The differences were significant except for the sample rate of 0.1Hz (for 0.1Hz sample rate, $b=-0.34$, $t(318)=-1.92$, $p=0.06$; for 1Hz sample rate, $b=-0.39$, $t(319)=-2.27$, $p<0.05$; for 2Hz sample rate, $b=-0.41$, $t(319)=-2.31$, $p<0.05$); see Figure 22. For negative affect, mean similarity was higher in low reliability condition than that in the hard and low reliability conditions (for 0.1Hz sample rate, $b=0.4$, $t(318)=2.29$, p<0.05; for 1Hz sample rate, b=0.49, t(319)=2.76, p<0.05; for 2Hz sample rate, b=0.49, t(319)=2.72, p<0.05); see Figure 23. The effects of other fixed effect variables were not significant.

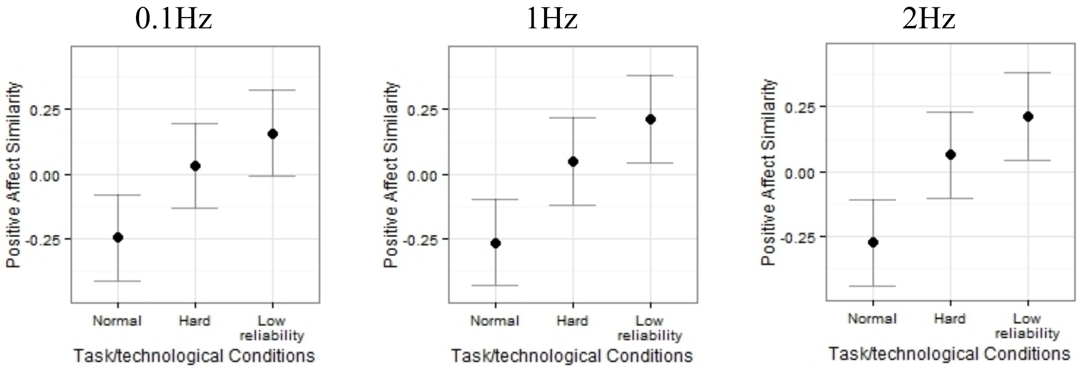


Figure 22. The predicted means and standard errors of similarities of positive integral affect. The original similarity index calculated from DTW technique was standardized and multiplied by -1, so that higher value indicates higher similarity. All the values were computed when the values of all the other variables held at their means.

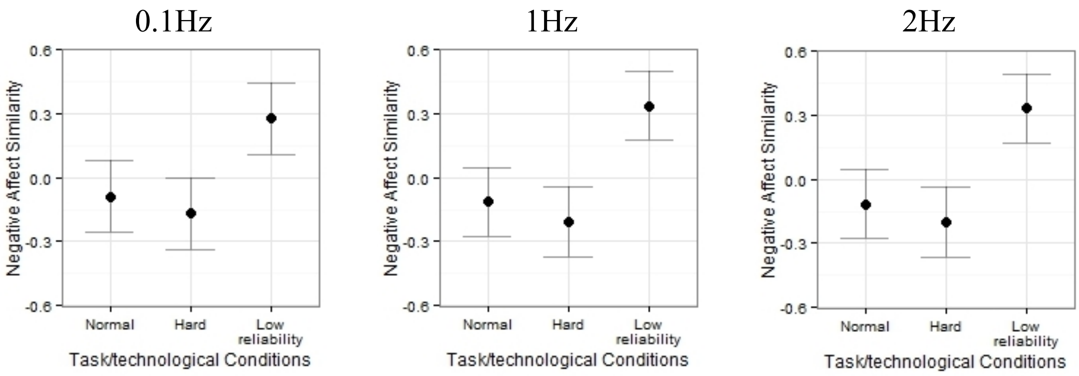


Figure 23. The predicted means and standard errors of similarities of negative integral affect. The original similarity index calculated from DTW technique was standardized and multiplied by -1, so that higher value indicates higher similarity. All the values were computed when the values of all the other variables held at their means.

4.2.4.3. Task/technological conditions and the mutual influence of the active user and the passive user

APIM framework was applied to the data to explore the effect of task/technological conditions on the mutual influence of the active user and the passive user. Similar procedure described in section 4.2.3 was used.

The tests showed that the partner effect X role X task/technological conditions three-way interaction effect was significant for positive affect sampled at 1Hz and 2Hz. Specifically, the active user's influence over the passive user was the weakest in low reliability condition compared to normal and hard conditions (for 1Hz, $b=0.09$, $t(3312)=6.7$, $p<0.05$; for 2Hz, $b=0.02$, $t(10408)=3.73$, $p<0.05$); and the comparative influence was stronger in the hard condition comparing to the normal condition (for 1Hz, $b=0.05$, $t(2741)=3.4$, $p<0.05$; for 2Hz, $b=0.03$, $t(4186)=4.57$, $p<0.05$). See Figure 24 and Figure 25 for visualizations. The test of simple effects of the partner effect X role interaction on different levels of task/technological conditions showed that the effect was significant on hard condition (for 1Hz, $b=0.08$, $t(7757)=6.83$, $p<0.05$; for 2Hz, $b=0.04$, $t(12163)=6.09$, $p<0.05$). These results indicated that the active user's influence over the passive user was stronger than the opposite under hard condition.

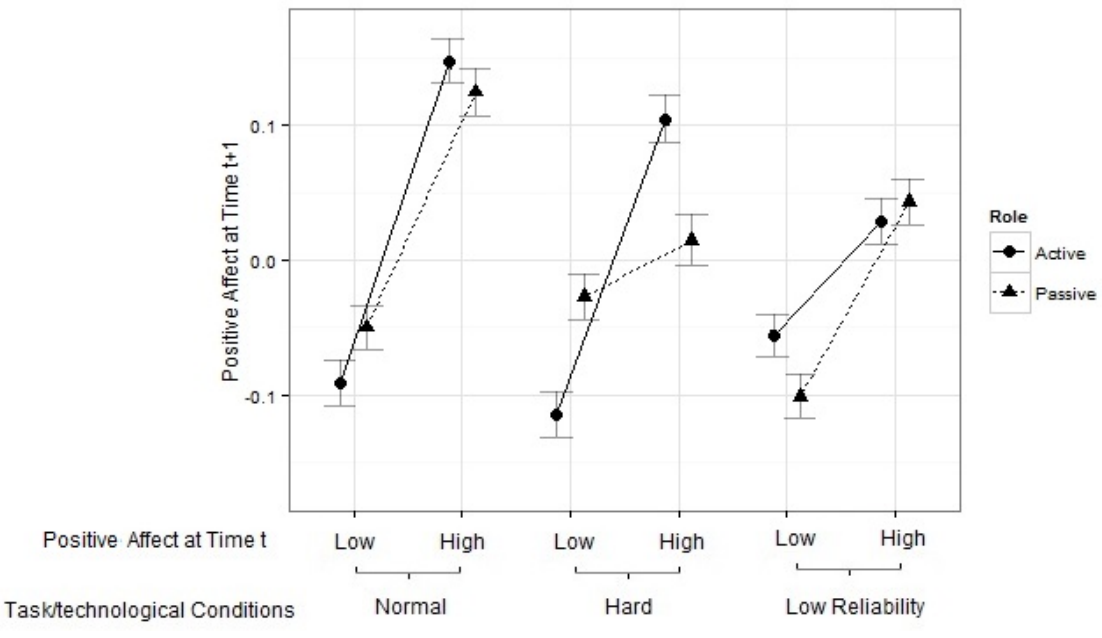


Figure 24. The predicted means and standard errors of actor's positive affect at time t+1 given partner's positive affect levels at time t (high was defined as the value of +1 SD and low was defined as the value of -1 SD), partner's role, and task/technological conditions. The sample rate is 1Hz. All the values were computed when the values of all the other variables held at their means.

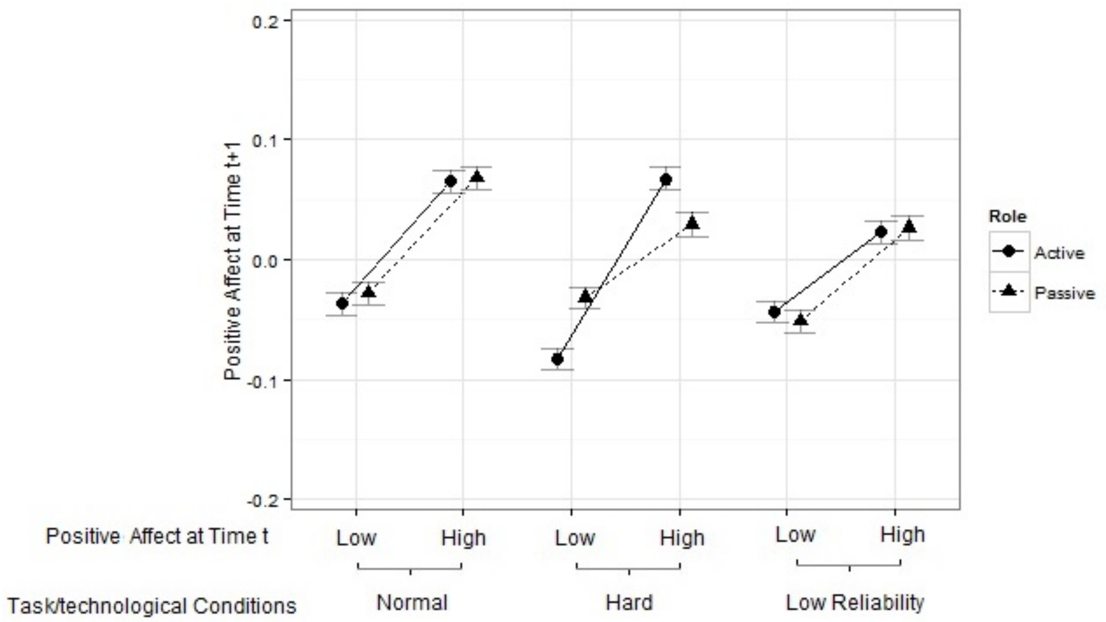


Figure 25. The predicted means and standard errors of actor's positive affect at time t+1 given partner's positive affect levels at time t (high was defined as the value of +1 SD and low was defined as the value of -1 SD), partner's role, and task/technological conditions. The sample rate is 2Hz. All the values were computed when the values of all the other variables held at their means.

This significant interaction effect was not observed for positive affect sampled in 0.1Hz (b=0.03, t(6871)=0.57, p=0.56), and negative affect sampled across all the sample rates (for 0.1Hz, b=0.03, t(5195)=0.72, p=0.46; for 1Hz, b=0.01, t(74266)=0.95, p=0.34; for 2Hz, b=0.002, t(29914)=0.22, p=0.82).

4.2.5. Summary of the findings

4.2.5.1. Summary of the findings for affect similarity

Figure 26 summarizes the main relationships found in this study. The correlations between these variables were statistically significant for sample rate at 1Hz and 2Hz. As the figure shows, team interaction increased the similarity of positive integral affect (hypothesis 1.1). Difficult task also increased the similarity of positive integral affect. Low technology reliability condition increased the similarity of both positive integral affect and negative integral affect.

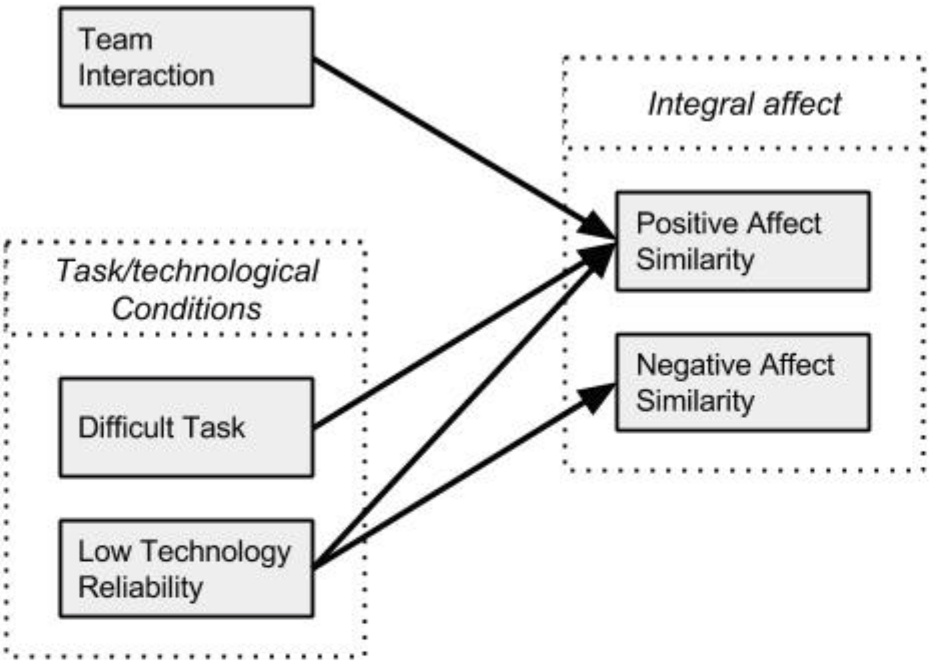


Figure 26. A summary of the relationships among team interaction, task/technological conditions, and similarity of integral affect within a team. These relationships were found at the sample rate of 1Hz and 2Hz.

At 0.1Hz sample rate, only two arrows were tested significant: team interaction increased similarity of positive affect and low technology reliability increased similarity of negative affect.

4.2.5.2.Summary of the findings for affect mutual influence within teams

For positive affect sampled under 1Hz and 2Hz, the findings are summarized in Figure 27 and Figure 28. If the passive user was not trained (low expertise condition), the task was normal (normal condition), or the technology was not reliable (low reliability condition), the mutual influence between the active user and the passive user in a team was similar (Figure 27).

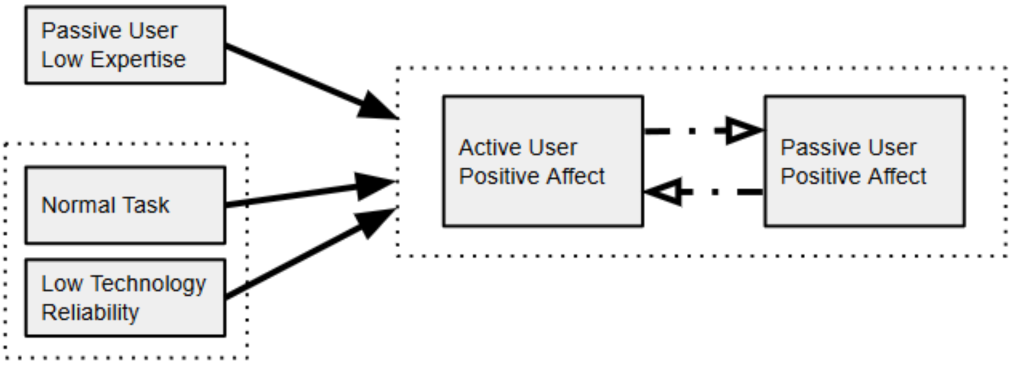


Figure 27. There were equivalent influences for positive affect between the active user and the passive user within a team under low expertise condition and normal and low reliability conditions, given the sample rate of 1Hz and 2Hz.

If the passive user was trained (high expertise condition) or the task was difficult (difficult condition), the active user's influence over the passive over was stronger than the passive user's influence over the active user (Figure 28).

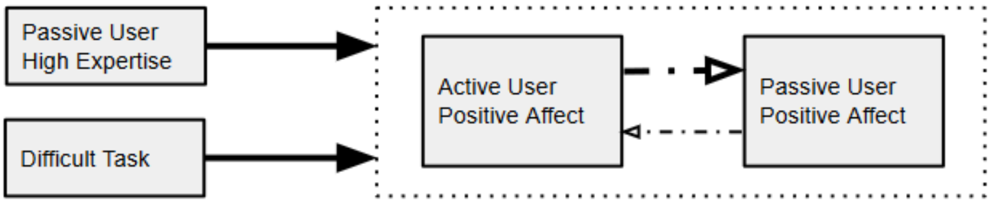


Figure 28. The active user's influence over the passive over was stronger than the passive user's influence over the active user for positive affect under high expertise condition and difficult task condition, given the sample rate of 1Hz and 2Hz.

However, the relationship mentioned in Figure 27 and Figure 28 was not found when the sample rate was 0.1Hz. Further, no mutual influence was found between the individuals in a team at this sample rate.

For negative affect sampled at 1Hz and 2Hz, the only significant relationship was that a team member's negative affect level negatively related to the other team member's negative affect level. In addition, this relationship was not found at 0.1Hz sample rate.

4.3. Discussion

4.3.1. Affective convergence of teams

This study identified factors that influenced the affective convergence of teams: team interaction, task demand, and technology reliability (see Figure 26). Through pseudo-couple analysis, the statistical tests were able to isolate the effect of task/technological conditions and test the effect of team interaction. Individuals worked in the same team did show higher similarity in positive affect with a relatively large effect size (on average, the distance measure of actual teams was smaller than that of pseudo teams by 0.38-0.41 unit of standard deviation depending on the sampling rate). When task demand was high (difficult condition) or technology reliability was low (low reliability condition), the team showed significantly higher similarity in positive affect, compared to normal condition. This could be a result of the team members' increased need to collaborate to cope with the demand under such operations. Due to the increased need for collaboration, the team members had to interact with each other more frequently. As emotional contagion is based on an unconscious process that individuals mimic others' facial expressions, vocal utterances, or postures (Hatfield, Carpenter, & Rapson, 2014), more frequent interaction could provide more opportunity for the team members to influence each other.

This finding resembled the bottom-up process of group affect discussed by Barsade et al. (Barsade & Gibson, 1998; Kelly & Barsade, 2001); see Figure 29. In this model, affect sharing forms the basis of group affect. Individual affect change is influenced by system characteristics during the task process; in this study, the system characteristics included task demand and technology reliability. System characteristics could also influence affect sharing. Additional examples of system characteristics that could influence affect sharing: physical layout of the task environment could be a facilitator or barrier for team communication (Stryker & Santoro, 2012) thus influence affect sharing; distance collaboration as a mode of cooperation could be a barrier for affect sharing depending on how the communication system is designed (J. T. Hancock, Gee, Ciaccio, & Lin, 2008).

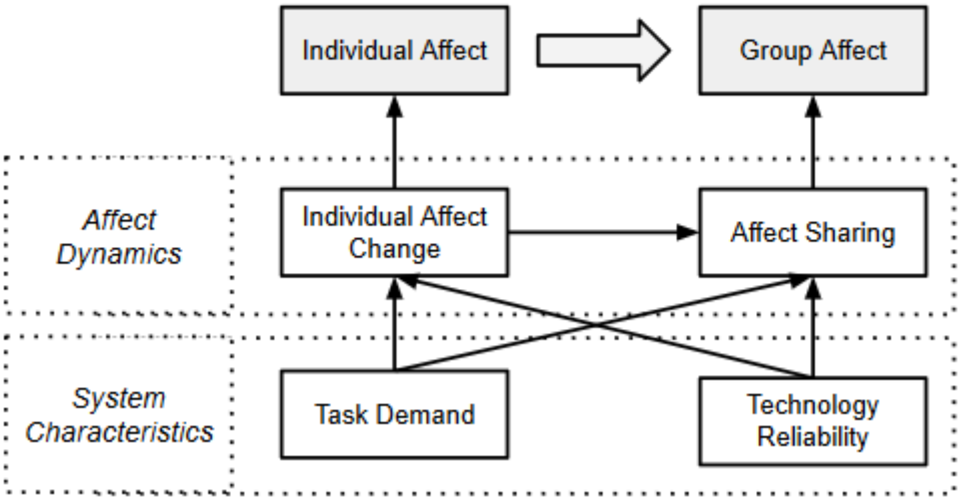


Figure 29. Bottom-up process of group affect found in this study.

However, the actual teams did not show higher similarity in negative affect compared to the pseudo teams. This result is counter-intuitive since research have shown that negative events are more salient and tend to elicit stronger and quicker emotional responses (Rozin & Royzman, 2001), and negative emotions may be more likely to lead to emotional contagion than positive emotions (Barsade, 2000). Given this results and a general lack of significance of in the tests associated with negative affect, one possible explanation is that the measurement of negative affect used in this study was not effective. Pooling the negative basic emotions (anger, disgust, and sad) as one negative affect measure may not be appropriate. In addition, basic emotions may not be the most instrumental way of categorizing emotions in the current task context. A review (D'Mello & Calvo, 2013) was conducted for studies in affective computing that tracked both basic and non-basic emotions. The results showed that non-basic emotions, including boredom, confusion, and frustration, occurred at five times the rate of basic emotions. Thus for the studies in related fields, instruments designed to measure those non-basic emotions should be used. In summary, follow up studies of this dissertation research should use tools that are capable of recognize facial expressions according to the valence dimension directly, or recognize non-basic emotions that are more relevant to technology use.

4.3.2. Using dynamic time warping (DTW) technique to calculate time series similarity

Developing similarity measures of time series data has been a challenge. In the research in physiological compliance (PC), similarity indexes are developed to measure the similarity of physiological signals (Elkins et al., 2009; Henning, Armstead, & Ferris, 2009; Montague et al., 2014). Examples of such PC similarity indexes included signal matching (SM), instantaneous derivative matching (IDM), directional agreement (DA), cross correlation (CC), and weighted coherence (WC). SM, IDM, and DA were developed by Elkins et al. (2009). CC and WC were used by Henning et al. (Henning, Boucsein, & Gil, 2001) to measure PC; WC was developed based on Porges et al.'s work (Henning et al., 2001; Porges et al., 1980). In SM, the differences between the areas of the two individuals' data curves are compared. The smaller the area between the curves, the lower the SM score. A low SM score indicates higher curve similarity and thus higher PC. In IDM, the slopes of the curves are compared. This is accomplished by averaging the differences between the instantaneous derivatives of corresponding points between the curves; the derivative of a point provides the tangent and thus the slope of the curve at that point. A low IDM score would indicate high similarity between the curves, and thus high PC. In the third PC measure, DA, the directional movement of the curve is assessed by comparing each point on the curve with its previous point. For example, an increasing directional movement would mean data point B is higher in value than data point A. A percentage of the two curves' directional agreement is then calculated; a higher percentage indicates higher PC. A CC coefficient is calculated to determine the covariance of corresponding data points in each physiological data curve at lag 0. While the PC indicators discussed above were derived from the original physiological signal, which recorded on the time domain, WC is a PC indicator that concerns frequency domain. WC quantifies the similarities of two individuals' physiology responses on a specified frequency band regardless of phase differences.

