

Emotional Understanding

The Adaptability of Accurate Emotional Predictions

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Psychologie

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät II

Humboldt-Universität zu Berlin

von

Frau Dipl.-Psych. Michaela Turß

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:

Prof. Dr. Elmar Kulke

Gutachter:

1. Prof. Dr. Wolfgang Scholl
2. Prof. Dr. Ralf Schulze
3. Prof. Gerald Matthews, PhD

eingereicht am: 06. März 2013

Tag der mündlichen Prüfung: 13. September 2013

Abstract

In the ability model of emotional intelligence by Mayer and Salovey (1997), emotional understanding is a prerequisite for emotion regulation. Knowing which emotions occur in which situations should be beneficial and adaptive. One of the subtests for emotional understanding asks for likely emotional reactions in hypothetical situations. In contrast, Gilbert and Wilson (2003) argue that characteristic biases in affective forecasting are adaptive. The current thesis aims to measure accuracy of emotional predictions in a natural setting and examines its adaptive value. In the anxiety study, public officials were asked to predict future emotions in an important test (N=143). The second study focused on freshman student work-groups (N=180 in 43 groups). Group members predicted interpersonal feelings for each other (affection, satisfaction with the collaboration, fun, and anger). In both studies, accuracy of emotional predictions is defined as low bias (i.e. Euclidean distance) and high correspondence (i.e. profile correlation). The round robin design in the work-group study also allows to decompose accuracy following Cronbach (1955). In both studies, a low bias was adaptive in terms of strong criteria, also incrementally over and above intelligence and personality alone. Accuracy was partly related to general knowledge but not to intelligence. Associations to emotional intelligence were inconsistent. Accuracy as correspondence is theoretically interesting but much less reliable. There is some evidence for its adaptive value on a group level but no indication of incremental validity. Future research should focus on specific situations and specific emotions. Also, processes underlying affective forecasts should be evaluated in detail.

Zusammenfassung

Im Rahmen des Leistungsansatzes von emotionaler Intelligenz sehen Mayer und Salovey (1997) Emotionsverständnis als Voraussetzung für Emotionsregulation. Es sollte nützlich sein zu wissen, wie man sich in bestimmten Situationen fühlen wird. Zur Messung werden unter anderem Vignetten eingesetzt, in denen Emotionen für hypothetische Situationen vorhergesagt werden. Im Gegensatz dazu postulieren Gilbert und Wilson (2003) charakteristische Fehler bei affektiven Vorhersagen, die motivational günstig sind. In der vorliegenden Arbeit wird die Akkuratheit emotionaler Vorhersagen im natürlichen Umfeld untersucht, um dessen adaptiven Wert zu beurteilen. Zunächst sollten Beamtenanwärter ihre Emotionen in einer bedeutenden Testsituation vorhersagen (N=143). Dann wurden studentische Arbeitsgruppen (180 Mitglieder in 43 Gruppen) gebeten, Gefühle zwischen den Mitgliedern zu prognostizieren (Zuneigung, Zufriedenheit mit der Zusammenarbeit, Freude und Ärger). Akkuratheit wurde als geringer Bias (euklidische Distanz) und hohe Korrespondenz (Profilkorrelation) definiert. Das Round Robin Design der zweiten Studie ermöglichte die Varianzzerlegung der Akkuratheit nach Cronbach (1955). In beiden Studien ist ein niedriger Bias adaptiv in Hinblick auf harte Kriterien, auch inkrementell über Intelligenz und Persönlichkeit hinaus. Bias hing teilweise mit Allgemeinwissen zusammen, aber nicht mit Intelligenz. Zusammenhänge zu emotionaler Intelligenz waren inkonsistent. Die Akkuratheit als Korrespondenz ist theoretisch interessant aber deutlich weniger reliabel. Auf Gruppenebene konnte die Korrespondenz Kriterien vorhersagen, aber es zeigte sich keine inkrementelle Validität. Zukünftige Forschung sollte sich auf spezifische Situationen und spezifische Emotionen konzentrieren sowie die Prozesse untersuchen, die emotionalen Vorhersagen zugrunde liegen.

Contents

1	General Introduction	1
2	Theoretical Background	4
2.1	Emotional Intelligence	4
2.2	Research on Emotional Predictions	9
2.3	Accuracy Research	13
2.4	Research Questions	16
3	Study 1: Anxiety Study	19
3.1	Methods	19
3.1.1	Sample	19
3.1.2	Procedure	20
3.1.3	Material	20
3.2	Results	24
3.2.1	Descriptive Statistics	24
3.2.2	Forecasts and Experiences over Time	25
3.2.3	Accuracy as Bias	27
3.2.4	Accuracy as Correspondence	31
3.3	Discussion	40
3.3.1	Bias, Correspondence, and other Abilities and Traits	40
3.3.2	The Adaptability of Bias and Correspondence	41
3.3.3	Connection to the Affective Forecasting Paradigm	42

4	Study 2: Work-Group Study	44
4.1	Methods	44
4.1.1	Sample	44
4.1.2	Design and Procedure	45
4.1.3	Material	45
4.2	Results	49
4.2.1	The Group as Unit of Analysis	49
4.2.1.1	Descriptive Statistics	49
4.2.1.2	Bias and Correspondence	50
4.2.1.3	Accuracy Decomposition	54
4.2.2	The Person as Unit of Analysis	56
4.2.2.1	Perceiver and Target Effects	56
4.2.2.2	Bias and Correspondence	57
4.3	Discussion	63
4.3.1	Stability of Bias and Correspondence	63
4.3.2	Bias, Correspondence, and other Abilities and Traits	64
4.3.3	The Adaptability of Bias and Correspondence	65
5	General Discussion	66
5.1	Theoretical Considerations	66
5.2	Strengths and Limitations	68
5.3	Implications for Future Research	69
5.4	Conclusion	72
	References	73