Although these different approaches are useful that they reflect unique characteristics of the synchronized time series, one common weakness of the previous listed similarity indexes (except WC) is that they cannot account for the lagged mutual influences of time series. In other words, when applied to the analysis of affect similarity, these indexes cannot account for emotional contagion. When emotional contagion happens, there will be a time lag between the occurrences of two similar affective values. Furthermore, this time lag is not a constant – it could be positive or negative. This is because any one of the two individuals in a team could be the one who initiate the contagion. The DTW approach can account for emotional contagion as its algorithm searches for a flexible alignment of two time series.

However, the DTW alignment algorithm has its limitations. A significant one was that it was designed to find an "optimal path" which minimizes the distance measure of the two time series (Berndt & Clifford, 1994); however, this optimal path may not be the "accurate" path that equal to the extract pattern corresponding to the emotional contagion.

A problem related to the calculation of affect similarity in this study was the sample rate of the affect data. The original sample rate of the time series was 12Hz, which was too high for affect sharing process. In the analysis, a "shotgun" approach was used that three different sample rates were included: 0.1Hz, 1Hz, and 2Hz. The results showed that 1Hz and 2Hz yielded very similar results while the results from 0.1Hz was different from the two. As the timing of facial mimicry was close to the sample rate of 1Hz and 2Hz (Dimberg et al., 2000; Mancini et al., 2013), these two sample rates should be more appropriate for analyses based on facial expression. Future research could identify the optimal sample rate base on measure instruments and research questions.

4.3.3. Training and mutual influence of affect

Hypothesis 1.3 stated that the active user's influence over the passive user is stronger than the passive user's influence over the active user. In the analysis, this hypothesis was only partially supported for positive affect (see Figure 18, Figure 19, and Table 9). Specifically, the hypothesis was not supported under low expertise condition where the mutual influence seemed to be equivalent for the two roles. Hypothesis 1.4 was not supported for positive affect as the active user's influence over the passive user was higher in the high expertise condition than that in the low expertise condition (see Figure 21).

This might be an effect of team cross-training. Team cross-training is designed to have team members trained in the duties of their teammates (Volpe, Cannon-Bowers, Salas, & Spector, 1996). Cross-training has three levels with different depth of training (Blickensderfer, Cannon-Bowers, & Salas, 1998). In the first level (positional clarification), the team receives verbal presentations of all the team members' jobs. In the second level (positional modeling), the team members observe each other's job in addition to verbal presentations. In the third level (positional rotation), the team members will go through hands-on experience of carrying out each other's role. Research have shown that cross-training enhances team shared mental models (Marks, Sabella, Burke, & Zaccaro, 2002) and improves team performance (Volpe et al., 1996). In the low expertise condition of this study, the passive user received minimum training before the task by reading a document that contained information about the goal of the task, how to control the computer, and the team members' roles. This was a positional clarification level of cross-training. In the high expertise condition, the passive user received the same hands-on training as the active user. This training offered the passive user the experience of acting as an operator, which was close to the experience of an active user. Thus this could be considered a positional rotation level of cross-training. Teams received effective cross training may lead to team members' increased ability to monitor each other's performance (Eduardo Salas, Sims, & Burke, 2005). The passive user in high expertise condition might have a higher awareness of the active user's status such as affective state. Due to the higher awareness of the active user's affective state, the passive user's affect was influenced by the active user to a greater extent.

Another effect of cross-training is that the team is more likely to use implicit communication (Espevik, Johnsen, & Eid, 2011). Under the context of this study, the increased use of implicit communication might lead to reduced awareness of team member's affective state. However, this might only apply to the active user. There are two reasons. First, since the active user's workload was higher than the passive user's workload (Montague & Xu, 2012), reduced need for explicit communication could cause the active user to pay more attention to operational task. Second, for the passive user, the implication of reduced need for explicit communication was that he/she can allocate more resource to monitor the task and the active user's status. These could explain the pattern of unequal mutual affective influence under the high expertise condition.

4.3.4. Additional limitation and future direction

One if the important limitations of this study was its generalizability given the context of the task. To evaluate the generalizability of the results of this study, taxonomies and theories in the field of computer supported collaborative work (CSCW) and general team/group research (Grudin & Poltrock, 2012) should be considered. In the Four-Square Map of Groupware Options (Johansen et al., 1991), groupware can be categorized in four types based on two dimensions: same/different time and same/different location. Nuanmaker et al. (1991) also considered group size as an important dimension as it often introduces qualitative differences in the interaction process. This study focused on tasks that required same time and same location collaboration. Such a scenario represented the typical way of active and passive use of shared technology. Also the interaction of one active user and one passive user was considered in this study, as active and passive use of shared technology often happens among dyads or small groups/teams. Dyads and groups with more than two members may have qualitative differences (Moreland, 2010; Williams, 2010). Moreland (2010) argued that emotions are stronger in dyads than in groups, because dyads often have closer relationships and emotions flows more directly in dyads. Thus the generalization of the results of this study to groups or teams with larger size should be taken with caution.

The nature of task performed with MATB should also be considered when generalizing the results. According to McGrath's group task circumplex model (Larson, 2010; McGrath, 1984; Mennecke & Wheeler, 1993; Straus, 1999), there are eight mutually exclusive and collectively exhaustive group tasks: planning tasks, creativity tasks, intellective tasks, decision making tasks, cognitive conflict tasks, mixed-motive tasks, competitive tasks, and psycho-motor tasks (see Figure 30).

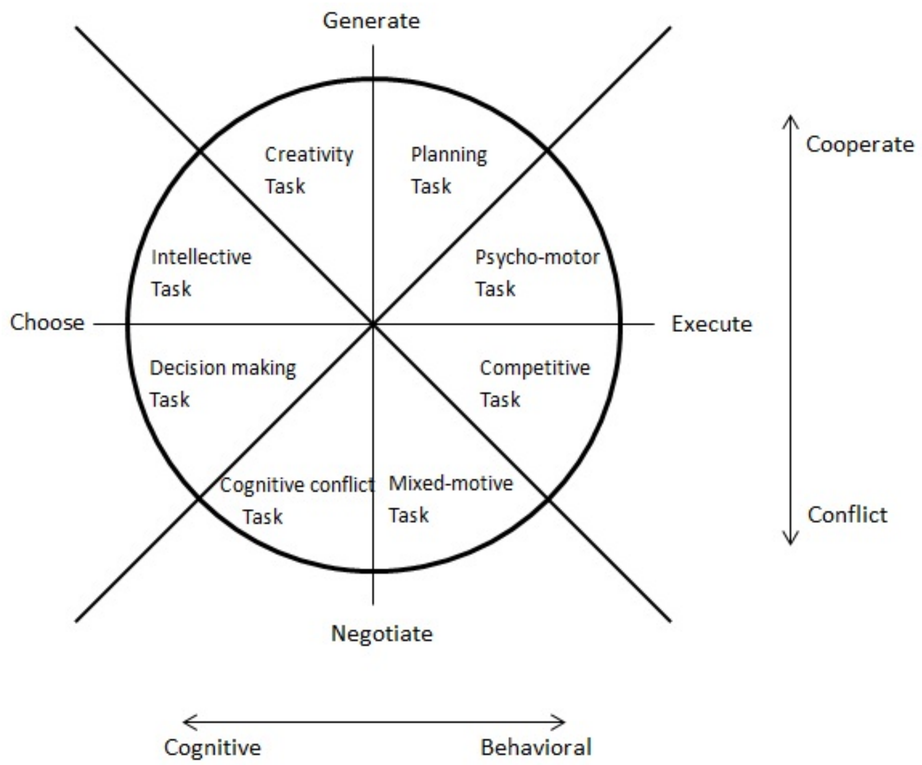


Figure 30. McGrath's group task circumplex model (1984).

In the case of active and passive use of shared technology, team tasks could be under the quadrants of generate, choose, or execute. The tasks in MATB are primarily psycho-motor tasks. Tasks under other quadrants may involve very different communication process. This different communication process could lead to differences in how affect is shared. Furthermore, affect may play different roles in such tasks. Future research should explore other types of tasks to expand the understanding of the dynamics of integral affect in teams.

5. Chapter V: Study 2 – Affect and Trust in Technology

5.1. Methods

5.1.1. Participants

The participants of this study were recruited from the University of Wisconsin – Madison through an online posting on the Student Job Center (https://jobcenter.wisc.edu/) website. 54 participants participated in this study in two-person teams (n=27). $10 cash compensation was offered for all the participants. As an incentive for team performance, three teams that achieved top performance was rewarded with $20 cash for each of their team members.

The participants ranged in age from 18 to 56 (mean = 22.31, SD = 5.42). 34 of the participants were female (62.96%). There were 12 female-female teams (44.44%), 6 male-male teams (22.22%), and 9 mixed gender teams (33.33%). 32 participants (59.26%) reported that they were Caucasians, and 19 participants (35.19%) reported that they were Asian. Other reported ethnicity groups included African American and Hispanic. In terms of school years, 7 participants (12.96%) were freshman, 13 (24.07%) were sophomore, 5 (9.26%) were junior, 17 (31.48%) were senior, 1 (1.85%) was fifth year student, and 11 (20.37%) were graduate school students. The participants majored in a wide variety of disciplines, for example, industrial engineering (n=4), business (n=4), psychology (n=3), accounting (n=2), biology (n=2), electrical engineering (n=2), economics (n=2), journalism (n=2), political science (n=2), etc. Finally, participants in 7 teams (25.93%) reported that they did not know their teammate prior to the experiment.

5.1.2. Design

This study was a mixed design with one within-subject variable and two between-subject variables. The first between-subject variable was the same as study one: role of participants (active user/passive user) that was manipulated on the individual level.

The second between-subject variable was initial mood of the teams, or incidental affect, which was manipulated on the team level. There were three conditions: positive, negative, and neutral. There are several methods developed for mood induction in the laboratory, for example, viewing film clips (J. J. Gross & Levenson, 1995) or pictures (Bradley & Lang, 2007), listening to music (Eich, Ng, Macaulay, Percy, & Grebneva, 2007; Västfjäll, 2002), performing guided dyadic interaction (Roberts, Tsai, & Coan, 2007), and performing guided facial actions (R. W. Levenson, Ekman, & Friesen, 1990). Emotional film clips and pictures were widely used in human factors research (Dzindolet et al., 2010; Helton & Russell, 2011; Lim, Woo, Bahn, & Nam, 2012; Merritt, 2011; Raddatz, Werth, & Tran, 2007; Stokes et al., 2010; Tran, Raddatz, & Werth, 2008). Previous meta-analysis showed that watching film clips is a very effective mood induction technique (Gerrards - Hesse, Spies, & Hesse, 1994; Westermann, Spies, Stahl, & Hesse, 1996). A recent study (Ellard, Farchione, & Barlow, 2012) suggested that the effectiveness of mood induction for viewing film clips and pictures are similar. In this study, mood states were induced using the International Affect Picture System (IAPS) (Lang et al., 2008). A set of 90 validated images (30 positive, 30 negative, and 30 neutral images) were used. The participants in a team viewed sets of images according to the treatment condition they received. In each condition, the 30 images were assigned to three sets with 10 images in each set. So there were three different sets of images in each condition. These different sets of images were used for mood induction before each task trials. The effectiveness of the mood induction was tested through a pilot study.

The within-subject variable were the task/technological conditions which were manipulated in the same way as study one. The three conditions were normal condition, hard condition, and low reliability condition. The sequence of presenting these three conditions was randomized across the teams.

The overall design of the study is shown in Figure 31.

Overall study:

Initial mood		
Positive	n = 18 (9 groups)	**Task/technological conditions** (randomized sequence) N, H, L
Negative	n = 18 (9 groups)	**Task/technological conditions** (randomized sequence) N, H, L
Neutral	n = 18 (9 groups)	**Task/technological conditions** (randomized sequence) N, H, L

Within each group:

Role	Active user	Passive user
Number of participants	1	1

Figure 31. Design of the experiment of study two.

5.1.3. Apparatus and laboratory setting

The laboratory setting of this study was the same as study one.

As reported previously, a total number of 90 images were selected from IAPS to use as materials for mood induction. The images were selected to make sure there were variety in the contents of the images. Positive, neutral, and negative images were selected based on the valence ratings based on IAPS publications (Lang et al., 2008). For positive images, the contents included families, sports and adventures, babies, animal cubs, and erotic couples. The average valence (ranged 1 to 9) was 7.51 and average arousal (ranged 1 to 9) was 4.93. For neutral images, the contents included household objects, mushrooms, and human faces with neutral expressions. The average valence was 5.04 and average arousal was 3.07. For negative images, the contents included accidents, contamination, attacking animals, attacking humans, mutilated bodies, and dead animals. The average valence was 2.32 and average arousal was 6.26. Table 13 shows the average valence and arousal values for each of the selected images sets.

Table 13: Average valence and arousal values for each of the selected images sets.

Mood induction condition	Set number	Average valence	Average arousal
Positive	Set #1	7.526	5.054
	Set #2	7.399	5.075
	Set #3	7.614	4.673
Neutral	Set #1	5.212	3.299
	Set #2	4.985	3.115
	Set #3	4.929	2.786
Negative	Set #1	2.219	6.109
	Set #2	2.378	6.305
	Set #3	2.352	6.369

A pilot study with a sample size of 18 participants (9 teams) were conducted to examine the effectiveness of the mood induction and to examine if the effect of different sets of images in the same condition had significant difference on the intensity of mood induced. The design of the pilot study was the same with the actual study except that the participants were asked to fill out PANAS after the mood induction rather than after the task trial. Statistical tests were conducted to compare the effects of mood induction between conditions as well as the effects of different sets of images within the same mood condition.

5.1.4. Procedure

The procedure of this study was similar with study one but with two differences (Figure 32). First, in the training session, both the active user and the passive user received the same level of hands-on training since expertise was not manipulated. Second, mood induction sessions were added to the procedure. Before each task trial, the participants were asked to view a same set of images from IAPS individually. The images were shown on the computer screen and each image was displayed for 10 seconds. During the image display period, the participants were instructed to answer a multiple choice question about if the image was taken with a digital camera or a film camera. After the mood induction session, the participants would proceed to the first task trial. Another set of images was used for the second trial. The last set of images were used for the last task trial. The sequence of showing the three sets of images were randomized across teams.

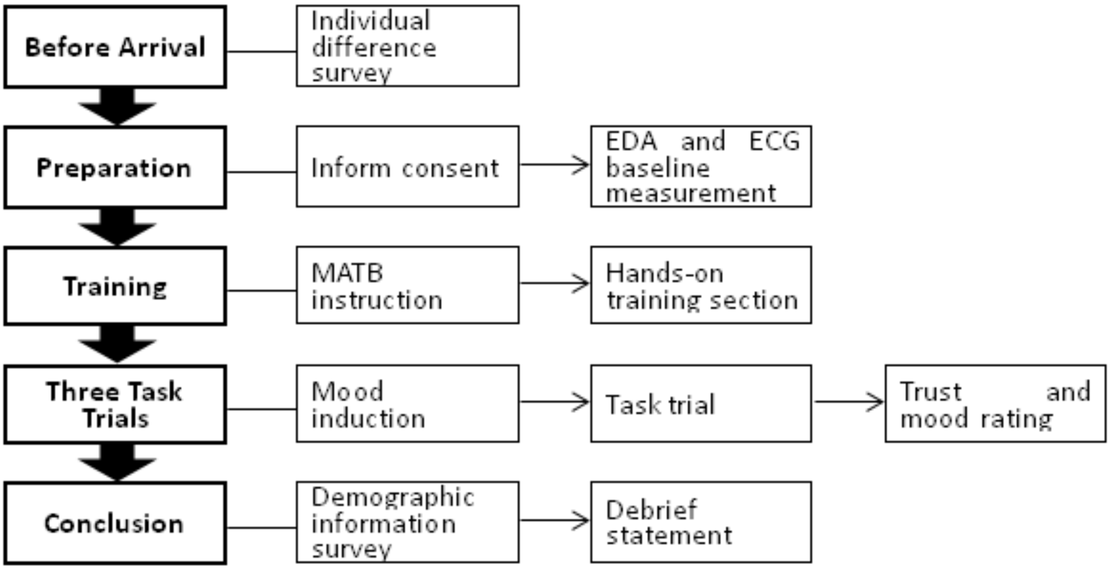


Figure 32. Procedure of the experiment of study two.

5.1.5. Measurements

5.1.5.1. Self-report measures

Trust in technology were measured by the trust in automated system scale developed by Jian, Bisantz, and Drury (Jian et al., 2000). This validated scale had been widely used in human automation interaction studies with high reliability (Madhavan et al., 2006; Merritt & Ilgen, 2008; Stokes et al., 2010). The Positive and Negative Affect Schedule (PANAS) were used as a measure of end-of-task mood state. These two scales were administered to the participants individually after each task trial.

At the end of the experiment, the participants finished a demographic information survey which included items such as age, gender, ethnicity, major, and year of school. The familiarity of the participants with each other in a team were measured with one item in a 1 (did not know each other prior to the experiment) to 5 (good friends) points scale. There was also a follow-up open ended question asking about how the participants met each other.

4.3.5.2. Affective states

Positive and negative integral affect was measured by facial expression recognition. Time series variables indicating the levels of positive negative affect were derived through the same method used in study one as described in section 4.2.5.1. Several indicators of integral affect were calculated, including average value of affect in the first 60 seconds of a recording, average value of affect in the last 60 seconds of a recording, average value of affect in the full 360 seconds of a recording, and the peak value of affect.

5.1.6. Data analysis

5.1.6.1. Hypothesis 2.1: initial mood state (incidental affect) influences an individual's affective state during the interaction process (integral affect)

To test this hypothesis, the effect of initial mood condition (positive/negative/neutral) on positive or negative integral affect was tested. An LME model was fitted to the data using integral affect as dependent variable, initial mood condition (positive/negative/neutral) as independent variable, and individuals nested in teams as random intercepts. In addition, task/technological conditions, role, and the interaction terms were controlled (see Table 14).

Table 14: LME model specifications for study two.

	Dependent variable	**Fixed effect variables**	**Random effect variables**
Hypothesis 2.1	1. Positive or negative integral affect	1. Initial mood condition 2. Task/technological conditions 3. Role 4. Interaction terms (all possible two-way interactions)	1. Individuals 2. Teams
Hypothesis 2.2	1. Trust in technology	1. Task/technological conditions 2. Initial mood condition 3. Task/technological conditions X Initial mood condition interaction 4. Role 5. Other interaction terms (all possible two-way interactions)	1. Individuals 2. Teams
Hypothesis 2.3 model 1	1. Positive or negative integral affect	1. Initial mood condition 2. Task/technological conditions 3. Role 4. Interaction terms (all possible two-way interactions)	1. Individuals 2. Teams
Hypothesis 2.3 model 2	1. Trust in technology	1. Positive or negative integral affect 2. Initial mood condition 3. Task/technological conditions 4. Role 5. Interaction terms (all possible two-way interactions)	1. Individuals 2. Teams

5.1.6.2.Hypothesis 2.2: initial mood state (incidental affect) affects trust in technology

The path diagram to guide the test of this hypothesis is shown in Figure 33. To test the effect of initial mood state on trust in technology, an LME model was fitted to the data using trust in technology as dependent variable, initial mood condition (positive/negative/neutral) and task/technological condition (normal/hard/low reliability) as independent variables, and individuals nested in teams as random intercepts. The effect of the role of the participants and the interaction terms were controlled (see Table 14).

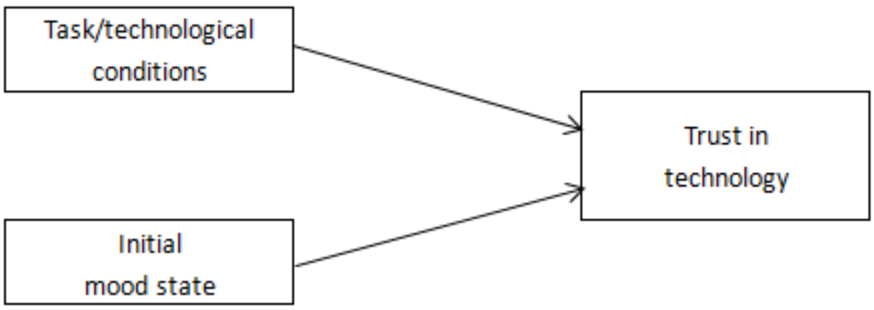


Figure 33. Path diagram of the model for hypothesis 2.2.

5.1.6.3. Hypothesis 2.3: affective state of an individual during the task (integral affect) mediates the effect of initial mood state (incidental affect) on trust in technology

A mediator is a variable that lies in a causal sequence of two variables (Baron & Kenny, 1986; Wu & Zumbo, 2008). If mediation effect exists, part or all of the effect of the predictor variable on the outcome variable will "go through" the mediator variable. This hypothesis stated that the affect state of the participant mediates the effects of two variables on trust in technology. A diagram that shows the mediation model is shown in Figure 34.

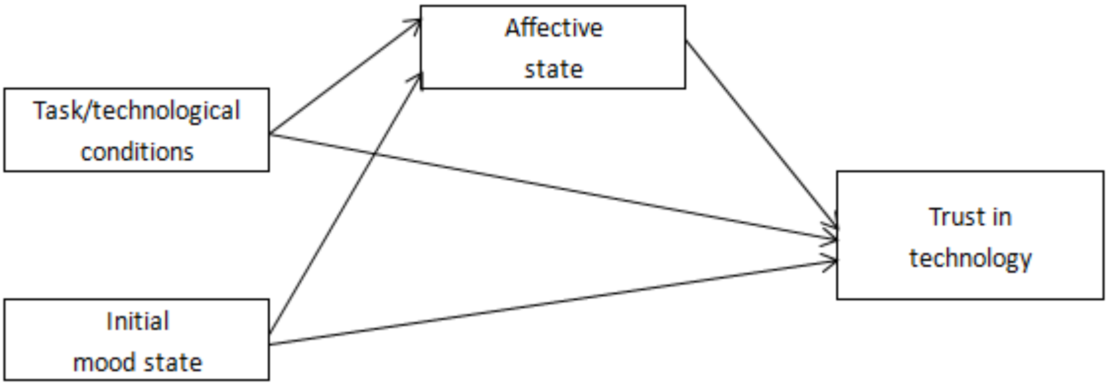


Figure 34. Path diagram of the mediation model for hypothesis 2.3.