List of Figures

3.1	Course of Predictions and Experiences ($N = 128 - 141$) . . .	27
3.2	Anxiety as a Moderator of the Relationship between Bias and Coping	31
3.3	Relationship between Emotional Predictions (x) and Experiences (y) for High and Low Anxiety, and High and Low Positive Coping	39

List of Tables

3.1	Anxiety Study: Course of the Data Collection	21
3.2	Sample Items for Positive and Negative Coping	24
3.3	Intercorrelations and Descriptive Statistics	26
3.4	Repeated-measures ANOVA and Contrasts with Positive and Negative Affect as Outcome Variables	28
3.5	First-order Correlations of Difference Scores and Criteria . . .	30
3.6	One-way ANOVA with Random Effects. Unrestricted Model	33
3.7	Estimating Experiences with Predictions	35
3.8	Moderators of Correspondence	36
3.9	Adaptability of Correspondence	38
4.1	Intercorrelations and Descriptive Statistics (Group Level, N=43)	51
4.2	Retest Reliability and Validity of Accuracy Measures (Group Level, N=43)	53
4.3	Validity of Accuracy Components (Group Level, N=43) . . .	55
4.4	Effects for Perceivers, Targets, and Targets of the Second Degree	58
4.5	Hierarchical Random Intercept Model for Task and Relation- ship Conflict	61
4.6	Hierarchical Random Intercept Model for Competence and Friendship	62

Chapter 1

General Introduction

In recent years, the concept of emotional intelligence (EI) has been receiving increasing attention. Publications on EI have grown exponentially over at least two decades, and economic interest creates considerable market demands regarding measurement and training. Whereas the concept is embraced in public, it is controversially debated, if not severely criticized in academia. It is generally plausible that people differ in their ability regarding emotions, i.e. that some are more likely to succeed in emotionally challenging situations. Also, it is easy to appreciate the fact that emotion-related skills and characteristics, like dealing with stress, managing conflict or staying optimistic, are valuable in life and many people may perceive someone to be emotionally gifted. EI proves useful to talk about success in everyday life, especially in social life. Nearly every useful addition to general intelligence can be projected into this concept. Also, it emphasizes strengths that go beyond scholastic achievement. When it comes to the theoretical basis, though, and to generally accepted definitions and measurement procedures, academic intelligence outclasses EI by far.

The current thesis aims to add to the comprehension of emotional understanding, one subbranch of the ability model of emotional intelligence (Mayer & Salovey, 1997). Currently, vignette items of hypothetical situ-

ations are used for its measurement. In the conducted studies, emotional understanding was assessed in a natural setting and hypothetical situations were replaced with real ones. Participants were asked to predict their emotions and, later on, the same participants reported on their actual experiences. Predictions and experiences are now combined to calculate accuracy scores in terms of a low bias (i.e. Euclidean Distance) and high correspondence (i.e. profile correlation). This accuracy is related to abilities, traits, emotionally relevant criteria.

The ability model aims to define emotional intelligence as intelligence in the content domain of emotions. To do so, performance should be evaluated as right or wrong, with right answers to be generally preferable. Since emotional understanding is defined as emotional knowledge, it is obvious that accurate knowledge should be worth thriving for. In this sense, Mayer and Salovey (1997) see emotional understanding as a prerequisite of emotional management, so that accurate knowledge can be used to effectively influence emotions in ourselves and others. In other words, accuracy should be adaptive.

Other lines of research have taken different approaches. The notion of depressive realism suggests that depression is associated with accurate expectations which consolidate the disease (Alloy & Abramson, 1979). Others emphasized that a 'rosy view' is preferable (Mitchell, Thompson, Peterson, & Cronk, 1997), and that there are certain biases in *affective forecasting* that can serve a purpose (Wilson & Gilbert, 2003). One of these biases is *immune neglect*, i.e. the lack of insight into mood repair processes and, therefore, the overestimation of the duration of emotions. This is supposed to be functional because it enforces to seek positive events and to avoid negative ones (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998).

Generally, for the predictions of one's own emotions, the adaptability of accuracy can be called into question. The current thesis explores this

domain with a special focus on predictions for different targets. Two studies were conducted. In the anxiety study, public officials predicted emotions in an important exam situation. The design is comparable to typical affective forecasting studies, and it focuses on a negative emotion toward a specific event. Results can be interpreted in context of affective forecasting studies, and it can be explored how interindividual differences in bias relate to criteria. Even if a general bias is functional, it probably should not be as high as possible. Also, an accurate view can be compared to a rosy one.

In contrast, the work-group study focuses on the prediction of interpersonal feelings and relationship development. Following the EI rationale, emotional knowledge can be used to establish positive and effective relationships. This adds a social perspective to emotional predictions and allows to examine emotional understanding as to the self and others. This is especially important since the original conceptualization of ability EI emphasized this distinction.

In both studies, accuracy of emotional predictions will be related to abilities, traits, and criteria. Construct validity and incremental value will be reported. Then, strengths and limitations will be discussed and possibilities for future research will be explored.

Chapter 2

Theoretical Background

The aim of this thesis is to explore the accuracy of emotional predictions and its relation to other abilities, traits, and criteria. First, research on *emotional intelligence* (EI) is introduced (2.1) and problems regarding its conceptualization and measurement are discussed. Secondly, research on emotional predictions is presented (2.2). Then, the calculation of accuracy is addressed (2.3). In a last step, these theoretical frameworks are combined and it is argued that emotional understanding, as one branch of ability EI, can be conceptualized as accuracy of emotional predictions (2.4).

2.1 Emotional Intelligence

Contrary to popular beliefs, the concept of emotional intelligence was not introduced by Goleman (1995) but by Salovey and Mayer (1990). The original idea was to broaden traditional views of intelligence and to further explore the content domain of emotions. It was Goleman's publication, however, that triggered an overwhelming public interest for the concept, led to broad media coverage, and established its role in the business world (e.g. Matthews, Zeidner, & Roberts, 2003). Academically the concept has been less enthusiastically perceived. EI is called obscure and unnecessary,

as well as mainly a label to make money (e.g. Brody, 2004; Asendorpf, 2002). Before the criticism is addressed, two opposing frameworks will be distinguished because they are associated with different problems.

Whereas the ability model aims to assess maximal performance regarding emotions (cf. Ackerman, 1994; Cronbach, 1984), mixed models rely on self report to capture typical performance as to different emotion-related traits and skills. Empirically, the ability model of EI and different kinds of mixed models are only weakly related, if at all. A meta-analysis of Van Rooy and Viswesvaran (2004) reports an association of .13 which does not support the assumption of a coherent construct. In a more recent meta-analysis, self-report and ability measures of EI only correlate with .12, even if they are both meant to capture the ability model. Zeidner, Shani-Zivotich, Matthews, and Roberts (2005) compared ability and mixed model EI in gifted and non-gifted high school students. Gifted students scored high on the MSCEIT but low on self-reports EI and non-gifted students low on the MSCEIT and high on the self-reports.

The mixed models capture different desirable and more or less emotion-related traits and competencies. Bar-On (1997), for example, defined EI as 'an array of non-cognitive capabilities, competencies, and skills that influence one's ability to succeed in coping with environmental demands and pressures' (p. 6). The main problem with these models is a lack of discriminant validity regarding standard personality measures (Van Rooy & Viswesvaran, 2004; Van Rooy, Viswesvaran, & Pluta, 2005; Joseph & Newman, 2010) and a high number of subscales with substantial conceptual overlap. Also, the validity of self reports of competencies must be questioned (e.g Kruger & Dunning, 1999). The Emotional Quotient inventory (EQ-i, Bar-On, 1997), for example, distinguishes 5 scales and 15 subscales. Problem solving is measured with items like *My approach in overcoming difficulties is to move step by step.* and social responsibility contains items like

I like helping people. According to a reanalysis of Matthews et al. (2003, p. 209f), '80- 90 % of the reliable variance in the scales can be explained by just three factors, [...] self-esteem, empathy, and impulse control.' Therefore, EI as measured with the mixed model approach does not warrant a new construct called emotional intelligence.

The ability model shows slightly more promising results to support a new construct of EI. In their early definition, Salovey and Mayer (1990) state that EI 'involves the ability to monitor one's own and others' feelings and emotions, to discriminate among them, and to use this information to guide one's thinking and action' (p. 189). Thus, intelligence is applied to emotions. Rather than self-reported typical performance, maximal performance (cf. Ackerman, 1994; Cronbach, 1984) is assessed in terms of right and wrong. Accordingly, the MSCEIT, as the prominent diagnostic tool of ability EI (Mayer, Salovey, & Caruso, 2001), is less susceptible to faking than self report measures (Day & Carroll, 2008). The relationship of ability EI to general intelligence is moderate and within expectations (.32 and .22, Van Rooy & Viswesvaran, 2004; Joseph & Newman, 2010), with a strong relationship to emotional understanding (.39, Joseph & Newman, 2010). Also, there was evidence for discriminant validity regarding personality measures (Van Rooy & Viswesvaran, 2004), which has been called into question since (Fiori & Antonakis, 2011). The test is severely criticized for its psychometric properties, lack of incremental predictive validity, and its scoring procedure (e.g. Matthews et al., 2003). Mayer, Salovey, Caruso, and Cherkasskiy (2011) argue that the MSCEIT represents an integrative ability model of EI as opposed to specific-ability models that 'examine a particular realm of emotional intelligence in depth – for example, perceiving emotion in faces' (p. 531). However, the current trend is to focus on specific abilities and to add theoretical depth (e.g. Matthews, Zeidner, & Roberts, 2012; Mayer, Roberts, & Barsade, 2008).

The MSCEIT measures four branches of ability EI: perceiving emotions, facilitating thought, understanding emotions and managing emotions (Mayer et al., 2001). Those branches reflect the definition of Mayer and Salovey (1997): EI is '...the ability to perceive emotions, to access and generate emotions so as to assist thought, to understand emotions and emotional knowledge, and to reflectively regulate emotions so as to promote emotional and intellectual growth' (p. 5).

Right and wrong answers in the MSCEIT are not as easily determined as in standard intelligence tests. There are three different scoring procedures: consensus scoring, expert scoring, and target scoring (Mayer, Salovey, Caruso, & Sitarenios, 2003). Consensus scoring is most common and applied in the current study. For each question, there is a frequency distribution in the norm sample that is used to score the answers. Participants who agree with the majority of the norm sample receive higher scores. With consensus scoring, it is logically impossible to diagnose emotional genius because the best answers are those of many. Roberts, Zeidner, and Matthews (2001) hypothesize that consensus scoring causes the MSCEIT to measure social conformity.