A number of procedures have been proposed for detecting mediation effect (Fritz & MacKinnon, 2007; MacKinnon, Fairchild, & Fritz, 2007). The classical treatment of mediation effect is called the causal steps test proposed by Baron and Kenny (1986). In this test, assuming there are one predictor variable X, one mediator M, and one outcome variable Y, three linear models are needed to be fitted to the data:

$$Y = i_1 + cX + e_1 \tag{1}$$

$$M = i_2 + aX + e_2 \tag{2}$$

$$Y = i_3 + c`X + bM + e_3 \tag{3}$$

To test the presence of the mediation effect, one should follow four steps. First, the total effect of X on Y must be significant (the coefficient c in formula (1)). Second, the effect of X on M must be significant (the coefficient a in formula (2)). Third, the direct effect of M on Y controlled for X must be significant (the coefficient b in formula (3)). Fourth, the effect of X on Y controlled for M must be smaller than the total effect of X on Y; that is to say, c` must be smaller than c. There are two limitations in this approach. First, the requirement of the significance of the total effect may exclude the situation that direct and indirect effects have opposite signs and cancel out each other in total effect (MacKinnon, Krull, & Lockwood, 2000). Also, if the causal process between X and Y was distal, the required sample size to detect a significant total effect will be large (Shrout & Bolger, 2002). Second, the power of detecting mediation effect is very low when this approach is used, according to previous simulation studies (Fritz & MacKinnon, 2007; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002).

Some of the other approaches of mediation analysis focused on measuring indirect effects in the linear models directly. One of such tests is product of coefficients test (MacKinnon & Dwyer, 1993). The product of coefficients test uses the product $\hat{a}\hat{b}$ to estimate indirect effect. A number of analytical procedures of estimating the standard error of this product have been developed (MacKinnon et al., 2002; Sobel, 1987). However, confidence intervals derived from the estimated standard error and assumed normal or student's t distribution are often inaccurate (Bollen & Stine, 1990; MacKinnon et al., 2002). The validity of the significant testing is also in doubt as a result. Computer intensive resampling method is recommended by some researchers (Bollen & Stine, 1990; MacKinnon, 2007; Preacher & Selig, 2012), since resampling methods do not require as many assumptions as other tests and provide a more accurate results.

To test hypothesis 2.3 of the current study, first, two LME models were fitted to the data. The first model resembled formula (2) in Baron and Kenny's approach. In this model, average level of positive or negative affective state were entered as the dependent variable; initial mood condition (positive/negative/neutral) and task/technological conditions (normal/hard/low reliability) were entered as independent variables; and individuals nested in teams were entered as random intercepts. In addition, the role of the participants (active/passive user) and the interaction terms were controlled (see Table 14). The vectors of confidents for initial mood state ($\widehat{a_1}$) and task/technological conditions ($\widehat{a_2}$) were obtained. The second model resembled formula (3) in Baron and Kenny's approach. In this model, trust in technology were entered as the dependent variable; average level of positive or negative affective state, initial mood condition (positive/negative/neutral), and task/technological conditions (normal/hard/low reliability) were entered as independent variables; and individuals nested in teams were entered as random intercepts. In addition, role of the participant (active/passive user) and interaction terms were controlled (see Table 14). After the model fit, the vector of confidents for average level of positive or negative affective state ($\hat{b}$) were obtained.

Second, the indirect effects were estimated by the products of coefficients ($\widehat{a_1}\hat{b}$ and $\widehat{a_2}\hat{b}$). Parametric bootstrapping (Pituch, Stapleton, & Kang, 2006; Zhang et al., 2008) was used to calculate the confidence interval and perform the significance test. The procedure was as follows: first, a simulated dataset was generated according to the model estimation from the data. Specifically, the design matrix of the original data was preserved, and the parameter estimates were used for the simulation. The sample size was equal to the sample size of the data. Second, LME models described above were fitted to the sampled data and the products of coefficients were calculated. Third, the first two steps were repeated 1,000 times thus 1,000 different estimations of the products of coefficients were obtained. 1,000 replications were chosen since previous research showed that this amount of replication offers good results if α was set to 0.05 (Manly, 2007; Marriott, 1981). The point estimations of the products of coefficients were the mean values of the 1,000 samples. The estimated standard errors were the standard deviations of the 1,000 samples. The 5% and 95% limits are defined as the 25th and 976th score in the distribution of the values of 1,000 samples. For a product of coefficient, if its 95% confidence interval did not contain zero, one can conclude that the indirect effect is significantly different from zero at $p<0.5$ (two tailed).

5.2. Results

5.2.1. Pilot study results

To test the effectiveness of mood induction, LME models were fitted to the pilot study data using positive affect and negative affect measured by PANAS as dependent variables, initial mood conditions (positive/neutral/negative) as fixed effect variable, and teams and individuals as random intercepts. The results indicated that both positive affect ($F(2, 15)=4.16$, $p<0.05$) and negative affect ($F(2, 6)=5.83$, $p<0.05$) were significantly influenced by initial mood conditions (see Figure 35). Specifically, positive affect value was higher in positive condition than the mean of neutral and negative condition ($b=1.14$, $t(15)=2.88$, $p<0.05$). Negative affect value was higher in negative condition than the mean of positive and neutral condition ($b=1.22$, $t(6)=3.33$, $p<0.05$).

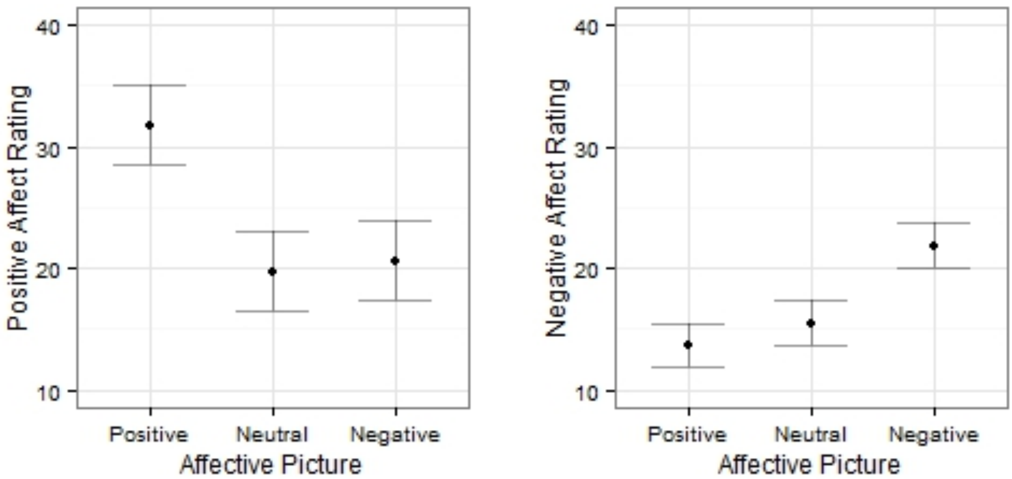


Figure 35. The mean and standard error of the predicted values of PANAS affect rating (positive and negative) on positive, neutral, and negative initial mood conditions.

5.2.2. Integral affect data description

There were a total number of 162 video recordings from the experiment (54 participants X 3 experimental trials). CERT was used to analyze the video recordings and generated time series for six basic emotions (anger, disgust, fear, joy, sad, and surprise). Each time series contained 4,320 data points as the videos had 12 frames per second and were 360 second in length.

Similar to first study, fear and surprise were dropped from further analysis due to very low average values and very low percentages of total emotional state values expressed by the participants (less than 5%). The value of joy was used as a measure of positive affective state.

The sum of the values of anger, disgust, and sad was used as a measure of negative affective state. The average percentage of missing values in the videos was 6.82% (minimum 0%, 1st quartile 0.26%, median 2.16%, 3rd quartile 11.14%, maximum 48.47), with a standard deviation of 9.64%. 14 out of 162 videos had a missing data point percentage of 20%. These numbers were very similar to study 1.

Average values of positive affect and negative affect were calculated for the first 60 seconds of a recording, the last 60 seconds of a recording, and the full 360 seconds of a recording. The peak values were also extracted from the full 360 seconds of recording. Missing cells were dropped in the calculations. Please refer to Table 15 for descriptions of the data. In the subsequent analysis, results were done with all four ways of deriving affective values (mean of first 60 seconds, mean of last 60 seconds, mean of full 360 seconds, and peak value). However, this report includes the results from the analysis done using mean of last 60 seconds and peak values.

Table 15. Averages and standard deviations of positive and negative affect derived from CERT outputs.

Affect valence	**Recording period**	**Average**	**Standard deviation**
Positive	360 seconds	0.048	0.056
	First 60 seconds	0.057	0.086
	Last 60 seconds	0.038	0.057
	Peak value	0.721	0.273
Negative	360 seconds	0.458	0.183
	First 60 seconds	0.449	0.189
	Last 60 seconds	0.468	0.193
	Peak value	0.940	0.077

5.2.3. The effect of incidental affect and task/technological conditions on integral affect

To test the effect of incidental affect and task/technological conditions on integral affect, LME models with specifications detailed in Table 14 were fitted to the data. The results indicated that both of the independent variables had significant effects on positive integral affect (for incidental affect, F(2, 23)=7.78, $p<0.05$; for task/technological conditions, F(2, 93)=6.25, $p<0.05$). Specifically, for incidental affect conditions, positive integral affect was higher in positive condition than the mean of neutral and negative condition (b=0.82, t(23)=3.89, $p<0.05$). For task/technological conditions, positive integral affect was lower in low reliability condition than the mean of normal and hard condition (b=0.41, t(94)=3.29, $p<0.05$). Figure 36 visualizes these effects under the condition that the values of all the other independent variables were held at their mean.

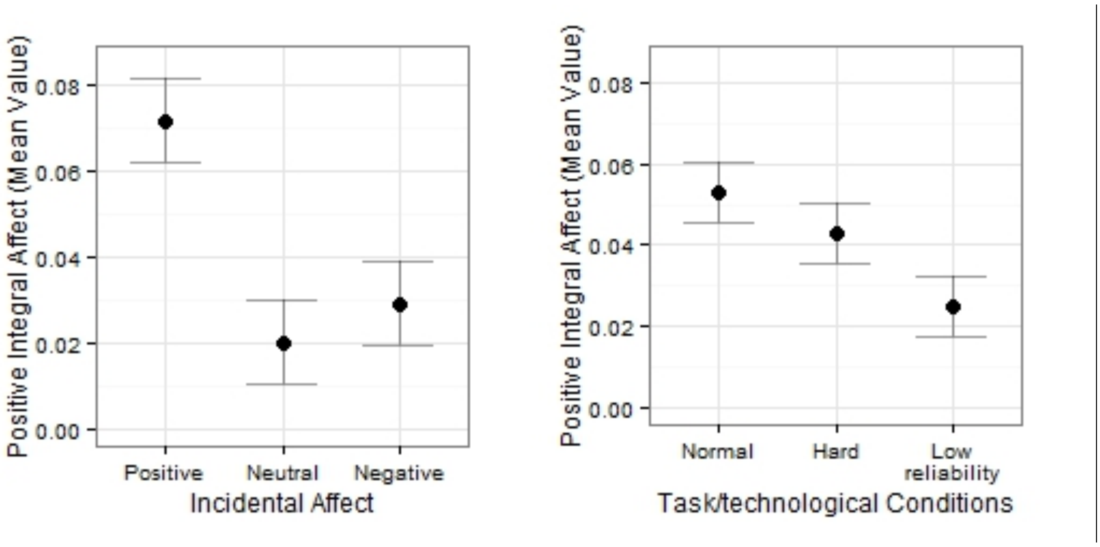


Figure 36. The mean and standard error of the predicted values of mean positive affect during the trial (last 60 seconds) on different initial mood conditions and different task/technological conditions. These plots were created given the values of all the other independent variables were held at their mean.

The tests also showed that the effect of incidental affect on the peak value of integral affect was significant (F(2, 24)=3.45, $p<0.05$). Peak value of positive integral affect was higher in positive condition than the mean of neutral and negative condition (b=0.79, t(49)=3.21, $p<0.05$). The effect of task/technological conditions was not significant (F(2, 98)=2.40, p=0.10). These effects are visualized in Figure 37.

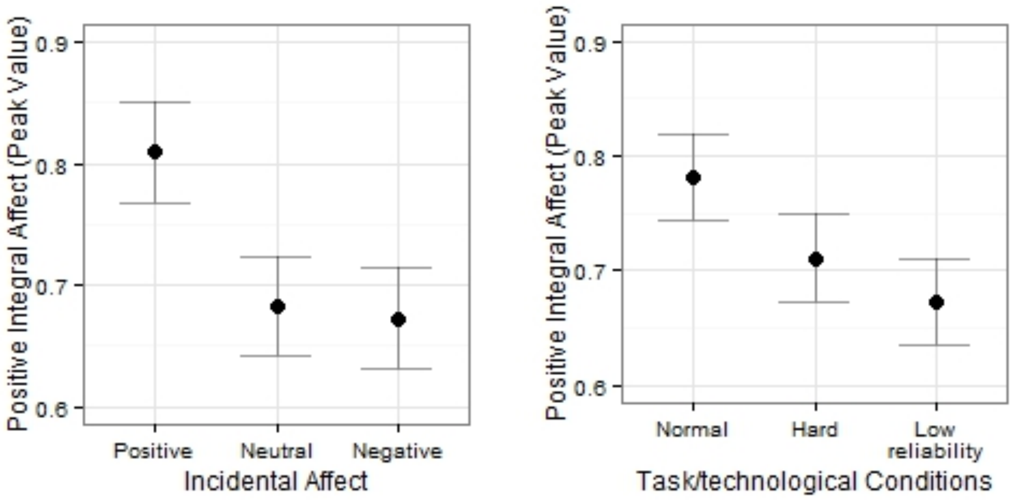


Figure 37. The mean and standard error of the predicted values of peak positive affect during the trial on different initial mood conditions and different task/technological conditions. These plots were created given the values of all the other independent variables were held at their mean.

For negative affect, no significant effect was found for incidental affect conditions ($F(2, 142)=0.37$, $p=0.69$). However, the effect of task/technological conditions was significant ($F(2, 142)=5.06$, $p<0.05$). Specifically, negative integral affect was lower in normal condition than the mean of hard and low reliability condition ($b=0.76$, $t(142)=2.69$, $p<0.05$). Figure 38 shows visualizations of the mean negative affect values given different incidental affect conditions and task/technological conditions.

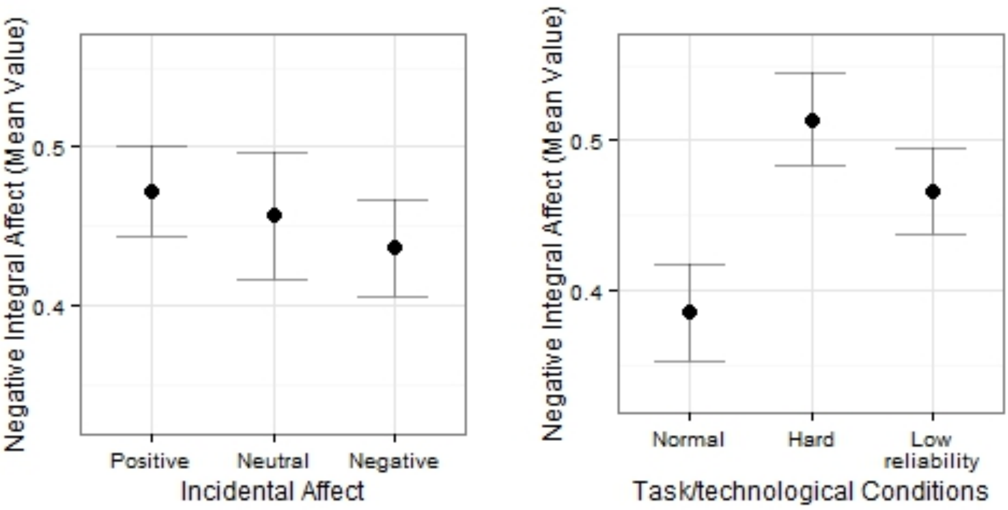


Figure 38. The mean and standard error of the predicted values of mean negative affect during the trial (last 60 seconds) on different initial mood conditions and different task/technological conditions. These plots were created given the values of all the other independent variables were held at their mean.

For the test on peak values of negative affect, neither the effect of incidental affect (F(2, 27)=0.57, p=0.57) nor the effect of task/technological conditions (F(2, 123)=2.45, p=0.09) was significant.

In summary, hypothesis 2.1, which states that incidental affect influences integral affect, was supported for the positive integral affect but not for the negative integral affect.

5.2.4. Affect and trust in technology

5.2.4.1. The effect of incidental affect and task/technological conditions on trust in technology

Using trust in technology as the dependent variable, LME model specified in Table 14 was fitted to the data. The results suggested that the main effect of task/technological conditions was significant (F(2, 100)=11.03, p<0.05) while the main effect of incidental affect was not significant (F(2, 24)=0.19, p=0.83). Trust in technology in the low reliability condition was lower than the mean of normal and hard conditions (b=-0.52, t(100)=-3.03, p<0.05). Please refer to Figure 39 for visualization.

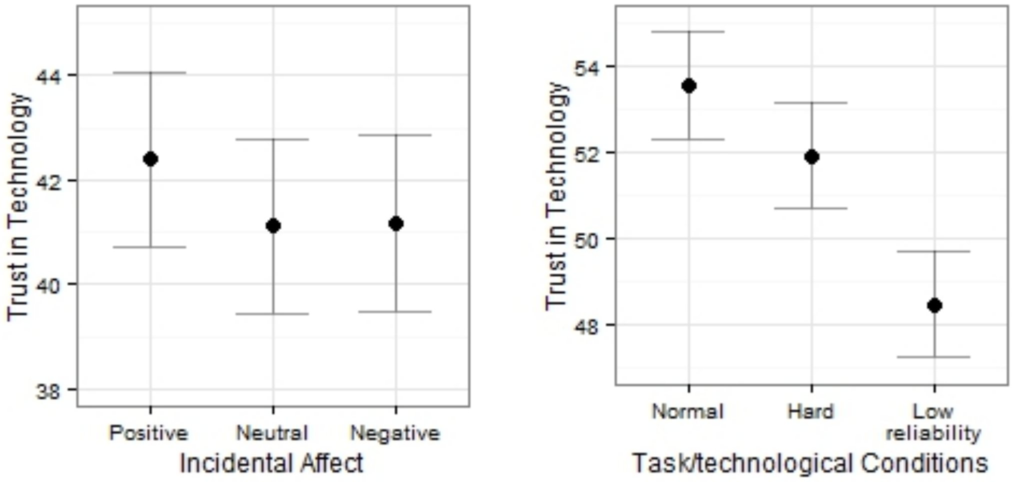


Figure 39. The mean and standard error of the predicted values of trust in technology on different initial mood conditions and different task/technological conditions.

These results suggested that hypothesis 2.2 was not supported. Incidental affect did not have a significant effect on trust in technology in this study.

5.2.4.2.Integral affect and trust

The significance (or lack thereof) of the main effects held true after the addition of positive and negative integral affect as independent variables to the LME model. One notable result was that the positive integral affect of the last 60 seconds had a significant positive main effect on trust in technology ($F(1, 115)=7.87$, $p<0.05$; $b=0.45$, $t(115)=2.81$, $p<0.05$).

5.2.5. Mediation analysis

Mediation analyses were conducted according to the bootstrapping approach described in section 5.1.6.3. The first test tested whether positive integral affect (mean for the last 60 seconds) mediated the relationship between incidental affect and trust in technology. A diagram of the mediation model is shown in Figure 40. As reported previously, positive incidental affect caused a higher level of positive integral affect ($a=0.75$), and positive integral affect related to a higher level of trust in technology ($b=0.45$). The calculation of bias-corrected bootstrap confidence interval (CI) indicated that the indirect effect was significant ($ab=0.37$; 95% CI: 0.10, 0.78). The coefficient of pathway c’ was not significantly different from 0 thus indicating that this was a full mediation. When the measure of positive integral affect was replaced by the peak value, the indirect effect was not significant (95% CI: -0.16, 0.01).

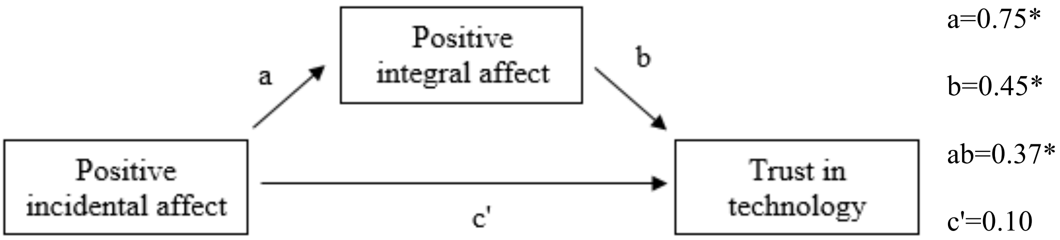


Figure 40. The diagram for the mediation relationship among positive incidental affect, positive integral affect (mean for the last 60 seconds), and trust in technology. Positive integral affect was measured at the mean level of the last 60 seconds of the task trials. a, b, ab, and c’ are standardized coefficients. * indicates the corresponding coefficient is statistically significantly different than 0 at $p<0.05$ level.

Hypothesis 2.3 was supported by the result of this mediation analysis.

The second test tested whether positive integral affect (mean for the last 60 seconds) mediated the relationship between task/technological conditions and trust in technology. A diagram of the mediation model is shown in Figure 41. Low reliability condition caused a lower level of positive integral affect (a=-0.48), and positive integral affect related to a higher level of trust in technology (b=0.45). The calculation of bias-corrected bootstrap confidence interval (CI) indicated that the indirect effect was significant (ab=-0.18; 95% CI: -0.43, -0.05). The coefficient of pathway c’ was also significant at c'=-0.33 thus indicating that this was a partial mediation.

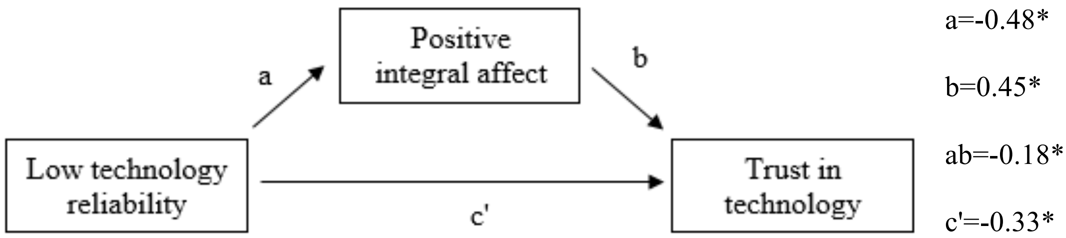


Figure 41. The diagram for the mediation relationship among low technology reliability, positive integral affect (mean for the last 60 seconds), and trust in technology. Positive integral affect was measured at the mean level of the last 60 seconds of the task trials. a, b, ab, and c' are standardized coefficients. * indicates the corresponding coefficient is statistically significantly different than 0 at $p<0.05$ level.