Expert scoring means that right and wrong answers are calculated based on the agreement with the experts' opinion, and target scoring defines a right answer as that of the target. In the MSCEIT, target scoring can only be applied to one subtest of the branch *perceiving emotions*. An emotion attributed to a facial expression will be considered right, if the actual person that is depicted wants to convey that emotion. For all other items in the MSCEIT, targets are hypothetical and non-existent. The paradigm of this thesis can be conceptualized as target scoring of emotional understanding because the actual experiences of certain targets serve as a scoring criterion for accurate predictions.

According to Mayer and Salovey (1997), EI is generally adaptive because

it guides thought and action. They state that people with high emotional ability should have good mental health. Their well-being is promoted by effective emotion regulation in self and others, so that 'the emotionally intelligent person is often a pleasure to be around and leaves others feeling better' (Salovey & Mayer, 1990, p. 201). This statement can be called into question. High cognitive ability as to emotions could be used to manipulate others (Matthews et al., 2003) or to end undesirable relationships. In some cases, there needs to be a trade-off between one's own and others' goals and between short- and long-term consequences. Not all criteria can be met in every situation. Someone might achieve their social goals, without, for example, being generally well-perceived. Thus, the relationship to different criteria is not easily predicted and highly context-specific (cf. Matthews et al., 2012).

The focus of the current thesis is on emotional understanding as a sub-branch of EI. It is defined as the ability '... to understand emotions and to utilize emotional knowledge' (Mayer & Salovey, 1997, p. 12). More precisely, according to Mayer et al. (2001) it includes understanding the relationships among emotions, knowing causes and consequences, understanding complex emotions and blends, and being aware of potential transitions among emotions. Emotional understanding should be based on emotional perception and facilitation, and it should enable emotional management. Joseph and Newman (2010) propose a cascading model of EI that drops emotional facilitation and states that perception influences understanding which influences regulation.

One subtest of *emotional understanding* presents hypothetical situational vignettes. Participants are asked which emotions will likely arise in these situations, and how these emotions are going to develop. Since emotional understanding promotes emotional regulation, this knowledge, i.e. knowing which emotions are elicited in certain situations, should facilitate

success for the self and for others. There is evidence for the adaptive value of emotional knowledge in children (e.g. Denham & Kochanoff, 2002). It is important to note, though, that emotional knowledge in this context is measured with items similar to the MSCEIT subbranch emotional perception. For example, Izard et al. (2001) showed photographs of emotional expressions. Then, children were asked to freely label these emotions and to recognize emotional labels in a multiple-choice-task.

In sum, Mayer and Salovey (1997) believe that accurate knowledge about emotions is generally desirable because it helps to understand causes and consequences and because it can guide behavior. In contrast, (e.g. Wilson & Gilbert, 2003) assume that biases in emotional predictions can be beneficial. Without a general focus on individual differences, numerous studies found evidence for such biases.

2.2 Research on Emotional Predictions

When emotional predictions are examined on a group level, there is evidence for many inaccuracies. Mountain climbers, for example, underestimate the torment involved in their endeavors (Mitchell et al., 1997). Which is not only assumed to be without negative consequences but to be functional. People are more likely to approach certain situations again, when they overlook negative side effects. This promotes positive experiences.

A lack of useful biases has been called depressive realism (Alloy & Abramson, 1979). It indicates that part of affective disorders is the accurate perception of negative aspects. This leads to avoidance of activity and enhances the symptoms of depression since positive experiences become less likely. Similarly, Gilbert (2006) claims that it might be accurate to pick up on social rejection and it might be true that most leisure activities have burdening aspects, but seeing this does not make us happy. Empirical evidence on depressive realism is mixed. Dunning and Story (1991) suggest

that it is not a universal phenomenon but can become relevant under certain conditions.

Besides the underestimation of negative affect in positive situations, as mentioned above, Gilbert et al. (1998) focus on the overestimation of pain and its duration in negative situations. They argue that people generally assume to be immensely and long-lastingly affected by traumatic life experiences, for example the death of a child, paralysis, job-loss, or severe illness. In contrast, the actual negative impact of such events on emotional well-being is more short-lived. After a negative event, many people associate something valuable with it, for example an opportunity for personal growth. People might appreciate life to a higher extent, value the social support that they received, or focus on positive aspects of an occupational change. All these considerations are examples of mood repair processes that come into play after the fact. Gilbert et al. (1998) assume that there is a psychological immune system that, similarly to the actual immune system, operates beyond awareness. This unawareness of mood repair is called *immune neglect*, and it explains why accomplishments of the psychological immune system are generally not foreseen. This bias is adaptive because people will work harder to prevent negative events if they overestimate the negative impact. This conceptualization of mood repair as mainly automatic and preferably beyond awareness contradicts the idea of emotional knowledge that is actively used to manage emotions.

Immune neglect is one topic examined in affective forecasting research (Wilson & Gilbert, 2003). In this paradigm, participants predict the valence of emotional reactions in certain situations. In most cases, another sample of participants report on their actual emotional experiences. Then, predictions and experiences are compared on a group level. Individual differences are seldom addressed but Hoerger (2012) found that interindividual differences in coping styles are indeed not taken into account when predicting emotions

for Valentine's Day with or without a date.

Originally, individual differences were overlooked in affective forecasting research, even if early studies already provided evidence for such differences. A study by Gilbert and Ebert (2002) showed that 66.3% of the participants made emotionally biased decisions in a dichotomous task which implies that 33.7% of the subjects, a rather large group, did not. Those studies that focus on interindividual differences rely on *prediction error scores*, i.e. the simple difference scores between predictions and criteria (Dunn, Brackett, Ashton-James, Schneiderman, & Salovey, 2007; Hoerger, Chapman, Epstein, & Duberstein, 2012; Hoerger, 2012). Also, Dunn et al. (2007) uses a one-item measure of affect, specifically valence on a 9-point scale.

Wilson and Gilbert (2003, p.354) argue that affective forecasting involves a series of steps. First, we mentally simulate the situation for which the forecast is made. *Construal* is the process that builds a mental *representation of the event*. Then, we *recall affective theories* to infer an *assessed affective reaction*, and we correct for *unique influences*. All these factors lead to the *affective forecast*.

Arguably the biggest source of unsystematic error is *misconstrual* of the event. Error due to the *recall of affective theories*, on the other hand, is valid since it reflects inadequate knowledge. In a similar way, the capability to correct for *unique influences* (f.e. one's own mood or person-specific characteristics) can vary and represents emotional knowledge. Apart from misconstrual and inaccurate theories, error in emotional predictions can be caused by motivated distortions (f.e. exaggerations of negative emotions for unwanted events), undercorrection (f.e. neglecting one's own mood), and focalism (to neglect other events and aspects that also influence our mood (Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 2002). With regard to the experience, Wilson and Gilbert (2003) distinguish the *initial affective experience* and the *affective experience over time*. *Sense making processes*

(like f.e. coping mechanisms) influence the transition from initial to later emotional response.

Another framework that deals with the prediction of emotional reactions (as *prospective self-report*) is the emotional self report model by Robinson and Clore (2002a). Accordingly, when asked about future or hypothetical emotional reactions, people can rely on four sources of information: the episodic experience itself, episodic memory, situation specific beliefs, and identity related beliefs. The first two are episodic in nature, experiential, and contextualized. In contrast, the latter two are semantic and taken out of context. An important assumption is that the emotion itself can never be stored in memory, nor be retrieved from it. The episodic experience is only available in online self-reports and thinking about emotions is always qualitatively different from experiencing them (Robinson & Clore, 2002b). When remembering an emotion, details of the situation are retrieved and the according emotion is inferred. Real emotions can be elicited during the retrieval process but these emotions are considered new emotions that are never identical to the original ones.

Within the model of Robinson and Clore (2002a) three principles apply, namely accessibility, dominance, and evanescence. Accessibility means that emotion reports are based on the most accessible knowledge. Dominance means that the sources of information are hierarchically ordered and early stages are more dominant than later ones. Shortly after an event, the episodic memory can be accessed and this information will be used over other sources. If situation specific information is available, it will be used over the person specific one, and if episodic information is available, it will be used over the situation specific one. The last principle, evanescence, means that episodic aspects fade rapidly, in general. When this model is applied, emotional understanding, as measured in the MSCEIT, primarily focuses on situation specific beliefs. There is no actual experience involved,

and the targets of predictions are non-existent.

When prospective and retrospective emotional self-reports are examined, temporal distance of the event is important (Liberman, Sagristano, & Trope, 2002). With a background of construal level theory, Liberman et al. (2002) assume that temporal distance influences the level of abstractness that is applied to the mental construal of the event. Participants imagined events in the near vs. distant future, meaning in a couple of days vs. in a year. Near events were represented in more detail whereas far events were more abstractly construed. The focus on detail and ambiguous information is therefore impaired for distant events. Judgments should be holistic and schema driven. For example, good days seemed better and bad days seemed worse when they were imagined a year from now vs. tomorrow. The authors compared these findings to the focalism hypothesis of Wilson, Wheatley, Meyers, Gilbert, and Axson (2000), arguing that focalism, or a disregard of detail should be more likely for distant events. This points to the context-dependency of accuracy of emotional predictions.

2.3 Accuracy Research

The main distinction regarding accuracy must be made between accuracy in terms of profile similarity and accuracy in terms of a lack of bias or error. Murphy and Balzer (1989) compared different accuracy and error scores and found no substantial correlations between different types of measures. This emphasizes the necessity to take both, profile similarity (here called *correspondence*) and *bias*, into account (Kenny & Acitelli, 2001).

Early accuracy research only operated with bias scores, or more specifically simple difference scores between predictors and criteria. This is also done in affective forecasting and when individual differences in affective forecasting accuracy are addressed (e.g Gilbert & Ebert, 2002; Hoerger et al., 2012). In an influential article, Cronbach (1955) severely criticized this

widespread approach. Bernieri (2001) attributes the substantial shift in focus, away from accuracy research, to this article. Kenny and Albright (1987) see it as a central reason for social psychology focusing on social cognition instead of social relations (for a critique of this phenomenon cf. Scholl, 2007).

Cronbach argued that difference scores fail to answer the general question of accuracy in interpersonal perception because these scores confound four distinguishable variance components of accuracy and thereby distort a clear interpretation. The first one is *elevation*, as the mean difference between predictors and criteria. The second type of accuracy, *differential elevation accuracy*, captures to what extent differences between the targets are accurately perceived by the judges. Then, *stereotype accuracy* represents the accuracy regarding different traits. The last component of accuracy is *differential accuracy* or the extent to which judges accurately capture target variation on certain traits, i.e. the uniqueness of a trait-target combination. According to Cronbach (1955), this final component represents the social perception accuracy that researchers are typically interested in. Borman (1977) not only agrees that differential accuracy is the most important part but states that differential elevation and stereotype accuracy are not interesting in general.