5.2.6. Summary of the findings

Figure 42 summarizes the relationships among incidental affect, task/technological conditions, integral affect, and trust in technology. In this figure, the integral affect was measure by the mean values of the last 60 seconds of the task. If the peak values were used, the only significant relationship was that positive incidental affect increased positive integral affect.

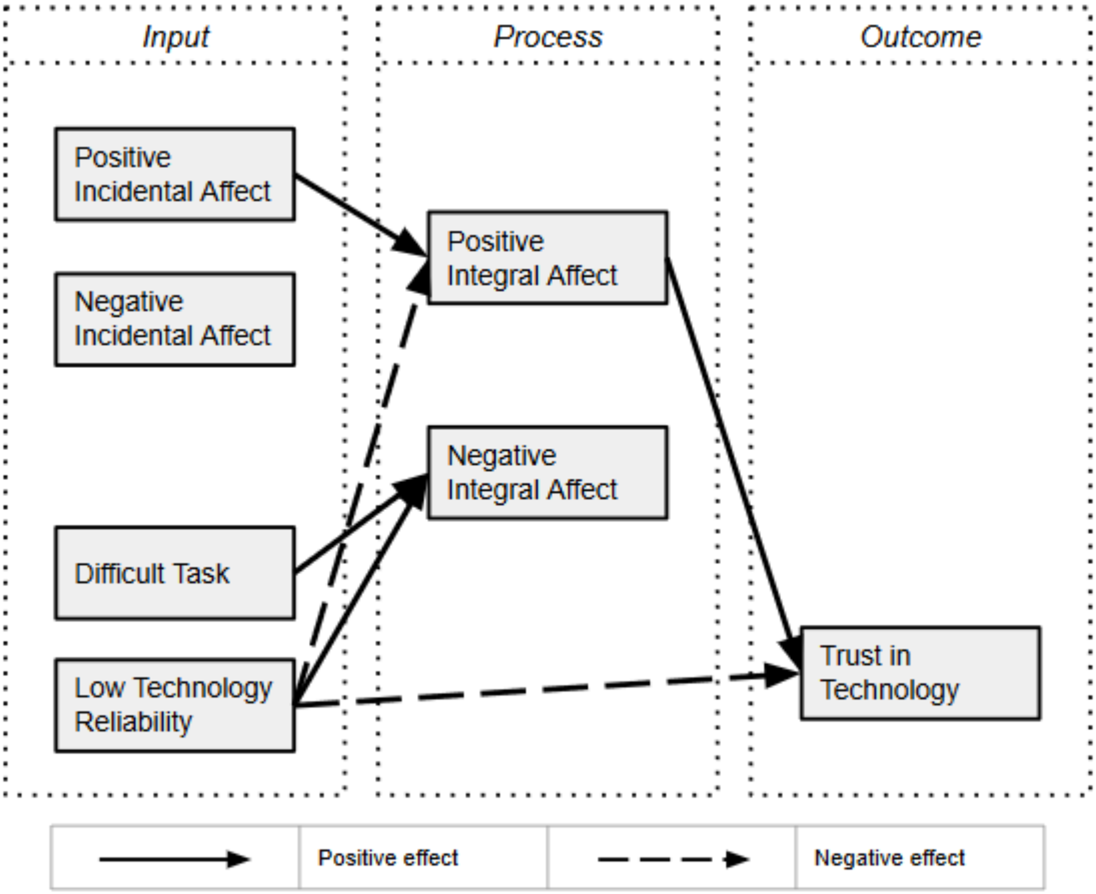


Figure 42. The relationships among incidental affect, task/technological conditions, integral affect, and trust in technology.

5.2.7. Additional analysis

5.2.7.1.Trust and distrust

Previous studies have found that the trust in technology scale used in this study consists of two distinct but related factors – trust and distrust (Safar & Turner, 2005; Spain, Bustamante, & Bliss, 2008). The effects of the independent variables on the two factors were similar to their effects on overall trust in technology score. The effects of incidental affect were not significant, while the effects of task/technological conditions were significant (for trust factor, $F(2, 100)=9.03$, $p<0.05$; for distrust factor, $F(2,100)=4.67$, $p<0.05$). The level of the trust factor was lower in the low reliability condition than that in the mean of normal and hard conditions ($b=0.53$, $t(100)=3.08$, $p<0.05$). The level of the distrust factor was higher in the low reliability condition than that in the normal condition (b=0.48, t(100)=2.08, p<0.05). Figure 43 visualizes the significant effects.

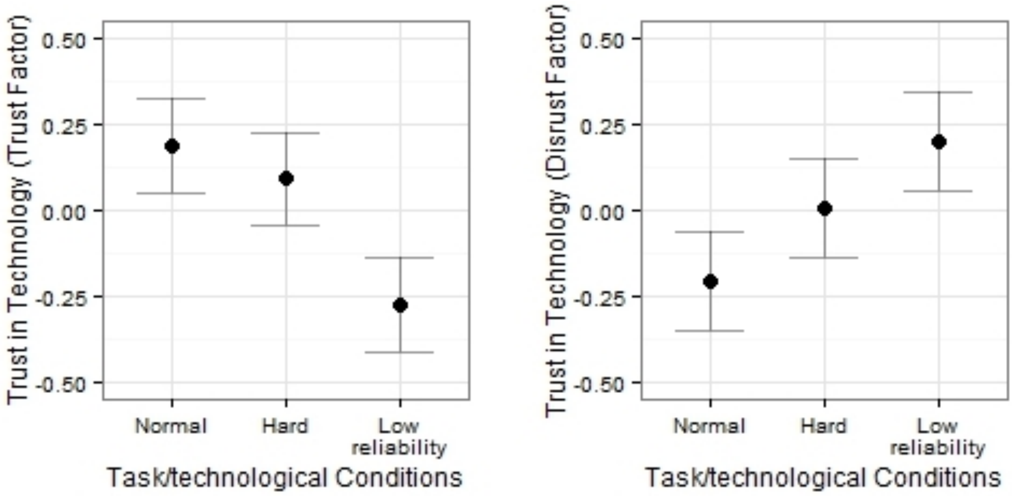


Figure 43. The mean and standard error of the predicted values of the two factors of trust in technology on different task/technological conditions. The values were standardized.

The approach used in section 5.2.4.2 was used for testing the relationship between integral affect and the two factors of trust. The results indicated that positive integral affect negatively correlated with trust factor (b=0.56, t(121)=3.15, p<0.05). Based on the additional analysis, Figure 42 is updated to use trust factor and distrust factor to replace overall trust in technology (Figure 44).

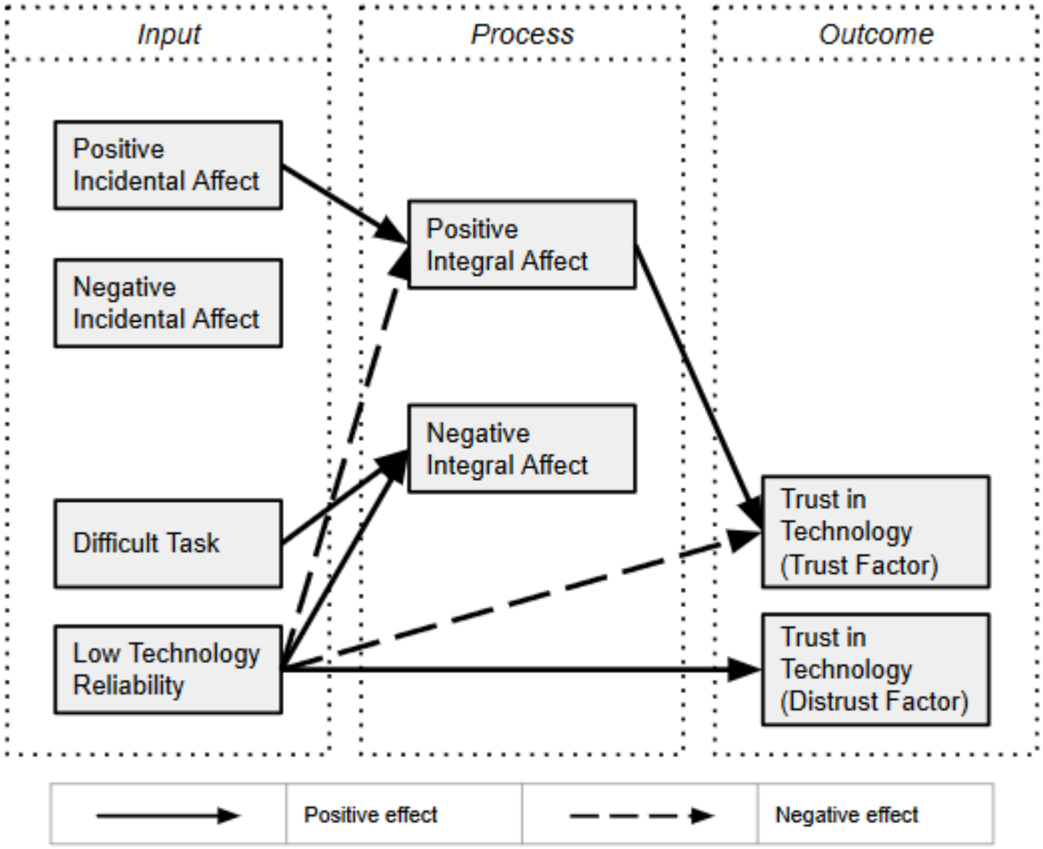


Figure 44. An updated version of Figure 42 to include trust factor and distrust factor of trust in technology.

5.2.7.2.End of task affect

5.2.7.2.1. End of task affect and integral affect

Additional analyses were performed regarding end of task affect measured using PANAS, which was administered at the end of the task. The correlations between the end of task affect and integral affect were tested in LME models controlling for the fixed effects of incidental affect, task/technological conditions, roles of the participants, and all the possible two way interactions, and random effects of teams and individuals. Two significant correlations were detected for the positive end of task affect. First, mean values of positive integral affect (last 60 seconds) was positively correlated with positive end of task affect ($b=0.16$, $t(117)=2.67$, $p<0.05$). Second, there was a negative correlation between peak values of negative integral affect with positive end of task affect ($b=0.09$, $t(104)=1.99$, $p<0.05$). However, negative end of task affect was not correlated with any of the integral affect measures.

5.2.7.2.2. Incidental affect, task/technological conditions, and end of task affect

Analyses similar to those described in section 5.2.3, 5.2.4, and 5.2.5 were performed, using end of task affect to replace integral affect. First, the effects of incidental affect and task/technological conditions on end of task affect were tested. The results showed that the effects of the independent variables were not significant on positive end of task affect (for incidental affect, $F(2,26)=0.87$, $p=0.43$; for task/technological conditions, $F(2,98)=3.00$, $p=0.05$). However, the effects on negative end of task affect were significant (for incidental affect, $F(2,24)=3.87$, $p<0.05$; for task/technological conditions, $F(2,97)=3.36$, $p<0.05$). Specifically, negative end of task affect was higher in negative incidental affect condition than the mean of positive and neutral conditions ($b=0.80$, $t(24)=2.68$, $p<0.05$). Mean negative end of task affect was higher in hard condition than the mean of normal and low reliability conditions ($b=0.18$, $t(96)=2.36$, $p<0.05$). These effects are visualized in Figure 45.

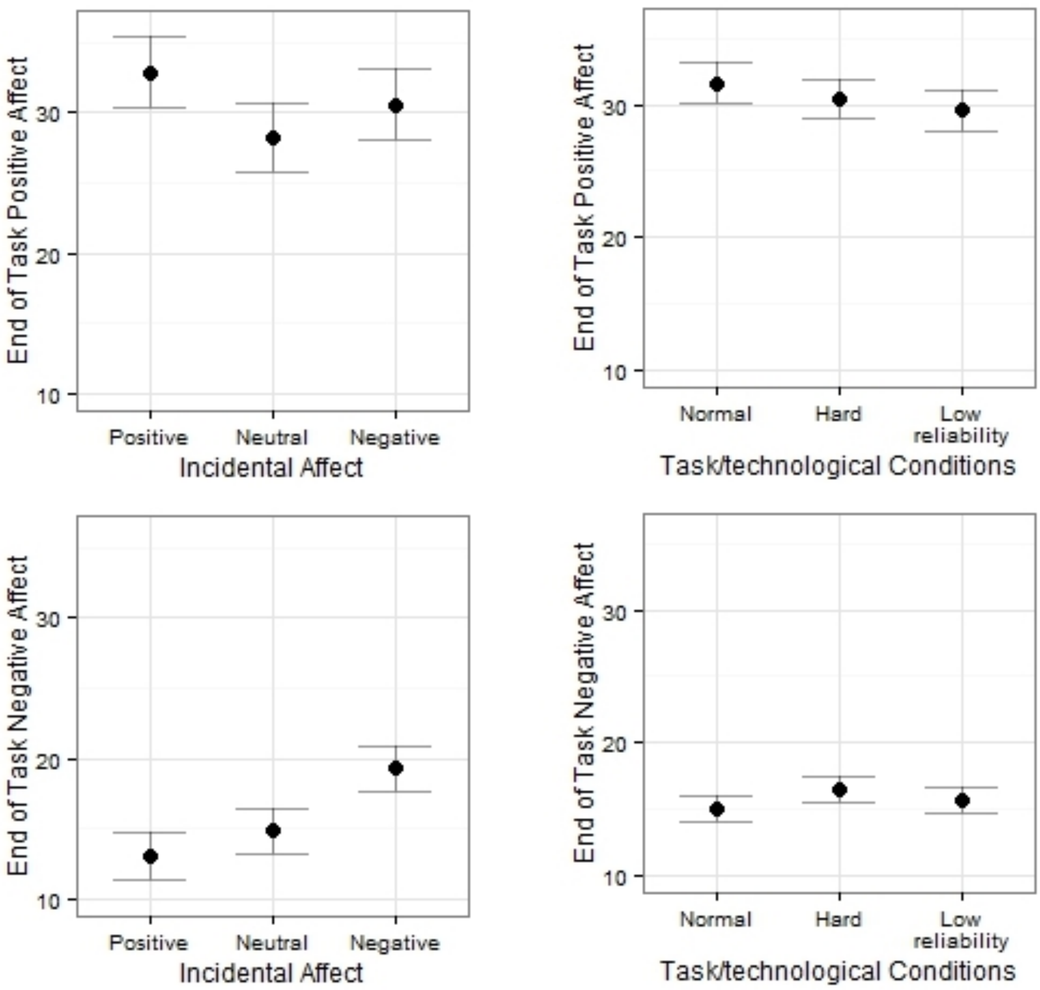


Figure 45. The mean and standard error of the predicted values of mean end of task affect on different initial mood conditions and different task/technological conditions. These plots were created given the values of all the other independent variables were held at their mean.

5.2.7.2.3. Trust in technology and end of task affect

The approach used in section 5.2.4.2 was used again for testing the relationship between end of task affect and trust in technology and the two factors of trust. Positive end of task affect had a significant positive correlation with overall trust in technology (b=2.33, t(119)=3.13, p<0.05) and trust factor (b=2.07, t(122)=3.78, p<0.05). Negative end of task affect had a significant positive correlation with distrust factor (b=1.25, t(3.122)=3.16, p<0.05). In addition to these significant correlations, there were two interesting significant interaction effects. The effect of positive end of task affect X incidental affect interaction was significant on both overall trust (F(2, 123)=6.17, p<0.05) and trust factor (F(2, 125)=7.90, p<0.05). Qualitatively, the effect could be interpreted as the positive correlation between positive end of task affect and trust disappeared under negative incidental affect condition. These effects are visualized in Figure 46.

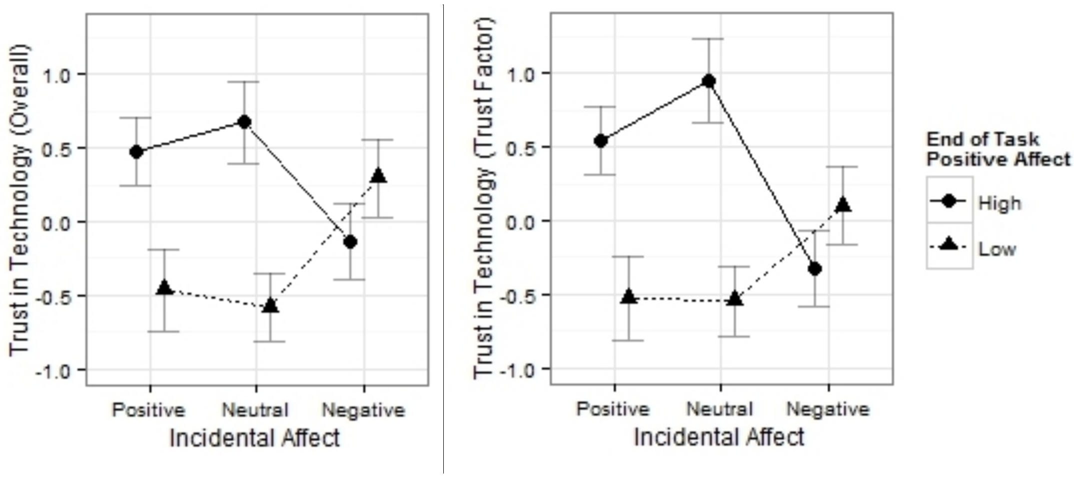


Figure 46. The mean and standard error of the predicted values of mean overall trust in technology and trust factor on different levels of positive end of task affect and initial mood conditions. The values were standardized. These plots were created given the values of all the other independent variables were held at their mean. High (low) positive end of task affect was defined as the value of +1 (-1) standard deviation off the mean.

The effect of negative end of task affect X task/technological conditions interaction was significant on both overall trust ($F(2, 96)=4.16$, $p<0.05$) and trust factor ($F(2, 97)=4.84$, $p<0.05$). Low negative end of task affect made the negative (or positive) effect of low reliability condition on overall trust (or distrust) disappeared. These effects are visualized in Figure 47.

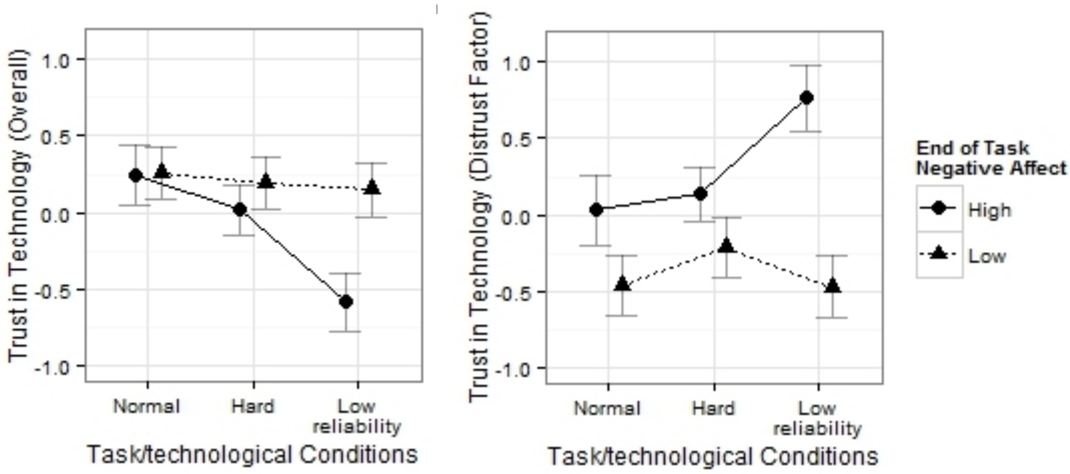


Figure 47. The mean and standard error of the predicted values of mean overall trust in technology and distrust factor on different levels of negative end of task affect and task/technological conditions. The values were standardized. These plots were created given the values of all the other independent variables were held at their mean. High (low) positive end of task affect was defined as the value of +1 (-1) standard deviation off the mean.

5.2.7.3.Summary of the relationships related to end of task affect

Figure 48 summarizes the main effects among incidental affect, task/technological conditions, end of task affect, and trust in technology. Negative incidental affect and difficult task increased negative end of task affect rating. None of the independent variables significantly predicted positive end of task affect rating. Positive affect rating positively correlated with overall trust in technology and trust factor. Negative affect rating positively correlated with distrust factor.

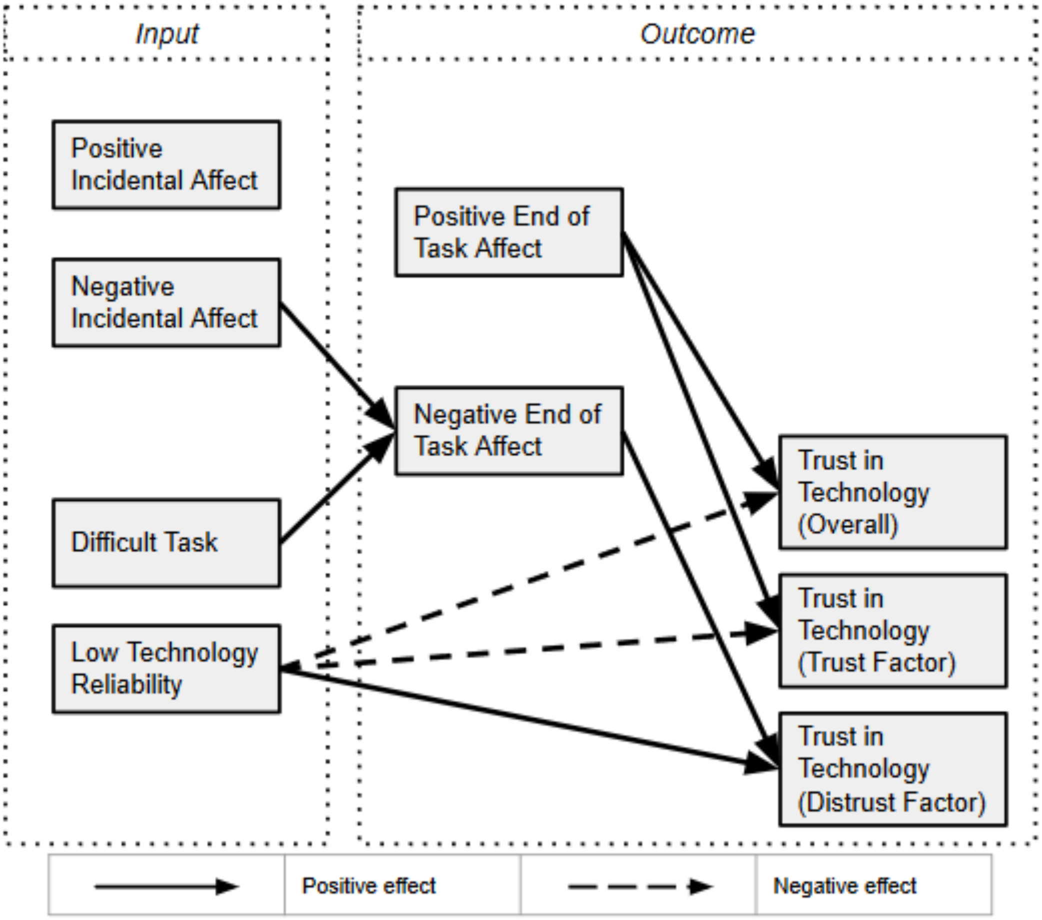


Figure 48. The main effects among incidental affect, task/technological conditions, end of task affect, and trust in technology.