Sulsky and Balzer (1988) emphasize the parallel nature of Cronbach's approach to analysis of variance. They state that elevation is comparable to the differential grand mean, or the general deviation of predictions and criteria. Differential elevation accuracy equals a main effect of rates or targets, i.e. the variance in the predictions that can be explained with target differences. Stereotype accuracy is the differential main effect of rating dimensions or traits, i.e. the variance in the predictions that can be explained with differences regarding the traits. Finally, differential accuracy refers to the interaction of targets and traits, i.e. variance in the predictions that

captures target differences on specific traits.

It is not always possible to calculate all the components of accuracy mentioned by Cronbach (1955). In order to calculate differential elevation accuracy, the research design needs to distinguish different targets for which predictions are made. Stereotype accuracy can only be calculated when different traits are distinguished for every target, and differential accuracy can only be distinguished from the error term when differential elevation accuracy and stereotype accuracy can both be calculated.

Social perception studies vary regarding their nomological or ideographic nature. They either focus on accuracy generally present in one group of participants, i.e. nomological accuracy, or on individual differences in accuracy, i.e. ideographic accuracy. Kenny and Albright (1987) generally argue for an examination of accuracy on a group level because of reliability deficits. The focus should be on *when* and *how* people are accurate and not on *who* is accurate. They say that the failure to replicate findings is due to insufficient variance. This is an important distinction from Cronbach (1955) whose original intention was to focus on interindividual differences in accuracy in order to distinguish good judges of personality (e.g. Colvin & Bundick, 2001; Christiansen, Wolcott-Burnam, Janovics, Burns, & Quirk, 2005). Later, Kenny and Winquist (2001) revised the initial statement to only focus on a group level. They suggested to combine the nomological and the ideographic approach. In one case, moderators of groups level accuracy are in focus, in the other case, individual accuracy scores are related to criteria.

Social relation analysis is a theoretical and methodological framework that can be used for multiple questions in interpersonal perception (Kenny, 1994). Here, it is used to decompose accuracy variance into perceiver and target effects, indicating variance due to the judges and variance due to the targets. When searching for a skill of emotional understanding in terms

of emotional prediction accuracy, there should be a considerable perceiver effect, indicating that people vary in their accuracy of their emotional predictions.

2.4 Research Questions

Emotional understanding was examined in emotionally challenging situations. From an emotional intelligence perspective, accuracy can easily be conceptualized as a corollary of *emotional understanding*, one branch of the ability model of EI by Mayer and Salovey (1997). Instead of using hypothetical situational vignettes (like in the MSCEIT), participants will predict emotions for actual upcoming situations. Their experiences in these situations will then serve as a scoring criterion. Accuracy will be calculated as *bias*, i.e. Euclidean distance, and *correspondence*, i.e. profile correlation. Also, the accuracy components of Cronbach (1955) will be calculated when possible.

The main assumption of this thesis is that accurate predictions should be based on acquired emotional knowledge which enables effective emotional management. It should, thereby, be adaptive in emotional encounters. Emotional knowledge is examined as to the self and to others, a distinction originally made by Salovey and Mayer (1990) in their definition of EI.

Two studies were conducted. The anxiety study focused on a specific emotion in a specific context. Public officials predicted their own emotions regarding an important test. Accuracy is related to independently measured distress, coping strategies, and satisfaction with the grade. This study is similar to affective forecasting studies but emphasizes individual differences and incremental validity.

The work-group study examined emotional understanding in a social context, a perspective that is often neglected. Freshman students gave a prognosis on the relationship-development in newly emerging groups, specif-

ically how interpersonal feelings were going to develop. Each group member predicted affection, satisfaction with the collaboration, fun, and anger that everyone in the group was going to feel for everyone else at the end of the semester. The end of the semester was a point of time when much work had to be done, presentations were held and part of the performance was graded. Emotional understanding should help to develop positive and successful relationships. This is arguably crucial, especially for freshman who should be interested in new relationships. The round robin design allows for accuracy decomposition following Cronbach (1955) and enables the calculation of potent peer ratings that capture how every individual is perceived by their fellow group members.

Accurate knowledge about interpersonal feelings (i.e. who likes whom, or who will be angry with whom) should provide important information when it comes to gaining affection or avoiding conflict. Term paper ratings, conflict levels, mutually perceived friendship and competence serve as criteria of successful goal attainment. Additionally, in the second study emotional predictions were done twice, so that retest reliabilities could be calculated.

The following research questions are addressed:

Bias, Correspondence, and other Abilities and Traits

As a proxy measure of emotional understanding, accuracy should be correlated to EI, especially to emotional understanding. As an ability, it should be moderately related to intelligence and general knowledge. Regarding personality, it should show discriminant validity, i.e. low associations.

The Adaptability of Bias and Correspondence

Accuracy should promote emotional management for the self and for others. Therefore, it should be related to less distress, better coping strategies, and positive relationship development. Satisfaction with the grade (in the anx-

iety study) and the term paper rating (in the work-group study) represent personal success that should be promoted by emotional knowledge, as well.

Connection to the Affective Forecasting Paradigm

This can only be tested in the anxiety study. On a group level, affective forecasting results should be replicated, specifically the overestimation of the duration of negative emotions. Regardless of a general bias, though, accuracy should again be adaptive as to the criteria because it enables emotional management.

Stability of Bias and Correspondence

This can only be tested in the work-group study. As a measure of individual differences, bias and correspondence should be stable in terms of retest reliability. When social relation analysis is applied, accuracy should show a substantial perceiver effect indicating that some people are generally better in predicting future emotions than others.