Moderation effects are summarized in Figure 50. Negative incidental affect moderated the relationship between overall trust in technology and trust factor. Negative affect rating moderated the relationship between overall trust in technology and distrust factor.

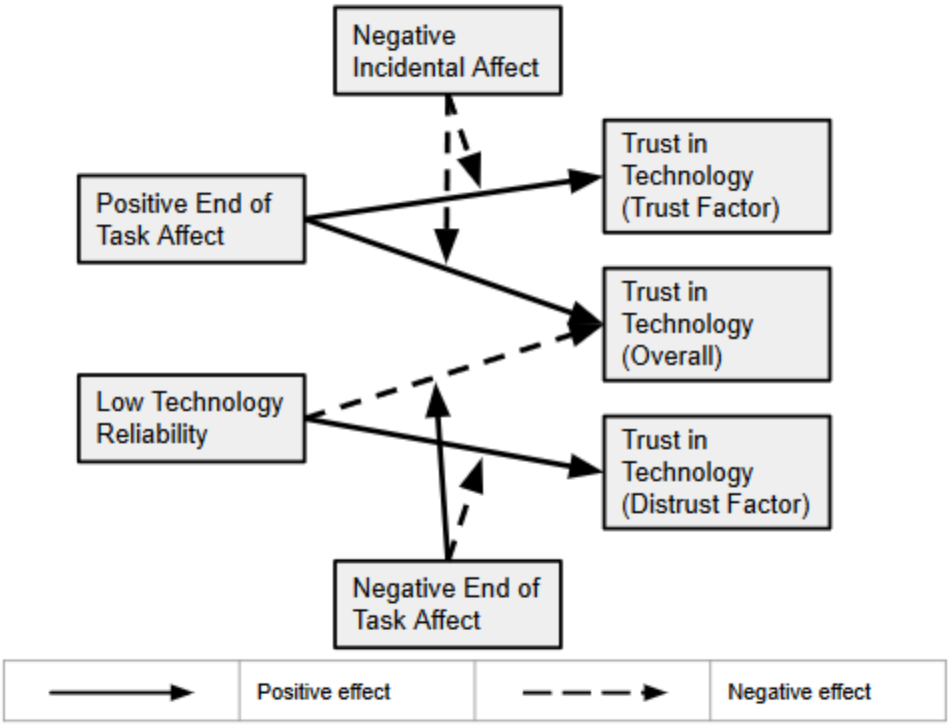


Figure 49. The moderation effects among incidental affect, task/technological conditions, end of task affect, and trust in technology.

5.2.7.4.Affect similarity within teams

The similarity values of integral affect within teams were calculated using the DTW approach used in study 1. First, multiple imputations were performed to create 5 imputed datasets. Second, the time series were all resampled at 1 Hz from the original 12 Hz. Finally, DTW technique was used to align the time series within teams to calculate similarity. To be consistent with analysis done in this study, only the last 60 seconds of the time series were included in the analysis. Mean Euclidean distances of the time series of positive/negative integral affect of the two individuals were calculated as indexes of dissimilarity. The values were than multiplied by -1 so that they could be used as indexes of similarity. All the tests performed by fitted LME models to the data and the results were pooled across 5 imputed datasets using Rubin's rule.

Tests showed that, for positive integral affect, the similarity values were higher in neutral condition than that the positive and negative conditions ($b=0.74$, $t(65)=3.03$, $p<0.05$); the similarity values were lower in normal condition than that the hard and low reliability conditions (b=0.59, t(65)=2.92, p<0.05). See Figure 50 for visualizations. For negative integral affect, higher similarity was observed in the neutral condition compared to the positive and negative conditions (b=0.65, t(65)=2.06, p<0.05). However, the effect of task/technological conditions was not significant. See Figure 51 for visualizations.

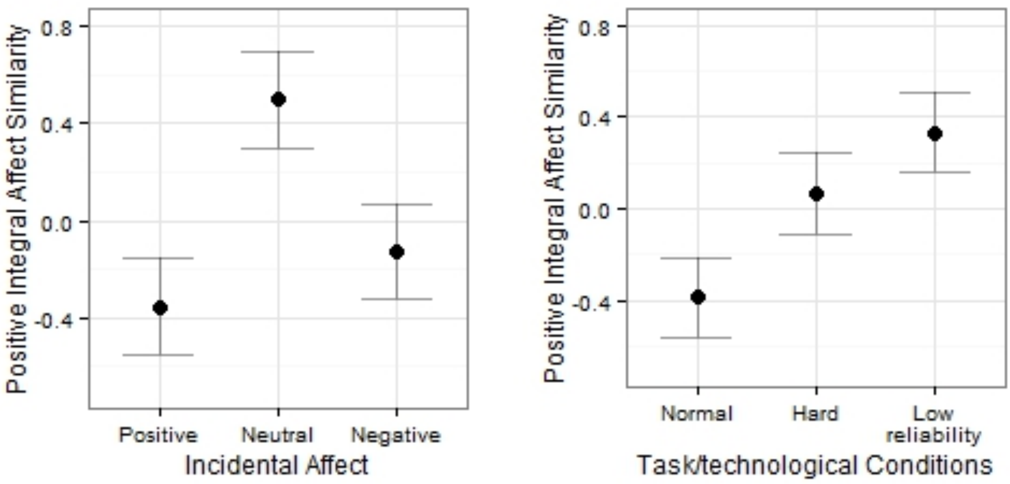


Figure 50. The mean and standard error of the predicted values of similarity of positive integral affect on different initial mood conditions and different task/technological conditions. These plots were based on the first imputation dataset and were created given the values of all the other independent variables were held at their mean.

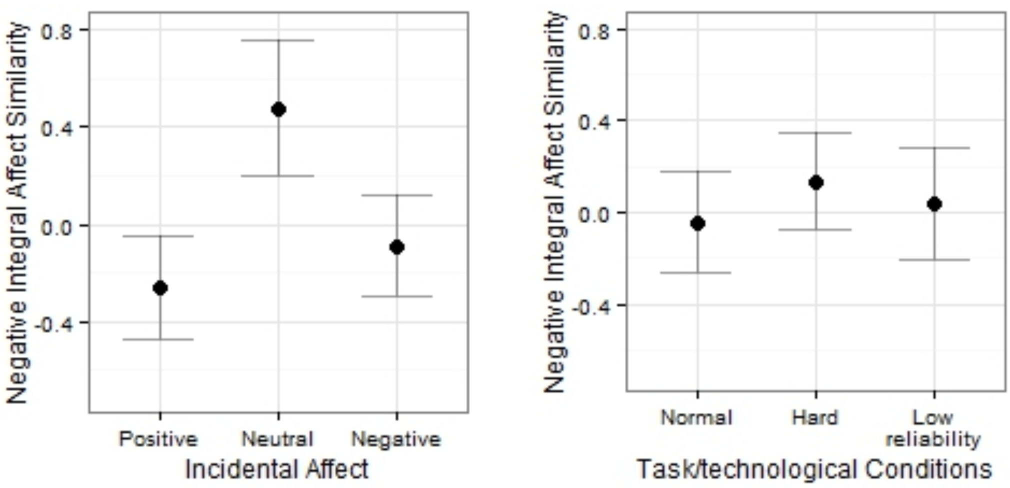


Figure 51. The mean and standard error of the predicted values of similarity of negative integral affect on different initial mood conditions and different task/technological conditions. These plots were based on the first imputation dataset and were created given the values of all the other independent variables were held at their mean.

Similarity values were calculated for overall trust in technology, trust in technology trust factor, trust in technology distrust factor, and end of task affect for individuals within a team. Tests were performed to explore if similarities of integral affect and end of task affect could predict similarities in trust. Two significant correlations were discovered: similarity of both the positive and negative integral affect significantly positively correlated with similarity of distrust (for positive affect, b=0.52, t(65)=2.92, p<0.05; for negative affect, b=0.27, t(65)=2.61, p<0.05).

A summary of the relationships found related to similarity of integral affects is depicted in Figure 52.

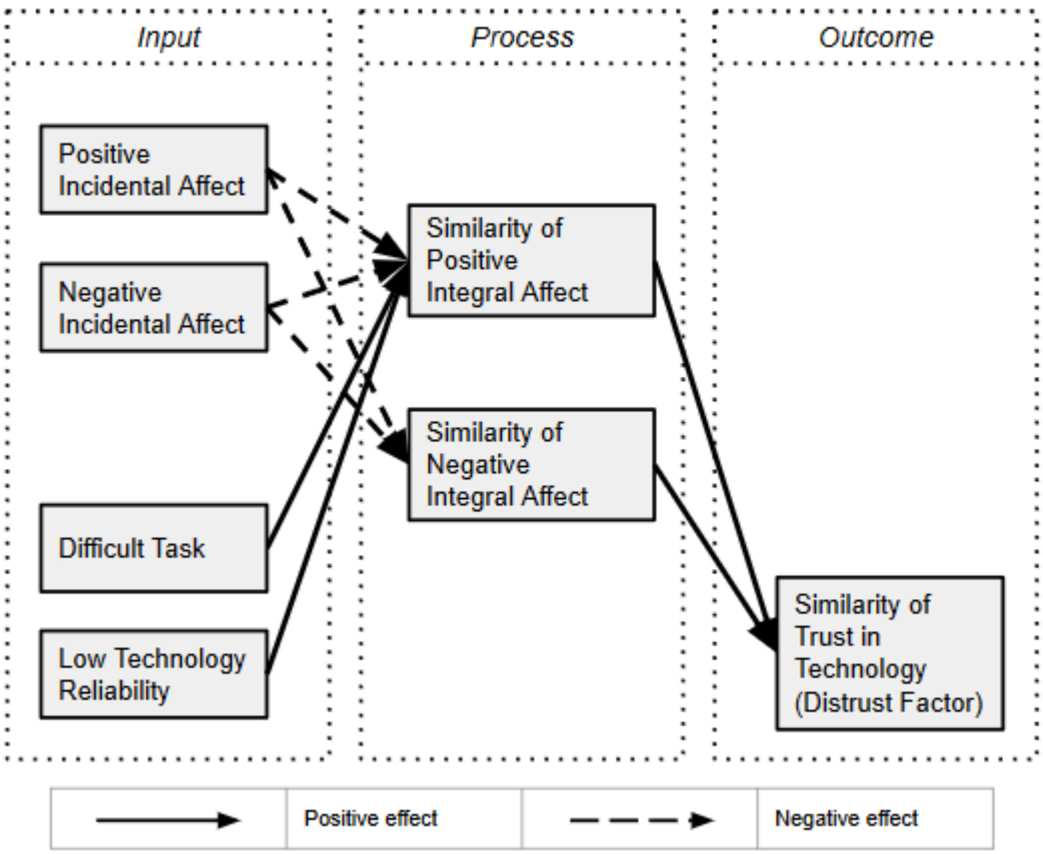


Figure 52. The main effects among incidental affect, task/technological conditions, similarity of integral affects, and trust in technology (distrust factor).

5.3. Discussion

5.3.1. Affect process as trust formation and calibration mechanism

The main contribution of this study is that it explored affective process as a mechanism of technological trust formation and calibration. By distinguishing incidental affect and integral affect, affect was treated as a “process” that continuously influences trust. Previously, affect has been recognized as a factor that influences the level of trust in technology (J. Y. Chen & Barnes, 2014; Hoff & Bashir, 2015; Merritt, 2011), however, most of the research limited the scope to incidental affect. In a recent review of trust in technology literature, Hoff and Bashir (2015) proposed that the formation and calibration of trust in technology has three layers: dispositional trust, situational trust, and learned trust. The authors stated that affect influences situational trust, which is defined as a component of trust that is influenced by the context of technology use. However, affect was not featured in the learned trust layer, which is a component of trust that is influenced by the actual interaction with the technology. This study implied that the model should be expanded to include incidental affect in situational trust, and integral affect in learned trust.

5.3.1.1.Incidental affect and trust

The second hypothesis of this study stated that there should be a main effect of incidental affect on trust in technology. However, this hypothesis was not supported. This is not consistent with previous studies (Merritt, 2011; Stokes et al., 2010). There are two possible explanations for this result. First, the duration of the effect of mood induction using IAPS might not be long enough. Before each task trial, the participants view 10 images and images were presented 10 seconds each. Although the pilot study showed that this mood induction procedure was effective right after the image viewing that it altered the reported positive and negative affect levels; its effect may be reduced after the 6-minute task. This is evident in the additional analysis where incidental affect was used to predict end of task affect rating. Incidental affect failed to significantly influence positive end of task affect. Although incidental affect significantly related to negative end of task affect, its effect size was reduced compared to the results in the pilot study (see Figure 35 and Figure 45). Future research could explore other mood induction procedures that have a longer effect duration.

The second explanation was that the link between integral affect and trust was stronger than the link between incidental affect and trust in this study. The mediation analysis showed that positive integral affect fully mediated the relationship between positive incidental affect and trust in technology. Statistically, this could happen if the effect of incidental affect on trust in technology was distal and the effect of integral affect on trust in technology was proximal (Shrout & Bolger, 2002). In such case, a larger sample size would be needed to detect the direct effect. This “distance” of the effect may be amplified by the setting of this study. Previous studies which found significant relationship between incidental affect and trust were all under settings that involved a single user interacted with a technology (Merritt, 2011; Stokes et al., 2010). However, this study was conducted in a teamwork setting, which involved face-to-face interactions between two team members. This interpersonal interaction process shaped the end of task affect and made it deviated significantly from incidental affect. This may have rendered the direct effect of incidental affect on trust in technology even more difficult to detect.

Although the main effect was not found, this study discovered an interesting moderation effect that negative incidental affect moderated the relationship between positive end of task affect and trust. Figure 46 shows the interesting pattern that it seems negative incidental affect neutralized the relationship between positive end of task affect and trust in technology. According to the affective infusion model (Forgas, 1995), an individual is more likely to use heuristic processing strategy when in a positive affective state, and the level of positive affect may be used as a piece of information that is considered to be relevant to the judgement task at hand (Clore et al., 2001). In this case, one could observe a positive correlation between positive end of task affect and level of trust. On the other hand, an individual is more likely to use substantive processing strategy when in a negative affective state. In this case, positive and negative affect may still influence decision and judgement that affective states could have a priming effect on information processing, such as attention, perception, memory retrieval, and decision selection (Niedenthal et al., 2006). However, since the individual is using more information to form the decision or judgement, the effect of affective state may be smaller than that in the heuristic processing route. In this study, the negative incidental affect could have made the participants more likely to use the substantive processing route when calibrating trust thus the effect of positive end of task affect was reduced under this condition.

5.3.1.2.Integral affect and trust

In this study, significant positive correlation was found between positive integral affect and trust in technology; however, the correlation between negative integral affect and trust was not significant. Although incidental affect instead of integral was manipulated and measured, Merritt (2011) and Stokes et al. (2010) both found that positive affect positively related to trust in technology. This is different from the research in interpersonal trust, where negative affect was found to be significantly related to trust (Dunn & Schweitzer, 2005; Ferrin et al., 2007). Although important differences in cognitive and behavioral factors that influence trust in technology and trust in human have been identified (Madhavan & Wiegmann, 2007), more study is need to understand this difference in affective factors.

The significant mediation effect of positive integral affect on technological/task condition and trust in technology showed the role of affect in the trust calibration process. Low technology reliability condition significantly decreased the level of trust in technology, and this effect was mediated by positive trust; interacting with low reliability technology decreased a user's positive integral affect, and positive integral affect positively correlated with level of trust. Note that in the mediation analysis, both the direct effect (c' in Figure 41) and the indirect effect (the multiplication of a and b in Figure 41) were significant. This indicated that affect is part of the trust calibration mechanism. However, there was a limitation in this study that it was difficult to accurately estimate the effect size of indirect effect in relation to the total effect due to the complex design.

5.3.1.3.Trust and distrust

There were controversies whether trust and distrust are opposites of the same construct or two different constructs (Saunders, Dietz, & Thornhill, 2014). Analyses of empirically defined trust scales suggested that trust and distrust are opposites (Jian et al., 2000; Montague, 2010). A functional neuroimaging (fMRI) study suggested that trust and distrust may associate with different neurological processes (Dimoka, 2010). A study on online trust found that trust had a stronger correlation with enhancing low risk internet behaviors than distrust did with lowing it; while distrust had a stronger correlation with lowering high risk internet behavior than trust did with enhancing it (Chang & Fang, 2013). In this study, it was found that low technology reliability decreased trust and increased distrust. However, it was also found that positive integral affect and positive end of task affect positively correlated with trust but not distrust, while negative end of task affect positively correlated with distrust but not trust. These results suggested that trust and distrust may involve different affective processes. The limitation of this study was that it was not designed to further explore how trust and distrust in technology influence behavioral outcomes in active and passive user system. Future research could investigate if trust and distrust in shared technology would influence communication, collaboration, or technology use.

5.3.1.4.Affective process, layers of trust, and design

Given the results of this study, An updated version of Hoff and Bashir's (2015) three-layered trust model is proposed (see Figure 53).

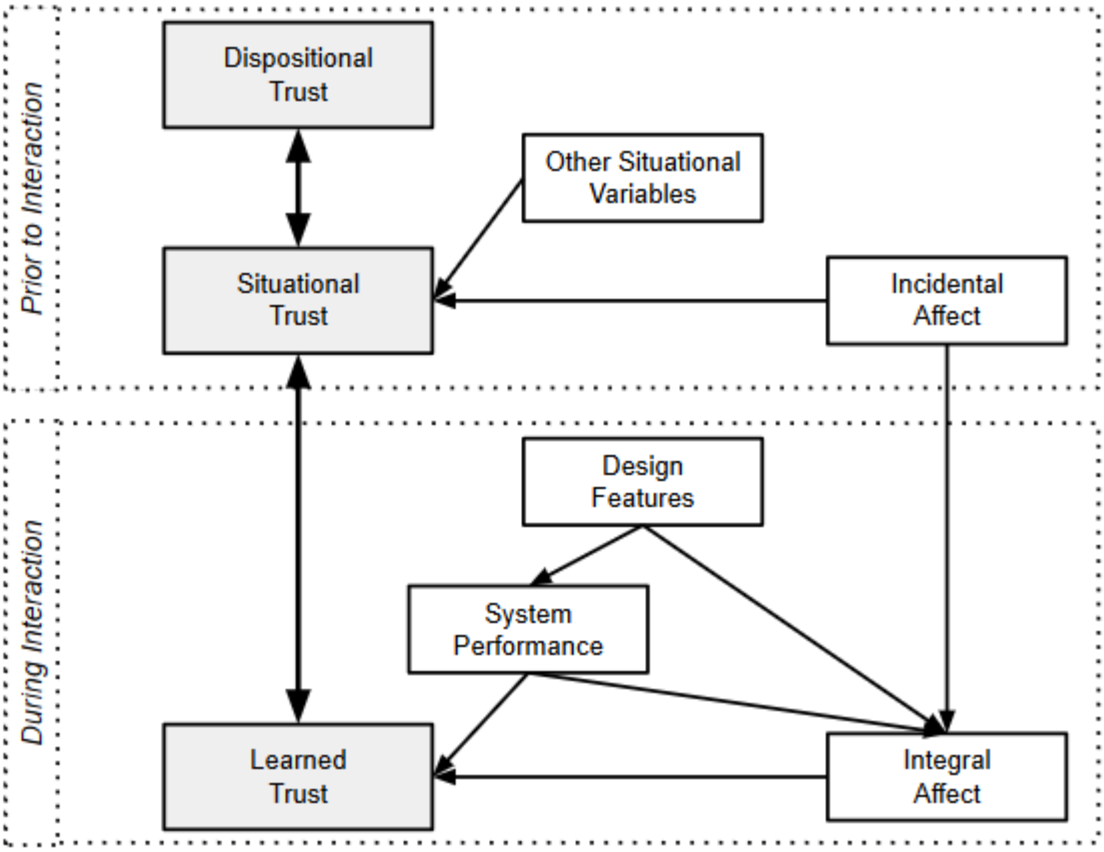


Figure 53. An updated three-layered trust model (Hoff & Bashir, 2015) with the addition of incidental affect and integral affect.

In this model, the main structure remains the same as the original three-layered trust model. Incidental affect is added as a factor to influence situational trust to replace "mood" in the original model. Same as "mood" in the original model, incidental affect is not directly related to the interaction with the technology – it is the affective state that a user "incidentally" experienced before the interaction. Integral affect is added as a factor to influence learned trust during the interaction with the technology. Integral affect is influenced by both design features of the technology (e.g., appearance, usability, level of control) and system performance (e.g., reliability).

In addition to theoretical implications, the findings also have practical implications for the design of systems that facilitate calibration for appropriate trust in technology. For example, to improve initial trust when a technology is first introduced, one could induce positive mood as incidental affect to facilitate the formation of higher level of trust (Merritt, 2011). At the same time, the use of appropriate emotional design, such as interfaces with humanoid features (Swangnetr, Zhu, Taylor, & Kaber, 2010), could induce positive integral affect to improve trust. Design features that facilitate positive interaction among team members may also promote trust in technology. Another point to note is that negative incidental affect might not be always a bad thing for trust calibration. As the results suggested, it might cause the users more likely to use substantive process strategy for trust calibration thus reducing the "unwanted" effect of integral affect on trust calibration. For example, this could help prevent over-trust caused by high level of positive affect when the task happened to be easier than usual when the technology is first introduced.

5.3.2. Integral affect and end of task affect

In the additional analysis, end of task affect replaced integral affect and the same procedure was applied to analyze the data. The analyses with integral affect and end of task affect showed notable similarities and differences. First, related to the effect incidental affect, positive incidental affect positively related to positive integral affect but not positive end of task affect; negative incidental affect positive related to negative end of task affect but not negative incidental affect. Second, related to the effect of task/technological conditions, difficult task increased both negative integral affect and negative end of task affect. Low reliability decreased positive integral affect and increased negative integral affect, however, it had no effect on end of task affect. Finally, related to trust in technology, both positive integral affect and positive end of task affect positively related to trust but not distrust. Distrust positively related to negative end of task affect but not negative integral affect.

A limitation of this study should be acknowledged: integral affect and end of task affect were measured using different instruments. Integral affect was measured using facial expression analysis, while end of task affect was measured using self-report rating scale. A literature review of affect measures indicated that the correlations among different measures of affect are moderate at best, small in typical studies, and inconsistent across studies" (Mauss & Robinson, 2009). According to the components processing perspective of defining emotion (K. R. Scherer, 2005), emotion is considered to be consisted of a number of distinct components, such as cognitive appraisal, physiological responses, action tendencies, facial and vocal expressions, and subjective feelings. However, these components may not be cohere (Niedenthal et al., 2006). As a result, theoretically affect cannot be reliably measured with one measurement instrument alone. In the context of this study, using facial expression to measure integral affect and self-report scale to measure end of task affect may be the optimum approach given the design and resource constrains; however, it is unknown that the similarities and differences in the results were an effect of the constructs themselves or the components that the instruments were measuring.

While this measurement issue seems to be difficult to resolve in the short term, there is a future direction that could be perused for this particular research. For integral affect, alterative measurement could be developed. In this study, the measure of integral affect was derived from facial expression recognition. Since the output of facial expression recognition was time series data, a number of different values could be calculated to represent the time series. In the analysis, four ways of deriving affective values were explored: mean of first 60 seconds, mean of last 60 seconds, mean of full 360 seconds, and peak value. For the mean values, only mean of the last 60 seconds were included for its proximity to the point when the participants were asked to evaluate trust. There are other ways to describe the time series data, for example, standard deviation can represent the variation in expressions; trend analysis can be applied to the time series to identify trends in the data; for positive affect, the frequency of smiles can be calculated. In addition, as previously discussed, a set of different algorithm can be applied to calculate positive and negative affect directly instead of pooling intensity values of basic emotions.

5.3.3. Additional future direction

A possible future direction is to investigate mechanisms of how affective process influence trust in technology using prominence-interpretation theory (Fogg, 2003). According to prominence-interpretation theory, the process of trust formation calibration involves two elements, namely prominence and interpretation. Prominence refers to the likelihood of a specific system element being perceived by a user. Interpretation refers to how a user evaluates the system element in terms of trust. The overall trust of the user towards the technology is the combined effect of the elements that are perceived by the user and the user's corresponding evaluation of the system factors. The prominence and interpretation are related to subjective perceptions of the user about the technology, and these perceptions are influenced by objective factors, such as the task being performed, user expertise, and individual differences, etc. (Fogg, 2003).

A previous study with a similar research design as the current study used qualitative methods to identify a list of factors that influenced a user's trust (Xu et al., 2014). These factors included technology factor (usability, competence, and appearance), user factor (confidence and personality), and task factor (task demand and potential outcome). The list can also be expanded by reviewing previous literature (e.g., Hancock et al., 2011; Hoff & Bashir, 2015; J. D. Lee & See, 2004; Madhavan & Wiegmann, 2007). The collection of these system elements that could potentially affect a user's trust in technology becomes a pool of potential antecedents of trust in technology. These system elements have to be perceived and evaluated by the user in order to have impact on the user's trust. Affect can be an important factor to guide the perception and interpretation of such factors; See Figure 54. Specifically, affective states can change the probability of certain system elements to be perceived and influence how the perceived elements to be interpreted.

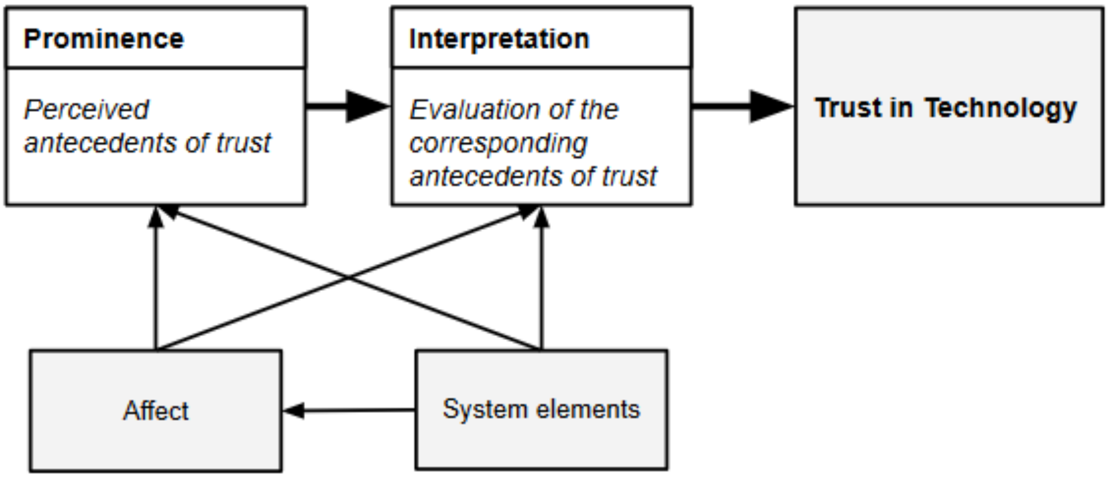


Figure 54. A proposed model of affect process in the formation and calibration of trust in technology.

References

Ackerman, M. S. (2000). The intellectual challenge of CSCW: The gap between social requirements and technical feasibility. *Human–Computer Interaction, 15*(2), 179-203.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin, 84*(5), 888.

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of personality and social psychology, 84*(5), 1054.

Anvari, M., McKinley, C., & Stein, H. (2005). Establishment of the world's first telerobotic remote surgical service: for provision of advanced laparoscopic surgery in a rural community. *Annals of surgery, 241*(3), 460.

Asan, O., & Montague, E. (2013). Technology-mediated information sharing between patients and clinicians in primary care encounters. *Behaviour & Information Technology*, 1-12.

Asan, O., Montague, E., & Xu, J. (2012). Assessing Patient and Doctor Eye Gaze Patterns Between Two Styles of Doctor Ehr Use in Primary Care Encounters. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 951-955.

Bafoutsou, G., & Mentzas, G. (2002). Review and functional classification of collaborative systems. *International Journal of Information Management, 22*(4), 281-305.

Barnard, J., & Rubin, D. B. (1999). Miscellanea. Small-sample degrees of freedom with multiple imputation. *Biometrika, 86*(4), 948-955.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology, 51*(6), 1173.

Barrett, L. F., & Russell, J. A. (1999). The Structure of Current Affect Controversies and Emerging Consensus. *Current Directions in Psychological Science, 8*(1), 10-14.

Barsade, S. G. (2000). The ripple effect: Emotional contagion in groups.

Barsade, S. G., & Gibson, D. E. (1998). Group emotion: A view from top and bottom. In D.Gruenfeld, E. Mannix & M. Neale (Eds.), *Research on managing groups and teams* (pp. 81-102). Stamford, CT: JAI Press.

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

Barsade, S. G., & Gibson, D. E. (2012). Group affect its influence on individual and group outcomes. *Current Directions in Psychological Science, 21*(2), 119-123.

Barsade, S. G., & Knight, A. P. (2015). Group affect. *Annu. Rev. Organ. Psychol. Organ. Behav., 2*(1), 21-46.

Bartel, C. A., & Saavedra, R. (2000). The collective construction of work group moods. *Administrative Science Quarterly, 45*(2), 197-231.

Bartlett, M. S., Littlewort, G. C., Frank, M. G., Lainscsek, C., Fasel, I. R., & Movellan, J. R. (2006). Automatic recognition of facial actions in spontaneous expressions. *Journal of Multimedia, 1*(6), 22-35.

Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bates, D. M., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using S4 classes.

Bensing, J. (1991). Doctor-patient communication and the quality of care. *Social Science & Medicine, 32*(11), 1301-1310.

Berndt, D. J., & Clifford, J. (1994). *Using Dynamic Time Warping to Find Patterns in Time Series.* Paper presented at the KDD workshop.

Biel, J.-I., Teijeiro-Mosquera, L., & Gatica-Perez, D. (2012). *Facetube: predicting* *personality from facial expressions of emotion in online conversational video.* Paper presented at the Proceedings of the 14th ACM international conference on Multimodal interaction.

Biel, J.-I., Tsiminaki, V., Dines, J., & Gatica-Perez, D. (2013). *Hi youtube!: Personality impressions and verbal content in social video.* Paper presented at the Proceedings of the 15th ACM on International conference on multimodal interaction.

Bisantz, A. M., & Seong, Y. (2001). Assessment of operator trust in and utilization of automated decision-aids under different framing conditions. *International Journal of Industrial Ergonomics, 28*(2), 85-97.

Blickensderfer, E., Cannon-Bowers, J. A., & Salas, E. (1998). Cross-training and team performance. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 299-311). Washington, DC: American Psychological Association.

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological methodology, 20*(1), 15-140.

Bosch, N., Chen, Y., & D'Mello, S. (2014). *It's written on your face: detecting affective states from facial expressions while learning computer programming.* Paper presented at the Intelligent Tutoring Systems.

Bosch, N., D'Mello, S., Baker, R., Ocumpaugh, J., Shute, V., Ventura, M., et al. (2015). *Automatic Detection of Learning-Centered Affective States in the Wild.* Paper presented at the Proceedings of the 2015 International Conference on Intelligent User Interfaces (IUI 2015). ACM, New York, NY, USA.

Boucsein, W., & Backs, R. W. (2000). Engineering psychophysiology as a discipline: Historical and theoretical aspects. In R. W. Backs & W. Boucsein (Eds.), *Engineering psychophysiology. Issues and applications* (pp. 3-30). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Bower, G. H. (1981). Mood and memory. *American psychologist, 36*(2), 129.

Bowers, C. A., Oser, R. L., Salas, E., & Cannon-Bowers, J. A. (1996). Team performance in automated systems. *Automation and human performance: Theory and applications*, 243-263.

Bradley, M. M., & Lang, P. J. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 29-46). New York: Oxford University Press, Inc.

Brave, S., & Nass, C. (2003). Emotion In Human-computer Interaction. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applicationsis* (pp. 81-96). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies, 62*(2), 161-178.

Butler, E. A. (2011). Temporal Interpersonal Emotion Systems The "TIES" That Form Relationships. *Personality and Social Psychology Review, 15*(4), 367-393.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 173-191). New York: Guildford Press.

Cadiere, G. B., Himpens, J., Vertruyen, M., & Favretti, F. (1999). The world's first obesity surgery performed by a surgeon at a distance. *Obesity Surgery, 9*(2), 206-209.

Carayon, P., Hundt, A. S., Karsh, B., Gurses, A., Alvarado, C., Smith, M., et al. (2006). Work system design for patient safety: the SEIPS model. *Quality and Safety in Health Care,* *15*(suppl 1), i50-i58.

Carr, L., Iacoboni, M., Dubeau, M., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, 100*(9), 5497-5502.

Chang, Y.-S., & Fang, S.-R. (2013). Antecedents and distinctions between online trust and distrust: Predicting high-and low-risk Internet behaviors. *Journal of Electronic Commerce Research, 14*(2), 149-166.

Chen, J. Y., & Barnes, M. J. (2014). Human–Agent Teaming for Multirobot Control: A Review of Human Factors Issues. *Human-Machine Systems, IEEE Transactions on, 44*(1), 13-29.

Chen, Y., Bosch, N., & D'Mello, S. (2015). Video-Based Affect Detection in Noninteractive Learning Environments.

Clegg, C. W. (1993). Social systems that marginalize the psychological and organizational aspects of information technology. *Behaviour & Information Technology, 12*(5), 261-266.

Clegg, C. W. (2000). Sociotechnical principles for system design. *Applied ergonomics, 31*(5), 463-477.

Clore, G. L., Wyer, R. S., Dienes, B., Gasper, K., Gohm, C., & Isbell, L. (2001). Affective feelings as feedback: Some cognitive consequences. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 27-62). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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. 106-123). New York: Oxford University Press, Inc.

Cole, M. S., Walter, F., & Bruch, H. (2008). Affective mechanisms linking dysfunctional behavior to performance in work teams: A moderated mediation study. *Journal of Applied Psychology, 93*(5), 945-958.

Cole, M. S., Walter, F., & Bruch, H. (2008). Affective mechanisms linking dysfunctional behavior to performance in work teams: a moderated mediation study. *Journal of Applied Psychology, 93*(5), 945.

Collins, A. L., Lawrence, S. A., Troth, A. C., & Jordan, P. J. (2013). Group affective tone: A review and future research directions. *Journal of Organizational Behavior, 34*(S1), S43-S62.

Comstock, J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research* (No. 104174). Hampton, VA: National Aeronautics and Space Administration.

Cook, W. L., & Kenny, D. A. (2005). The Actor–Partner Interdependence Model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development, 29*(2), 101-109.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies, 58*(6), 737-758.

Corsini, R. J. (1956). Understanding and similarity in marriage. *The Journal of Abnormal and Social Psychology, 52*(3), 327.

D'Mello, S., & Calvo, R. A. (2013). *Beyond the basic emotions: what should affective computing compute?* Paper presented at the CHI'13 Extended Abstracts on Human Factors in Computing Systems.

Delbanco, T. L. (1992). Enriching the Doctor-Patient Relationship by Inviting the Patient's Perspective. *Annals of Internal Medicine, 116*(5), 414-418.

Deriso, D., Susskind, J., Krieger, L., & Bartlett, M. (2012). *Emotion mirror: a novel intervention for autism based on real-time expression recognition.* Paper presented at the Computer Vision–ECCV 2012. Workshops and Demonstrations.

Deriso, D., Susskind, J., Tanaka, J., Winkielman, P., Herrington, J., Schultz, R., et al. (2012). *Exploring the facial expression perception-production link using real-time automated facial expression recognition.* Paper presented at the Computer Vision–ECCV 2012. Workshops and Demonstrations.

Desanctis, G., & Gallupe, R. B. (1987). A foundation for the study of group decision support systems. *Management Science, 33*(5), 589-609.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science, 11*(1), 86-89.

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *Mis Quarterly, 34*(2), 373-396.

Donkor, R., Burnett, G., & Sharples, S. (2014). Measuring the emotional validity of driving simulators. *Advances in Transportation Studies, 1*(SPECIAL ISSUE).

Dunn, J. R., & Schweitzer, M. E. (2005). Feeling and believing: the influence of emotion on trust. *Journal of personality and social psychology, 88*(5), 736.

Dzindolet, M. T., Spencer, M., Malcolm, L., Wigley, B., Glazer, C., & Pierce, L. (2010). *The Effect of Affective State on Virtual and Face-to-Face Group Performance.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Eckhardt, M., Fasel, I., & Movellan, J. (2009). Towards practical facial feature detection. *International Journal of Pattern Recognition and Artificial Intelligence, 23*(03), 379-400.

Eich, E., & Metcalfe, J. (1989). Mood dependent memory for internal versus external events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(3), 443.

Eich, E., Ng, J. T. W., Macaulay, D. P. A. D., Percy, A. D., & Grebneva, I. (2007). Combining music with thought to change mood. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 124-136). New York: Oxford University Press, Inc.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement.*

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system*. Salt Lake City, UT: A Human Face.

Elkins, A. N., Muth, E. R., Hoover, A. W., Walker, A. D., Carpenter, T. L., & Switzer, F. S. (2009). Physiological compliance and team performance. *Applied ergonomics, 40*(6), 997-1003.

Ellard, K. K., Farchione, T. J., & Barlow, D. H. (2012). Relative Effectiveness of Emotion Induction Procedures and the Role of Personal Relevance in a Clinical Sample: A Comparison of Film, Images, and Music. *Journal of Psychopathology and Behavioral Assessment, 34*(2), 232-243.

Ellsworth, P. C. (1994). Some reasons to expect universal antecedents of emotion. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 193-210). New York: John Wiley & Sons.

Endsley, M. R., & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*, 381-394.

Escoto, K. H., Karsh, B. T., & Beasley, J. W. (2006). Multiple user considerations and their implications in medical error reporting system design. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(1), 48.

Espevik, R., Johnsen, B. H., & Eid, J. (2011). Outcomes of shared mental models of team members in cross training and high-intensity simulations. *Journal of Cognitive*

Engineering and Decision Making, 5(4), 352-377.

Ferrin, D. L., Bligh, M. C., & Kohles, J. C. (2007). Can I trust you to trust me? A theory of trust, monitoring, and cooperation in interpersonal and intergroup relationships. *Group & Organization Management, 32*(4), 465-499.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fiedler, K., Asbeck, J., & Nickel, S. (1991). Mood and constructive memory effects on social judgement. *Cognition & Emotion, 5*(5-6), 363-378.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. *CHI'03 extended abstracts on Human factors in computing systems*, 722-723.

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., & Tauber, E. R. (2003). *How do users evaluate the credibility of Web sites?: a study with over 2,500 participants.* Paper presented at the Proceedings of the 2003 conference on Designing for user experiences, San Francisco, CA.

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychological bulletin, 117*(1), 39.

Forgas, J. P. (2001). The affect infusion model (AIM): An integrative theory of mood effects on cognition and judgments. In L. L. Martin & G. L. Clore (Eds.), *Theories of mood and cognition: A user's guidebook* (pp. 99-134). Mahwah, NJ: Erlbraum.

Frijda, N. H. (1994). Varieties of affect: Emotions and episodes, moods, and sentiments. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotions: Fundamental questions* (pp. 197-202). New York: Oxford University Press.

Frijda, N. H. (2008). The psychologists' point of view. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 68-87). New York: The Guilford Press.

Fritz, M. S., & MacKinnon, D. P. (2007). Required sample size to detect the mediated effect. *Psychological science, 18*(3), 233-239.

Garrison, G. M., Bernard, M. E., & Rasmussen, N. H. (2002). 21st-century health care: the effect of computer use by physicians on patient satisfaction at a family medicine clinic. *Family Medicine, 34*(5), 362-368.

Gerrards ‐ Hesse, A., Spies, K., & Hesse, F. W. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British journal of psychology, 85*(1), 55-78.

Gibson, C. B. (2003). The Efficacy Advantage: Factors Related to the Formation of Group Efficacy. *Journal of Applied Social Psychology, 33*(10), 2153-2186.

Gino, F., & Schweitzer, M. E. (2008). Blinded by anger or feeling the love: how emotions influence advice taking. *Journal of Applied Psychology, 93*(5), 1165.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of statistical Software, 31*(7), 1-24.

Golembiewski, R. T., & McConkie, M. (1975). The centrality of interpersonal trust in group processes. In C. L. Cooper (Ed.), *Theories of group processes* (pp. 131-185). New York: Wiley.

Gottman, J. M., & Levenson, R. W. (1985). A valid procedure for obtaining self-report of affect in marital interaction. *Journal of consulting and clinical psychology, 53*(2), 151.

Gregory, A. H., Worrall, L., & Sarge, A. (1996). The development of emotional responses to music in young children. *Motivation and Emotion, 20*(4), 341-348.

Gross, J. J. (1999). Emotion and emotion regulation. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 525-552). New York: Guildford.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion, 9*(1), 87-108.

Grudin, J. (1994). Computer-supported cooperative work: History and focus. *Computer, 27*(5), 19-26.

Grudin, J., & Poltrock, S. (2012). Taxonomy and theory in computer supported cooperative work. In S. W. J. Kozlowski (Ed.), *The Oxford Handbook of Industrial and Organizational Psychology* (Vol. 2, pp. 1323-1348). New York: Oxford University Press.

Gunes, H., & Schuller, B. (2012). Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*.

Hackman, J. R., & Oldham, G. R. (1975). Development of the job diagnostic survey. *Journal of Applied Psychology, 60*(2), 159.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational behavior and human performance, 16*(2), 250-279.

Hall, I., & Higgins, S. (2005). Primary school students' perceptions of interactive whiteboards. *Journal of Computer Assisted Learning, 21*(2), 102-117.

Hancock, J. T., Gee, K., Ciaccio, K., & Lin, J. M.-H. (2008). *I'm sad you're sad: emotional contagion in CMC.* Paper presented at the Proceedings of the 2008 ACM conference on Computer supported cooperative work.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(5), 517-527.

Hancock, P. A., Pepe, A. A., & Murphy, L. L. (2005). Hedonomics: The Power of Positive and. *Ergonomics in Design, 13*(1), 8-14.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload, 1*, 139–183.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1992). Primitive emotional contagion. *Review of personality and social psychology, 14*, 151-177.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Science, 2*(3), 96-99.

Hatfield, E., Carpenter, M., & Rapson, R. L. (2014). Emotional contagion as a precursor to collective emotions. In C. von Scheve & M. Salmela (Eds.), *Collective Emotions* (pp. 108-124). Oxford, UK: Oxford University Press.

Helton, W. S., & Russell, P. N. (2011). The effects of arousing negative and neutral picture stimuli on target detection in a vigilance task. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(2), 132-141.

Hendrick, H. W. (2002). Macroergonomics Methods: Assessing Work System Structure. In H. W. Hendrick & B. M. Kleiner (Eds.), *Macroergonomics: Theory, Methods, and Applications* (pp. 45-66). Mahwah, NJ: Lawrence Erlbaum Associates.

Henning, R. A., Armstead, A. G., & Ferris, J. K. (2009). Social psychophysiological compliance in a four-person research team. *Applied ergonomics, 40*(6), 1004-1010.

Henning, R. A., Boucsein, W., & Gil, M. C. (2001). Social-physiological compliance as a determinant of team performance. *International journal of psychophysiology, 40*(3), 221-232.

Hoff, K. A., & Bashir, M. (2015). Trust in Automation Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and*

Ergonomics Society, 57(3), 407-434.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI), 7*(2), 174-196.

Honaker, J., & King, G. (2010). What to do about missing values in time‐series cross‐section data. *American Journal of Political Science, 54*(2), 561-581.

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software, 45*(7), 1-47.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science, 19*(3), 265-288.

iMotions. (2013). FACET Module Guide.

Inbar, O., & Tractinsky, N. (2009). The incidental user. *interactions, 16*(4), 56-59.

Inbar, O., & Tractinsky, N. (2010). *Interface-to-face: sharing information with customers in service encounters.* Paper presented at the Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems, Atlanta, Georgia, USA.

Inbar, O., & Tractinsky, N. (2012). Lowering the line of visibility: incidental users in service encounters. *Behaviour & Information Technology, 31*(3), 245-260.

Inkpen, K., Mcgrenere, J., Booth, K., & Klawe, M. (1997). *The effect of turn-taking protocols on children's learning in mouse-driven collaborative environments.* Paper presented at the Proceedings of the conference on Graphics interface '97, Kelowna, British Columbia, Canada.

Itoh, K., Andersen, H. B., Madsen, M. D., Østergaard, D., & Ikeno, M. (2006). Patient views of adverse events: Comparisons of self-reported healthcare staff attitudes with disclosure of accident information. *Applied ergonomics, 37*(4), 513-523.