Chapter 3

Study 1: Anxiety Study

3.1 Methods

3.1.1 Sample

The participants in this study were public officials attending a specialized school for law, city administration, and social sciences. The educational program lasted for 3 years and participants had attended for 1 month when they entered the study. Tenured positions were only available for the top 20 % of the students, based on the grades in the final exams. The other students had to seek employment elsewhere.

The student population consisted of 149 subjects in five different classes. Participation was voluntary, anonymous, and rewarded with detailed written feedback on the conducted tests. Data was collected in class, and missing data occurred due to nonattendance in certain lessons. 143 students took part and 112 students completed all measurements. The mean age was 24.27 years (range: 18 - 41, $SD = 5.52$) and 55.9 % were female.

3.1.2 Procedure

The study took approximately six months from October 2005 to March 2006 and covered the first semester in school. Each of the five classes completed the questionnaires in the same week, either on Tuesday, Wednesday, or Thursday. All the questionnaires were conducted in class with a teacher accessible for questions. This teacher collected the data for a diploma thesis.

The course of the data collection is presented in Table 3.1. The emotionally challenging event was a social science test in the 1st week of December, which was the first in a series of critical examinations. One month prior to this test, participants were asked to predict their own and the group's average emotional reaction immediately before, immediately after, and one hour after the test. At the time of prediction, the students had completed three weeks of the program and were familiar with the procedures and the classmates. General mood was measured directly before the predictions were made. Intelligence, personality, and trait anxiety were measured afterwards. Emotional intelligence was measured later due to time constraints in class. On the day of the test itself, students reported their emotional experiences analogous to the predictions, i.e. immediately before, immediately after, and one hour after the test. To measure criteria of adaptability or emotional management, participants completed a coping inventory and reported their emotional distress one week prior to the main examinations. For reasons of data protection, there was no data on actual grades available but participants rated the satisfaction with their grade in the social science test.

3.1.3 Material

Psychometric Tests. *Intelligence* was measured with the short version of the Culture Fair Intelligence Test (CFT 3, German adaptation Cattell & Weiß, 1971) which focuses on general fluid intelligence. The item material consists

Table 3.1: Anxiety Study: Course of the Data Collection

Date	Landmarks and <i>Tests Conducted</i>
2 nd Week in October	Beginning of the School Program
1 st Week in November	<i>Demographic Data</i> <i>Mood</i> <i>Emotional Predictions</i> <i>Intelligence</i> <i>Personality</i> <i>Trait Anxiety</i>
1 st Week in December	Social Science test <i>Emotional Experiences</i> (before, after, one hour later)
2 nd Week in December	<i>Coping</i>
3 rd Week in December	<i>Distress</i>
4 th Week in December	Examinations
1 st Week in January	Results of the Social-Science-Test <i>Satisfaction with the Grade</i>
2 nd Week in January	<i>Emotional Intelligence</i>

of figural matrices and takes approximately 15 minutes. The test provides two parallel forms. Here, the short version of test form A was administered.

Emotional intelligence was measured with the The Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) V 2.0 (Mayer et al., 2001), an instrument based on the ability approach of emotional intelligence introduced by Salovey and Mayer (1990). Here, the German translation was used (Schütz, Hertel, & Schröder, 2002). It entails 141 items measuring eight subscales; two four each branch: emotional perception, emotional facilitation, emotional understanding, and emotional management. The test scores are calculated based on a consensus-scoring of an accumulated German sample (N=295). Internal consistencies of the branches were $\alpha = .88$ for perception, $\alpha = .67$ for facilitation, $\alpha = .68$ for understanding and $\alpha = .54$ for management. Mayer et al. (2003) claim, that the MSCEIT items were het-

erogeneous and that consistencies underestimate the reliability as opposed to retests.

Personality was measured with the Big Five Inventory (BFI, Lang, Luedtke, & Asendorpf, 2001). It measures neuroticism, extraversion, conscientiousness, agreeableness, and openness with 42 items in five to ten minutes. Additionally, trait anxiety was measured using the German version of the State-Trait-Anxiety-Inventory (STAI, Laux, Glanzmann, Schaffner, & Spielberger, 1981). The scale consists of 20 items and takes approximately 5 minutes.

Emotional Predictions. The participants were asked to imagine the upcoming social science test and predict how they themselves and the average classmate will feel immediately *before* the test, *after*, and one hour *later*. The ratings were conducted with 10 affective adjectives on a scale ranging from 1 (not at all) to 5 (absolutely). These adjectives were intended to cover a variety of possible affective reactions. The multidimensional state questionnaire (MBDF, Steyer, Schwenkmezger, Notz, & Eid, 1997) measures three affective dimensions: valence, tranquility and wakefulness. These dimensions can be compared to the affective aspect of stress as measured with the Distress Scale of the DSSQ (Matthews et al., 2002). For each dimension, the two items with the highest part-whole correlations were chosen. Valence is represented by happy and unsatisfied (cf. hedonic tone); tranquility by calm and nervous (cf. tense arousal); and wakefulness by awake and tired (cf. energetic arousal). Additionally, the subjects predicted four relevant emotional categories, namely anger, pride, anxiety, and sadness.

Accuracy of emotional predictions was calculated in terms of bias and correspondence of the predictions to the experiences (see Section 3.2). An oblimin principal axis analysis over the aggregated experience scores shows a strong first factor. A general affective index accounts for 38.19% of the variance with a consistency of $\alpha = .83$. A two factor solution explains 47.7% of the variance and divides the affective adjectives into positive affect (pride,

relaxed, happy, calm) and negative affect (anxious, unsatisfied, nervous, angry, sad and tired). Those factors intercorrelated with $-.23$ and had a consistency of $.64$ for positive affect and $.84$ for negative affect.