Izard, C. E. (2009). Emotion theory and research: Highlights, unanswered questions, and emerging issues. *Annual Review of Psychology, 60*, 1.

Jentsch, F., Barnett, J., Bowers, C. A., & Sales, E. (1999). Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors, 41*(1), 1-14.

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53-71.

Johansen, R., Sibbet, D., Benson, S., Martin, A., Mittman, R., & Saffo, P. (1991). *Leading business teams: How teams can use technology and group process tools to enhance performance*: Addison-Wesley Longman Publishing Co., Inc.

Johansson, G. (1989). Job demands and stress reactions in repetitive and uneventful monotony at work. *International journal of health services: planning, administration, evaluation, 19*(2), 365-377.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102-138). New York: The Guilford Press.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Macmillan.

Kaplan, S. H., Greenfield, S., & Ware, J. E. J. (1989). Assessing the effects of physician-patient interactions on the outcomes of chronic disease. *Medical Care, 27*(3), 110-127.

Karsh, B. (2006). *Meso-ergonomics: a new paradigm for macroergonomics research.* Paper presented at the Proceedings of the International Ergonomics Association 2006 Congress, Maastricht, Netherlands.

Karsh, B., Waterson, P., & Holden, R. J. (2014). Crossing levels in systems ergonomics: A framework to support 'mesoergonomic' inquiry. *Applied Ergonomics, 45*(1), 45-54.

Kelly, J. R. (2001). Mood and emotion in groups. In M. A. Hogg & S. Tindale (Eds.), *Blackwell handbook of social psychology: Group processes* (pp. 164-181). Malden, MA: Blackwell Publishers Inc.

Kelly, J. R., & Barsade, S. G. (2001). Mood and emotions in small groups and work teams. *Organizational Behavior and Human Decision Processes, 86*(1), 99-130.

Kelly, J. R., & Spoor, J. R. (2006). Affective Influence in Groups. In J. P. Forgas (Ed.), *Affect in Social Thinking and Behavior*. New York, NY: Psychology Press.

Kenny, D. A. (1996). Models of non-independence in dyadic research. *Journal of Social and Personal Relationships, 13*(2), 279-294.

Kenny, D. A., & Acitelli, L. K. (1994). Measuring similarity in couples. *Journal of Family Psychology, 8*(4), 417.

Kenny, D. A., Kashy, D., & Cook, W. L. (2006). *Dyadic data analysis*: Guilford Press.

Kim, J., & Moon, J. Y. (1998). Designing towards emotional usability in customer interfaces—trustworthiness of cyber-banking system interfaces. *Interacting with Computers, 10*(1), 1-29.

Kim, K. H., Bang, S. W., & Kim, S. R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical and biological engineering and computing, 42*(3), 419-427.

Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration:: Theory, design, and results. *Interacting with Computers, 14*(2), 119-140.

Kleiner, B. M. (2006). Macroergonomics: Analysis and design of work systems. *Applied ergonomics, 37*(1), 81-89.

Koufaris, M., & Hampton-Sosa, W. (2002). Customer trust online: examining the role of the experience with the Web-site. *Department of Statistics and Computer Information Systems Working Paper Series, Zicklin School of Business, Baruch College, New York*.

Kugler, T., Connolly, T., & Ordóñez, L. D. (2012). Emotion, decision, and risk: Betting on gambles versus betting on people. *Journal of Behavioral Decision Making, 25*(2), 123-134.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2012). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package).

Lalot, F., Delplanque, S., & Sander, D. (2014). Mindful regulation of positive emotions: a comparison with reappraisal and expressive suppression. *Frontiers in psychology, 5*.

Lang, P. J. (1995). The emotion probe. *American Psychologist, 50*(5), 372-385.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*.

Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. In M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 180-195). New York: Guildford Press.

Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference characteristic: A review. *Journal of Research in Personality, 21*(1), 1-39.

Larson, J. R. (2010). *In search of synergy in small group performance*. New York: Psychology Press.

Lee, J. D. (2006). Affect, attention, and automation. In A. F. Kramer, D. A. Wiegmann & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 73-89). New York: Oxford University Press, Inc.

Lee, J. D. (2007). Affect, attention, and automation. In A. F. Kramer, D. A. Wiegmann & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 73-89). New York: Oxford University Press, Inc.

Lee, J. D., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics, 35*(10), 1243-1270.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 46*, 50-80.

Lee, M. K. O., & Turban, E. (2001). A trust model for consumer internet shopping. *International journal of electronic commerce, 6*(1), 75-92.

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

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: physiological linkage and affective exchange. *Journal of personality and social psychology, 45*(3), 587-597.

Levy, B. I. (1984). Research into the psychological meaning of color. *American Journal of Art Therapy, 23*, 58-62.

Li, W., Harris, D., Li, L., & Wang, T. M. (2009). The Differences of Aviation Human Factors between Individualism and Collectivism Culture

Human-Computer Interaction. Interacting in Various Application Domains. In J. Jacko (Ed.), (Vol. 5613, pp. 723-730): Springer Berlin / Heidelberg.

Ligges, U. (2015). signal: Signal processing (Version 0.7-5).

Lim, S. I., Woo, J. C., Bahn, S., & Nam, C. S. (2012). *The Effects of Individuals' Mood State and Personality Trait on the Cognitive Processing of Emotional Stimuli.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Littlewort, G. C., Bartlett, M. S., Fasel, I. R., Chenu, J., Kanda, T., Ishiguro, H., et al. (2003). *Towards Social Robots: Automatic Evaluation of Human-robot Interaction by Face Detection and Expression Classification.* Paper presented at the NIPS.

Littlewort, G. C., Bartlett, M. S., Salamanca, L. P., & Reilly, J. (2011). *Automated measurement of children's facial expressions during problem solving tasks.* Paper presented at the Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on.

Littlewort, G. C., Whitehill, J., Wu, T. L., Fasel, I. R., Frank, M. G., Movellan, J. R., et al. (2011). *The computer expression recognition toolbox (CERT).* Paper presented at the Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on.

Lottridge, D., Chignell, M., & Jovicic, A. (2011). Affective Interaction Understanding, Evaluating, and Designing for Human Emotion. *Reviews of Human Factors and Ergonomics, 7*(1), 197-217.

Lyons, J., Stokes, C., Garcia, D., Adams, J., & Ames, D. (2008, 16-18 July 2008). *Trust and Decision Making: An Empirical Platform (CCP 204).* Paper presented at the Aerospace and Electronics Conference, 2008. NAECON 2008. IEEE National.

MacKinnon, D. P. (2007). *Introduction to statistical mediation analysis*. New York: Lawrence Erlbaum Associates.

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review, 17*(2), 144-158.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173-181.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological methods, 7*(1), 83.

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science, 8*(4), 277-301.

Madhavan, P., Wiegmann, D. A., & Lacson, F. C. (2006). Automation failures on tasks easily performed by operators undermine trust in automated aids. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(2), 241-256.

Malmir, M., Forster, D., Youngstrom, K., Morrison, L., & Movellan, J. R. (2013). *Home Alone: Social Robots for Digital Ethnography of Toddler Behavior.* Paper presented at the Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on.

Mancini, G., Ferrari, P. F., & Palagi, E. (2013). Rapid facial mimicry in geladas. *Scientific reports, 3*.

Manly, B. F. J. (2007). *Randomization, bootstrap and Monte Carlo methods in biology*. Boca Raton, FL: Chapman & Hall/CRC.

Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology, 87*(1), 3.

Markus, M. L. (1987). Toward a "critical mass" theory of interactive media. *Communication Research, 14*(5), 491.

Marriott, F. H. C. (1981). Barnard's Monte Carlo tests: how many simulations. *Applied Statistics, 28*, 75-79.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52*(1), 397-422.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion, 23*(2), 209-237.

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of personality and social psychology, 55*(1), 102.

McAllister, D. J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 24-59.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Mennecke, B. E., & Wheeler, B. C. (1993). *Tasks matter: modeling group task processes in experimental CSCW research.* Paper presented at the Proceeding of the Twenty-Sixth Hawaii International Conference on System Sciences, Wailea, HI.

Merritt, S. M. (2011). Affective Processes in Human–Automation Interactions. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(4), 356-370.

Merritt, S. M., & Ilgen, D. R. (2008). Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*(2), 194-210.

Montague, E. (2010). Validation of a trust in medical technology instrument. *Applied ergonomics, 41*(6), 812-821.

Montague, E., & Lee, J. D. (2012). Trust in Health Technologies. In P. Carayon (Ed.), *Handbook of human factors and ergonomics in health care and patient safety* (2nd ed., pp. 281-292). New York, NY: CRC Press.

Montague, E., Winchester III, W. W., & Kleiner, B. M. (2010). Trust in medical technology by patients and healthcare providers in obstetric work systems. *Behaviour & Information Technology, 29*(5), 541-554.

Montague, E., & Xu, J. (2012). Understanding active and passive users: The effects of an active user using normal, hard and unreliable technologies on user assessment of trust in technology and co-user. *Applied Ergonomics, 43*(4), 702-712.

Montague, E., Xu, J., & Chiou, E. (2014). Shared Experiences of Technology and Trust: An Experimental Study of Physiological Compliance between Active and Passive Users in Technology Mediated Collaborative Encounters. *IEEE Transactions on Human-Machine Systems, 44*(5), 614-624.

Moreland, R. L. (2010). Are dyads really groups? *Small Group Research, 41*(2), 251-267.

Mori, M., Ikeda, K., Hagiwara, G., Saga, M., Uehara, T., & Kita, H. (2009). Socialized Computers and Collaborative Learning. In H. Hattori, T. Kawamura, T. Idé, M. Yokoo & Y. Murakami (Eds.), *New Frontiers in Artificial Intelligence* (Vol. 5447, pp. 48-61): Springer Berlin / Heidelberg.

Movellan, J. R., Malmir, M., & Forester, D. (2012). HRI as a Tool to Monitor Socio-Emotional Development in Early Childhood Education.

Muir, B. M. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics, 37*(11), 1905-1922.

Müller, M. (2007). Dynamic time warping. In *Information retrieval for music and motion* (pp. 69-84). New York: Springer.

Myers, D., & Tingley, D. (2011). The influence of emotion on trust. *Cambridge, MA: Harvard University*.

Nardi, B. A. (1996). Studying context: A comparison of activity theory, situated action models, and distributed cognition. *Context and consciousness: Activity theory and human-computer interaction*, 69-102.

Niedenthal, P. M., Krauth-Gruber, S., & Ric, F. (2006). *Psychology of emotion: Interpersonal, experiential, and cognitive approaches*. New York: Psychology Press.

Niedenthal, P. M., & Setterlund, M. B. (1994). Emotion congruence in perception. *Personality and Social Psychology Bulletin, 20*(4), 401-411.

Niedenthal, P. M., Setterlund, M. B., & Jones, D. E. (1994). Emotional organization of perceptual memory. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 87-113). San Diego, CA: Academic Press.

Niederée, U., Jipp, M., Teegen, U., & Vollrath, M. (2012). *Effects of Observability, Mood States, and Workload on Human Handling Errors When Monitoring Aircraft Automation.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Noldus Information Technology. (2009). The Observer XT (Version 9.0). Leesburg, VA.

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

Nunamaker, J. F., Dennis, A. R., Valacich, J. S., Vogel, D., & George, J. F. (1991). Electronic meeting systems. *Communications of the ACM, 34*(7), 40-61.

O'Kane, P., & Hargie, O. (2007). Intentional and unintentional consequences of substituting face-to-face interaction with e-mail: An employee-based perspective. *Interacting with Computers, 19*(1), 20-31.

Ong, L. M. L., De Haes, J., Hoos, A. M., & Lammes, F. B. (1995). Doctor-patient communication: a review of the literature. *Social Science & Medicine, 40*(7), 903-918.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization science*, 398-427.

Pal, J., Pawar, U. S., Brewer, E. A., & Toyama, K. (2006). *The case for multi-user design for computer aided learning in developing regions.* Paper presented at the Proceedings of the 15th international conference on World Wide Web, Edinburgh, Scotland.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 39*, 230-253.

Pearson, S. D., & Raeke, L. H. (2000). Patients' trust in physicians: many theories, few measures, and little data. *Journal of General Internal Medicine, 15*(7), 509-513.

Perednia, D. A., & Allen, A. (1995). Telemedicine technology and clinical applications. *JAMA: the journal of the American Medical Association, 273*(6), 483.

Persson, A., Wanek, B., & Johansson, A. (2001). Passive versus active operator work in automated process control - a job design case study in a control centre. *Applied ergonomics, 32*(5), 441-451.

Picard, R. W. (1999). *Affective computing for HCI.* Paper presented at the Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Ergonomics and User Interfaces.

Picard, R. W. (2000). *Affective computing*: MIT press.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 23*(10), 1175-1191.

Pituch, K. A., Stapleton, L. M., & Kang, J. Y. (2006). A comparison of single sample and bootstrap methods to assess mediation in cluster randomized trials. *Multivariate Behavioral Research, 41*(3), 367-400.

Planalp, S. (1998). Communicating emotion in everyday life: Cues, channels, and processes. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion: Research, theory, applications, and contexts* (pp. 29-48). San Diego, CA: Academic Press.

Porges, S. W., Bohrer, R. E., Cheung, M. N., Drasgow, F., McCabe, P. M., & Keren, G. (1980). New time-series statistic for detecting rhythmic co-occurrence in the frequency domain: the weighted coherence and its application to psychophysiological research. *Psychological Bulletin, 88*(3), 580-587.

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures, 6*(2), 77-98.

R Core Team. (2014). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

R., B., & N., C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. New York, NY: Cambridge University Press.

Raddatz, K. R., Werth, A., & Tran, T. Q. (2007). *The influence of emotional state and pictorial cues on perceptual judgments.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Reese, K., Zheng, Y., & Elmaghraby, A. (2012). *A comparison of face detection algorithms in visible and thermal spectrums.* Paper presented at the Int'l Conf. on Advances in Computer Science and Application.

Rice, S., Trafimow, D., Keller, D., & Bean, N. (2012). Confluence theory: uniting two houses divided. *Theoretical Issues in Ergonomics Science, 13*(4), 411-429.

Robert B. Lount, J. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *Journal of personality and social psychology, 98*(3), 420.

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: Oxford University Press, Inc.

Rossi, F., Fasel, I., & Sanfey, A. G. (2011). Inscrutable games? Facial expressions predict economic behavior. *BMC Neuroscience, 12*(Suppl 1), P281.

Roter, D. L., Frankel, R. M., Hall, J. A., & Sluyter, D. (2006). The expression of emotion through nonverbal behavior in medical visits. *Journal of General Internal Medicine, 21*, 28-34.

Rozell, E. J., & Gardner III, W. L. (2000). Cognitive, motivation, and affective processes associated with computer-related performance: a path analysis. *Computers in Human*

Behavior, 16(2), 199-222.
Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and social psychology review, 5*(4), 296-320.
Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition and emotion, 23*(7), 1259-1283.
Safar, J. A., & Turner, C. W. (2005). *Validation of a two factor structure for system trust.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.
Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on, 26*(1), 43-49.
Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: does crew resource management (CRM) training work? *Human Factors: The Journal of the Human Factors and Ergonomics Society, 43*(4), 641-674.
Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a “Big Five” in teamwork? *Small group research, 36*(5), 555-599.
Saunders, M. N., Dietz, G., & Thornhill, A. (2014). Trust and distrust: Polar opposites, or independent but co-existing? *Human Relations, 67*(6), 639-665.
Schaafsma, J., Krahmer, E., Postma, M., Swerts, M., Balsters, M., & Vingerhoets, A. (2014). Comfortably Numb? Nonverbal Reactions to Social Exclusion. *Journal of Nonverbal Behavior, 39*(1), 25-39.
Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Lodon: Chapman & Hall.
Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social science information, 44*(4), 695-729.
Scherer, S., Stratou, G., & Morency, L.-P. (2013). *Audiovisual behavior descriptors for depression assessment.* Paper presented at the Proceedings of the 15th ACM on International conference on multimodal interaction.
Schwarz, N. (1990). Feelings as information: informational and motivational functions of affective states. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of Motivation and Cognition: Foundation of Social Behavior* (pp. 527-561). New York: Guilford Press.
Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion, 14*(4), 433-440.
Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of personality and social psychology, 45*(3), 513.
Sheridan, T. B. (2002). *Humans and automation: System design and research issues*. Santa Monica, CA: John Wiley & Sons, Inc.
Sherif, C. W., & Sherif, M. (1967). *Attitude, ego-involvement, and change*. New York: Wiley
Sherif, M., & Hovland, C. I. (1981). *Social judgment: Assimilation and contrast effects in communication and attitude change*. Westport: Greenwood.
Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological methods, 7*(4), 422.
Singh, I. L., Molloy, R., & Parasuraman, R. (1993). Automation-induced" complacency": Development of the complacency-potential rating scale. *The International Journal of Aviation Psychology, 3*(2), 111-122.
Smith, M. J., & Carayon-Sainfort, P. (1989). A balance theory of job design for stress reduction. *International Journal of Industrial Ergonomics, 4*(1), 67–79-67–79.

Smith, M. J., & Carayon, P. (2001). Balance Theory of job design. In W. Karwowski (Ed.), *The International Encyclopedia of Ergonomics and Human Factors* (pp. 1181-1184). London: Taylor & Francis.

Smith, M. J., & Sainfort, P. C. (1989). A balance theory of job design for stress reduction. *International Journal of Industrial Ergonomics, 4*(1), 67-79.

Sobel, M. E. (1987). Direct and indirect effects in linear structural equation models. *Sociological Methods & Research, 16*(1), 155-176.

Spain, R. D., Bustamante, E. A., & Bliss, J. P. (2008). *Towards an empirically developed scale for system trust: Take two.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Stokes, C. K., Lyons, J. B., Littlejohn, K., Natarian, J., Case, E., & Speranza, N. (2010). *Accounting for the human in cyberspace: Effects of mood on trust in automation.* Paper presented at the 2010 International Symposium on Collaborative Technologies and Systems (CTS).

Stoner, J. A. (1961). A comparison of individual and group decision involving risk. Unpublished Unpublished master's thesis. Massachusetts Institute of Technology.

Straus, S. G. (1999). Testing a Typology of Tasks An Empirical Validation of McGrath's (1984) Group Task Circumplex. *Small Group Research, 30*(2), 166-187.

Stryker, J. B., & Santoro, M. D. (2012). Facilitating face-to-face communication in high-tech teams. *Research-Technology Management, 55*(1), 51-56.

Sumwalt, R. L., Thomas, R. J., & Dismukes, K. (2002). *Enhancing flight-crew monitoring skills can increase flight safety.* Paper presented at the 55th International Air Safety Seminar, Dublin, Ireland.

Sunstein, C. R. (2002). The law of group polarization. *Journal of political philosophy, 10*(2), 175-195.

Swangnetr, M., Zhu, B., Taylor, K. B., & Kaber, D. B. (2010). Assessing the Effects of Humanoid Robot Features on Patient Emotion during a Medicine Delivery Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 54*(4), 349-353.

Teijeiro-Mosquera, L., Biel, J.-I., Alba-Castro, J.-L., & Gatica-Perez, D. (2014). What Your Face Vlogs About: Expressions ofEmotion and Big-Five Traits Impressions inYouTube. *IEEE Transactions on Affective Computing*.

Tellegen, A., Watson, D., & Clark, L. A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of personality and social psychology, 54*(6), 1063-1070.

Totterdell, P., & Holman, D. (2003). Emotion regulation in customer service roles: testing a model of emotional labor. *Journal of occupational health psychology, 8*(1), 55.

Totterdell, P., Kellett, S., Teuchmann, K., & Briner, R. B. (1998). Evidence of mood linkage in work groups. *Journal of personality and social psychology, 74*(6), 1504-1515.

Trafimow, D. (2009). Reeder's MIM as a special case of confluence theory.

Tran, T. Q., Raddatz, K. R., & Werth, A. (2008). *Improving Perceptual Judgments under Emotional States.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

van Buuren, S. (2013). *Flexible Imputation of Missing Data*. Boca Raton, FL: CRC Press.

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7*(3), 487-494.

van Ginkel, J. R., & Kroonenberg, P. M. (2014). Analysis of variance of multiply imputed data. *Multivariate behavioral research, 49*(1), 78-91.

Van Swol, L. M. (2009). Extreme members and group polarization. *Social Influence, 4*(3), 185-199.

Västfjäll, D. (2002). Emotion induction through music: A review of the musical mood induction procedure. *Musicae Scientiae, 5*(1 suppl), 173-211.

Vespa, P. (2005). Multimodality monitoring and telemonitoring in neurocritical care: from microdialysis to robotic telepresence. *Current opinion in critical care, 11*(2), 133.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision, 57*(2), 137-154.

Volpe, C. E., Cannon-Bowers, J. A., Salas, E., & Spector, P. E. (1996). The impact of cross-training on team functioning: An empirical investigation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 38*(1), 87-100.

Von Hippel, P. T. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology, 39*(1), 265-291.

Watson, D., & Vaidya, J. (2003). Mood measurement: Current status and future directions. In I. B. Weiner, J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (Vol. 2, pp. 351-375). Hoboken, NJ: John Wiley & Sons, Inc.

Weinstock, A., Oron-Gilad, T., & Parmet, Y. (2012). The effect of system aesthetics on trust, cooperation, satisfaction and annoyance in an imperfect automated system. *Work: A Journal of Prevention, Assessment and Rehabilitation, 41*, 258-265.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta‐analysis. *European Journal of Social Psychology, 26*(4), 557-580.

Whitworth, B. (2009). The social requirements of technical systems. In B. Whitworth & A. De Moor (Eds.), *Handbook of Research on Socio-Technical Design and Social Networking Systems* (pp. 3-22). Hershey, PA: IGI.

Wickens, C. D., Lee, J. D., Liu, Y., & Gordon, S. E. (2004). *An introduction to human factors engineering*. Upper Saddle River, N.J.: Pearson Prentice Hall.

Wiegmann, D. A., & Shappell, S. A. (2001). *Applying the human factors analysis and classification system (HFACS) to the analysis of commercial aviation accident data*. Paper presented at the the 11th International Symposium on Aviation Psychology.

Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis : the human factors analysis and classification system*. Burlington, VT: Ashgate.

Wiles, J. A., & Cornwell, T. B. (1991). A review of methods utilized in measuring affect, feelings, and emotion in advertising. *Current Issues and Research in Advertising, 13*(1-2), 241-275.

Williams, K. D. (2010). Dyads can be groups (and often are). *Small Group Research, 41*(2), 268-274.

Wilson, J. M., Straus, S. G., & McEvily, B. (2006). All in due time: The development of trust in computer-mediated and face-to-face teams. *Organizational Behavior and Human Decision Processes, 99*(1), 16-33.

Wittenbaum, G. M., Hollingshead, A. B., Paulus, P. B., Hirokawa, R. Y., Ancona, D. G., Peterson, R. S., et al. (2004). The functional perspective as a lens for understanding groups. *Small Group Research, 35*(1), 17-43.

Woolley, F. R., Kane, R. L., Hughes, C. C., & Wright, D. D. (1978). The effects of doctor-patient communication on satisfaction and outcome of care. *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology, 12*, 123-128.

Wu, A. D., & Zumbo, B. D. (2008). Understanding and using mediators and moderators. *Social Indicators Research, 87*(3), 367-392.

Xu, J., Le, K., Deitermann, A., & Montague, E. (2013). How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology. Unpublished Manuscript submitted for publication.

Xu, J., Le, K., Deitermann, A., & Montague, E. (2014). How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology. *Applied Ergonomics, 45*(6), 1495-1503.

Xu, J., & Montague, E. (2012). Psychophysiology of the passive user: Exploring the effect of technological conditions and personality traits. *International Journal of Industrial Ergonomics, 42*(5), 505-512.

Xu, J., & Montague, E. (2013a). Group polarization of trust in technology. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 57*(1), 344-348.

Xu, J., & Montague, E. (2013b). Working with an Invisible Active User: Understanding Trust in Technology and Co-User from the Perspective of a Passive User. *Interacting with Computers, 25*(5), 375-385.

Yu, Z., Scherer, S., Devault, D., Gratch, J., Stratou, G., Morency, L.-P., et al. (2013). *Multimodal prediction of psychological disorders: Learning verbal and nonverbal commonalities in adjacency pairs.* Paper presented at the Semdial 2013 DialDam: Proceedings of the 17th Workshop on the Semantics and Pragmatics of Dialogue.

Yuviler-Gavish, N., & Gopher, D. (2011). Effect of descriptive information and experience on automation reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(3), 230-244.

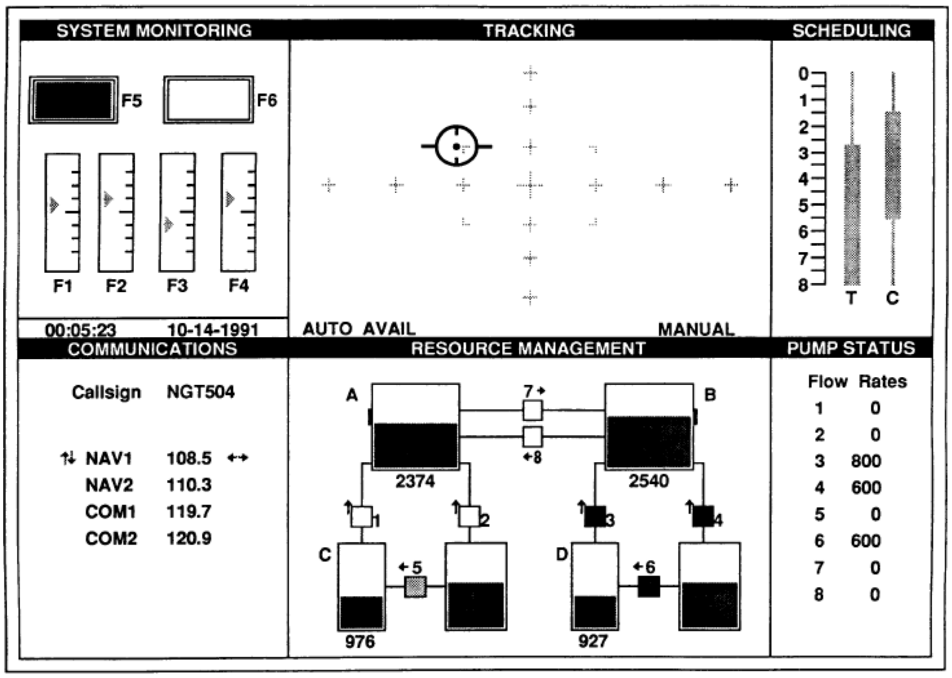
Zhang, Z., Wang, L., Shigemasu, K., Okada, A., Imaizumi, T., & Hoshino, T. (2008). Methods for evaluating mediation effects: Rationale and comparison. *New trends in psychometrics*, 595-604.

Appendix A: MATB Instructions

A.1. MATB instructions for active users

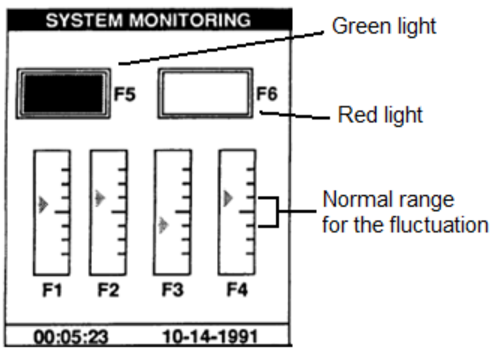
The Multi Attribute Task Battery is an upcoming technology with a potential of being used to do various tasks. This technology is in its early stages and hence different versions of the same technology will be used during different sections of this study. The MAT interface is composed of four different modules: monitoring, tracking, communication, scheduling as well the resource management task. For the purposes of our study, we will not be using the communication and scheduling module.

This is the interface of MAT you will see in the computer screen:



MAT is designed for helping the user to perform multiple tasks at the same time. The main modules and instructions are as follows:

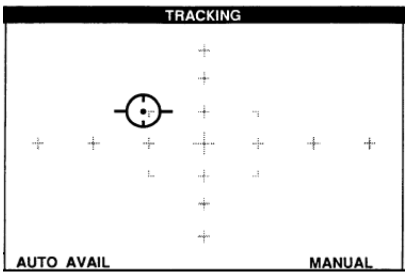
(1) Task one: system monitoring



The monitoring module of the interface presents you with tasks in the upper left window of the display screen. You need to monitor the various gauges as well as warning lights in this module. The absence of the green light, presence of the red light as well as four moving pointer dials for deviation from midpoint, requires your input. The two boxes in the upper portion of this module are the warning lights. One of the lights is normally “on” which is indicated by its green color. You should detect the absence of this light by tapping the “F5” key when the light green light turns off. The second light is normally turned “off”. When this light turns red, you are again required to detect this change by tapping the “F6” key. This same module also consists of the moving pointer dials task which consists of four vertical scales with moving indicators. Normally, the pointer dial fluctuates around the center of the scale within plus or minus one unit from the center. The pointer dial would move away from the middle of the vertical display to more than one unit above or below the center. You are required to detect this change and act by pressing the corresponding function key (F1 – F4). When you detect this change and presses the corresponding key, the pointer automatically corrects itself and moves immediately back to the center. Also upon doing so, a bar at the bottom of the dial is illuminated in yellow. If you fail to detect these changes in the warning lights and the pointer dials, they will return to their normal setting after a certain time period.

Your performance in this task will be calculated by <u>reaction time</u>, <u>misses</u>, and <u>false alarms</u>.

(2) Task two: tracking

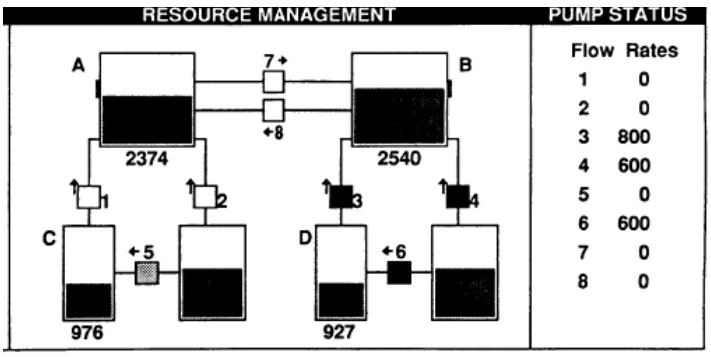


The second module of the MAT battery system is the tracking module which is located in the upper middle part of the screen. The target will move around randomly by itself. The main task for you is to keep the target in the center of the window. You should move the joystick around to make sure the cursor is within the dotted lines of the rectangle.

Your performance in this task will be calculated by the <u>deviation of the target from the center of the window</u>.

(3) Task three: resource management

The last module that we will be using for this study is the resource management module which is located in the lower middle part of the screen. There are a total of 6 rectangular fuel tanks which contain the green colored fuel controlled by the pumps. The maximum capacity for tanks A and B is 4000 units whereas the tanks C and D can only take up to 2000 units of fuel. The final two tanks have unlimited capacity. Your main task is to maintain tanks A and B at 2500 units each which is done by turning the pumps on or off. The lines where the tanks are interconnected have pumps which can be controlled and turned on by tapping on the corresponding number key. You should tap the same key again to turn the pump off. Each pump can only pump the fuel in one direction as indicated by the arrow above the pump itself. Pump failures may also occur at various times which are indicated by a red area on the failed pump. The flow rate of each pump is also displayed in the pump status window in the bottom right hand corner.



Your performance in this task will be calculated by the <u>deviation of A and B's fuel level from the optimum level of 2500</u>.

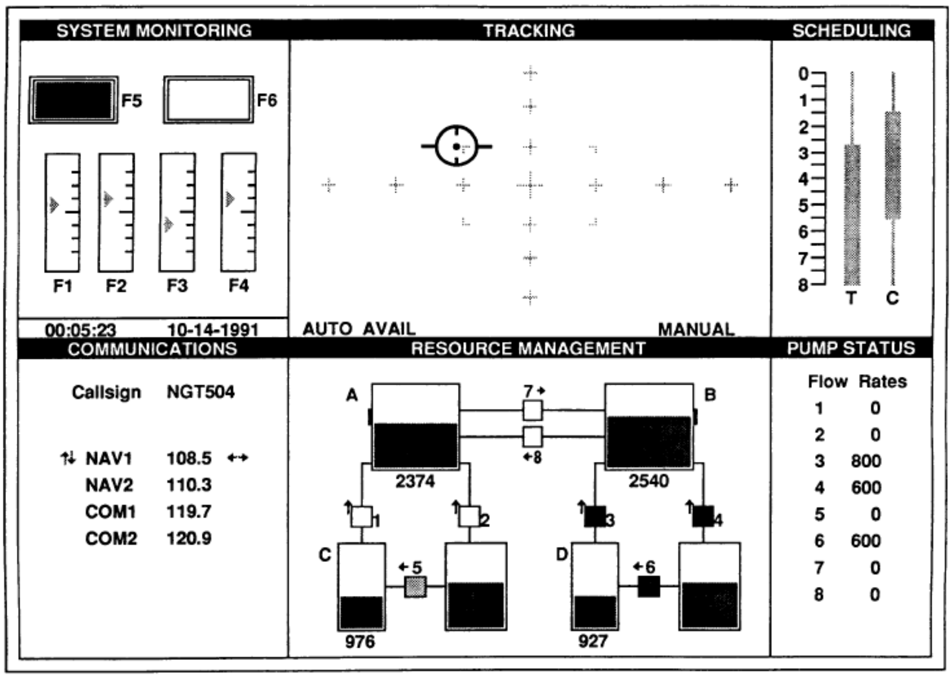
(4) Your role

There will be three versions of MAT for you and your teammate to operate, 6 minutes for each. The mission for your team is to keep high performance in all 3 tasks. The rule is, your teammate is not allowed the control the computer, but he/she can help you by talking to you. After each session, you will need to fill out a questionnaire before going to the next.

A.2. MATB instructions for passive users

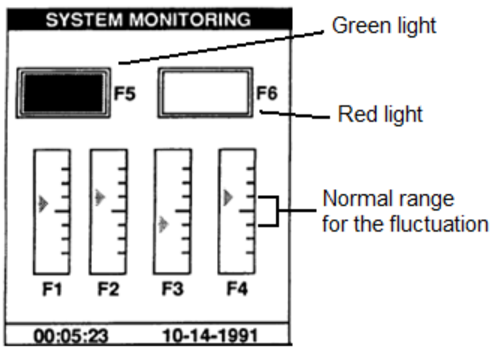
The Multi Attribute Task Battery is an upcoming technology with a potential of being used to do various tasks. This technology is in its early stages and hence different versions of the same technology will be used during different sections of this study. The MAT interface is composed of four different modules: monitoring, tracking, communication, scheduling as well the resource management task. For the purposes of our study, we will not be using the communication and scheduling module.

This is the interface of MAT you will see in the computer screen:



MAT is designed for helping the user to perform multiple tasks at the same time. The main modules and instructions are as follows:

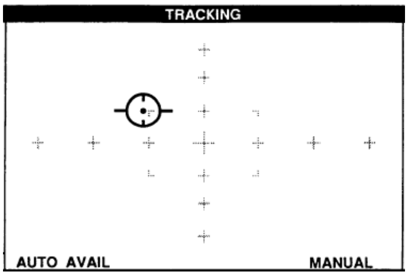
(1) Task one: system monitoring



The monitoring module of the interface presents you with tasks in the upper left window of the display screen. You need to monitor the various gauges as well as warning lights in this module. The absence of the green light, presence of the red light as well as four moving pointer dials for deviation from midpoint, requires your input. The two boxes in the upper portion of this module are the warning lights. One of the lights is normally "on" which is indicated by its green color. You should detect the absence of this light by tapping the "F5" key when the light green light turns off. The second light is normally turned "off". When this light turns red, you are again required to detect this change by tapping the "F6" key. This same module also consists of the moving pointer dials task which consists of four vertical scales with moving indicators. Normally, the pointer dial fluctuates around the center of the scale within plus or minus one unit from the center. The pointer dial would move away from the middle of the vertical display to more than one unit above or below the center. You are required to detect this change and act by pressing the corresponding function key (F1 – F4). When you detect this change and presses the corresponding key, the pointer automatically corrects itself and moves immediately back to the center. Also upon doing so, a bar at the bottom of the dial is illuminated in yellow. If you fail to detect these changes in the warning lights and the pointer dials, they will return to their normal setting after a certain time period.

Your performance in this task will be calculated by reaction time, misses, and false alarms.

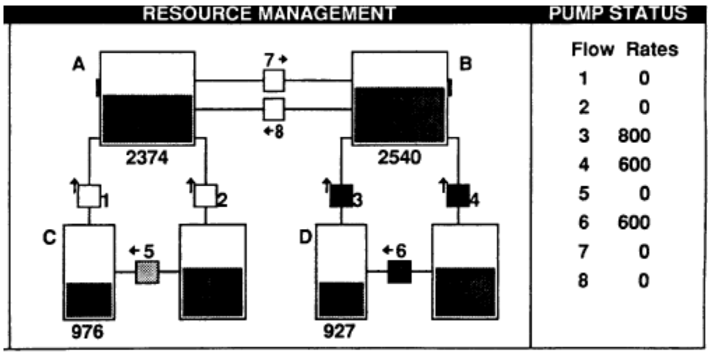
(2) Task two: tracking



The second module of the MAT battery system is the tracking module which is located in the upper middle part of the screen. The target will move around randomly by itself. The main task for you is to keep the target in the center of the window. You should move the joystick around to make sure the cursor is within the dotted lines of the rectangle.

Your performance in this task will be calculated by the deviation of the target from the center of the window.

(3) Task three: resource management



The last module that we will be using for this study is the resource management module which is located in the lower middle part of the screen. There are a total of 6 rectangular fuel tanks which contain the green colored fuel controlled by the pumps. The maximum capacity for tanks A and B is 4000 units whereas the tanks C and D can only take up to 2000 units of fuel. The final two tanks have unlimited capacity. Your main task is to maintain tanks A and B at 2500 units each which is done by turning the pumps on or off. The lines where the tanks are interconnected have pumps which can be controlled and turned on by tapping on the corresponding number key. You should tap the same key again to turn the pump off. Each pump can only pump the fuel in one direction as indicated by the arrow above the pump itself. Pump failures may also occur at various times which are indicated by a red area on the failed pump. The flow rate of each pump is also displayed in the pump status window in the bottom right hand corner.

Your performance in this task will be calculated by the deviation of A and B's fuel level from the optimum level of 2500.

(4) Your role

There will be three versions of MAT for you and your teammate to operate, 6 minutes for each. The mission for your team is to keep high performance in all 3 tasks. The rule is, you are not allowed the control the computer, but you can help your teammate by talking to him/her. After each session, you will need to fill out a questionnaire before going to the next.

Appendix B: Rating Scales

A.1. Propensity to trust technology

	Strongly disagree				Strongly agree
1. I think that automated devices used in medicine, such as CT scans and ultrasound, provide very reliable medical diagnosis.	1	2	3	4	5
2. Automated devices in medicine save time and money in the diagnosis and treatment of disease.	1	2	3	4	5
3. If I need to have a tumor in my body removed, I would choose to undergo computer-aided surgery using laser technology because it is more reliable and safer than manual surgery.	1	2	3	4	5
4. Automated systems used in modern aircraft, such as the automatic landing system, have made air journeys safer.	1	2	3	4	5
5. ATMs provide a safeguard against the inappropriate use of an individual's bank account by dishonest people.	1	2	3	4	5
6. Automated devices used in aviation and banking have made work easier for both employees and customers.	1	2	3	4	5
7. Even though the automatic cruise control in my car is set at a speed below the speed limit, I worry when I pass a police radar speed trap in case the automatic control is not working properly.	1	2	3	4	5
8. Manually sorting through card catalogues is more reliable than computer-aided searches for finding items in a library.	1	2	3	4	5
9. I would rather purchase an item using a computer than have to deal with a sales representative on the phone because my order is more likely to be correct using the computer.	1	2	3	4	5
10. Bank transactions have become safer with the introduction of computer technology for the transfer of funds.	1	2	3	4	5
11. I feel safer depositing my money at an ATM than with a human teller.	1	2	3	4	5
12. I have to tape an important TV program for a c1ass assignment. To ensure that the correct program is recorded, I would use the automatic programming facility on my VCR rather than manual taping.	1	2	3	4	5

A.2. Big Five personality

I am someone who…	Strongly disagree				Strongly agree
1. ______ Is talkative	1	2	3	4	5
2. ______ Tends to find fault with others	1	2	3	4	5
3. ______ Does a thorough job	1	2	3	4	5
4. ______ Is depressed, blue	1	2	3	4	5
5. ______ Is original, comes up with new ideas	1	2	3	4	5
6. ______ Is reserved	1	2	3	4	5
7. ______ Is helpful and unselfish with others	1	2	3	4	5
8. ______ Can be somewhat careless	1	2	3	4	5
9. ______ Is relaxed, handles stress well.	1	2	3	4	5
10. ______ Is curious about many different things	1	2	3	4	5
11. ______ Is full of energy	1	2	3	4	5
12. ______ Starts quarrels with others	1	2	3	4	5
13. ______ Is a reliable worker	1	2	3	4	5
14. ______ Can be tense	1	2	3	4	5
15. ______ Is ingenious, a deep thinker	1	2	3	4	5
16. ______ Generates a lot of enthusiasm	1	2	3	4	5
17. ______ Has a forgiving nature	1	2	3	4	5
18. ______ Tends to be disorganized	1	2	3	4	5
19. ______ Worries a lot	1	2	3	4	5
20. ______ Has an active imagination	1	2	3	4	5
21. ______ Tends to be quiet	1	2	3	4	5
22. ______ Is generally trusting	1	2	3	4	5
23. ______ Tends to be lazy	1	2	3	4	5
24. ______ Is emotionally stable, not easily upset	1	2	3	4	5
25. ______ Is inventive	1	2	3	4	5
26. ______ Has an assertive personality	1	2	3	4	5
27. ______ Can be cold and aloof	1	2	3	4	5
28. ______ Perseveres until the task is finished	1	2	3	4	5
29. ______ Can be moody	1	2	3	4	5
30. ______ Values artistic, aesthetic experiences	1	2	3	4	5
31. ______ Is sometimes shy, inhibited	1	2	3	4	5
32. ______ Is considerate and kind to almost everyone	1	2	3	4	5
33. ______ Does things efficiently	1	2	3	4	5
34. ______ Remains calm in tense situations	1	2	3	4	5
35. ______ Prefers work that is routine	1	2	3	4	5
36. ______ Is outgoing, sociable	1	2	3	4	5
37. ______ Is sometimes rude to others	1	2	3	4	5
38. ______ Makes plans and follows through with them	1	2	3	4	5
39. ______ Gets nervous easily	1	2	3	4	5
40. ______ Likes to reflect, play with ideas	1	2	3	4	5
41. ______ Has few artistic interests	1	2	3	4	5
42. ______ Likes to cooperate with others	1	2	3	4	5
43. ______ Is easily distracted	1	2	3	4	5
44. ______ Is sophisticated in art, music, or literature	1	2	3	4	5

A.3. Trust in technology

Below is a list of statement for evaluating trust between people and automation. There are several scales for you to rate intensity of your feeling of trust, or your impression of the system while operating the computer.

	Strongly disagree						Strongly agree
1. The system is deceptive.	1	2	3	4	5	6	7
2. The system behaves in an underhanded manner.	1	2	3	4	5	6	7
3. I am suspicious of the system's intent, action, or outputs.	1	2	3	4	5	6	7
4. I am wary of the system.	1	2	3	4	5	6	7
5. The system's actions will have a harmful or injurious outcome.	1	2	3	4	5	6	7
6. I am confident in the system.	1	2	3	4	5	6	7
7. The system provides security.	1	2	3	4	5	6	7
8. The system has integrity.	1	2	3	4	5	6	7
9. The system is dependable.	1	2	3	4	5	6	7
10. The system is reliable.	1	2	3	4	5	6	7
11. I can trust the system.	1	2	3	4	5	6	7
12. I am familiar with the system.	1	2	3	4	5	6	7

A.4. Mood state

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this way right now, that is, at the present moment. Use the following scale to record your answers.

	Very slightly or not at all				extremely
1. interested	1	2	3	4	5
2. distressed	1	2	3	4	5
3. excited	1	2	3	4	5
4. upset	1	2	3	4	5
5. strong	1	2	3	4	5
6. guilty	1	2	3	4	5
7. scared	1	2	3	4	5
8. hostile	1	2	3	4	5
9. enthusiastic	1	2	3	4	5
10. proud	1	2	3	4	5
11. irritable	1	2	3	4	5
12. alert	1	2	3	4	5
13. ashamed	1	2	3	4	5
14. inspired	1	2	3	4	5
15. nervous	1	2	3	4	5
16. determined	1	2	3	4	5
17. attentive	1	2	3	4	5
18. jittery	1	2	3	4	5
19. active	1	2	3	4	5
20. afraid	1	2	3	4	5

A.5. Familiarity

	Don't know at all				Good friends
How well do you know your teammate before the experiment?	1	2	3	4	5

How do you know your teammate? ____