Emotional Experiences. Emotional experiences were measured as self reports with the same 10 affective adjectives. Experiences of the general other were calculated as the mean experience per item.

Coping. Coping was measured using six of the 20 subscales of the questionnaire for coping with stress (SVF 78, Janke & Erdmann, 2002). Each of these subscales contained six items on a 5-point scale. The standard instruction asks for coping with stressors in general. Here, participants were asked how they handle stress in regard to the upcoming exams. The authors propose a distinction between positive and negative strategies in terms of attenuation and enhancement of stress. Following this logic, the conducted subscales represent four positive strategies (positive self-instruction, playing down, tone down, and reaction control) and two negative ones (resignation, rumination). Sample items are given in Table 3.2. An oblimin principal axis analysis of the scale scores supports a slightly different two factor solution with 66.88% variance explanation. Here, playing down (reversely coded) shifts to the negative strategies. The factor intercorrelation between positive and negative strategies is $.16$, all communalities are between $.58 - .79$ and all loadings are over $.70$. Indices of positive and negative coping are calculated accordingly 3.2. Internal consistencies are $\alpha = .74$ and $\alpha = .71$.

Distress prior to the Exams. This distress measure was measured independently from predictions and experiences in anticipation of the exams. The already described shortened version of the MBDF (Steyer et al., 1997) was conducted one week prior to the exams in order to measure valence, tranquility, and calmness. The one-factor solution in an oblimin principal axis analysis accounts for 53.10 % of the variance and the consistency of this index is $\alpha=.87$.

Table 3.2: Sample Items for Positive and Negative Coping

Scale	Sample Item
Positive Strategies	
Positive Self-Instruction	I tell myself that I won't give up.
Reaction Controll	I try to pull myself together.
Tone Down	I tell myself that it is nothing to worry about.
Negative Strategies	
Playing Down (reverse coded)	I manage the situation better than most others.
Rumination	I tend to think about it over and over.
Resignation	I tend to give up quickly.

Note. Own translation.

Satisfaction with Grades. Participants received their grade in the social science test in January. They rated the satisfaction with their grade on a five-point scale.

3.2 Results

First, descriptive statistics and first order correlations of the conducted tests and ratings are presented. Then, repeated measures ANOVA is used to evaluate the data in terms of the affective forecasting paradigm (Gilbert et al., 1998). Then, individual differences in the accuracy of emotional predictions are explored. After addressing validity and adaptive value of bias, correspondence as profile similarity will be examined.

3.2.1 Descriptive Statistics

Table 3.3 presents the intercorrelations and descriptive statistics of the conducted measurements. For predictions and experiences composite scores are reported¹ However, the correlational pattern is comparable when positive

¹The correlation between positive and negative affect is $r=.52$ for predictions and $r=.46$ for experiences. Different points of time (before, after, later) correlate with a range of $r=.66$ to $r=.82$ for predictions and $r=.55$ to $r=.66$ for experiences.

and negative affect are distinguished. The relationship between predicted and experienced affect is fairly high, already indicating substantial correspondence. Also, as expected in terms of discriminant validity, one's own experiences are more closely related to the predictions for the self than to those for others.

As to be expected, participants who predict more positive emotional experiences score lower on neuroticism and state anxiety. Those predictions are given in a better mood, and they are associated with less negative coping strategies and less distress. Positive coping strategies are unrelated. This pattern repeats itself to a lesser degree with predictions for others. Also, more positive predictions for others are associated with higher overall emotional intelligence. Emotional intelligence scores are higher for women and intelligence scores higher for men. Coping strategies are related to distress and higher emotional intelligence goes along with higher satisfaction with the grade.

When predictions for different points of time are distinguished, relationships to personality and mood are comparable. Only the correlation to coping is stronger for earlier points of time, i.e. participants with more positive coping strategies expect to feel better, especially before and immediately after the test.

3.2.2 Forecasts and Experiences over Time

The course of predictions and experiences over time is examined with a 3×2 repeated measures ANOVA, treating time (before, after, later) and type (prediction, experience) as factors, and emotional ratings as outcomes (see Figure 3.1). Positive and negative affect are distinguished because they are distinguished in affective forecasting research (Gilbert et al., 2002, 1998). The within subjects effects and the contrasts are presented in Table 3.4. Repeated contrasts are reported for *time* and simple contrasts for *type*. Time

Table 3.3: Intercorrelations and Descriptive Statistics

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
1. Prediction Self ^a	.61**																						
2. Prediction Others	.59**	.39**																					
3. Experience Self	.12	-.07	.02																				
4. Age	.11	.03	.22*	.26**																			
5. Gender ^b	.03	-.10	-.02	-.28**	-.03																		
6. Intelligence	.10	.21*	.13	-.05	-.24**	.04																	
7. EI Overall ^c	-.12	.02	-.03	-.07	-.31**	-.04	.48**																
8. EI Perception	-.05	-.01	.05	-.12	-.22**	.25**	.38**	.54**															
9. EI Facilitation	-.04	-.02	.10	.05	-.30**	.01	.34**	.45**	.53**														
10. EI Understanding	-.03	.07	.08	-.07	-.34**	.10	.73**	.79**	.82**	.72**													
11. EI Management	-.37**	-.14	-.32**	-.14	-.25**	-.11	.07	.19*	.07	.11	.14												
12. Neuroticism	.10	.17*	.15	-.07	-.11	.02	.26**	.14	.11	.04	.19*	-.15											
13. Extraversion	-.03	-.05	.05	.28**	-.07	-.29**	.07	.16	.09	.18*	.15	-.13	.14										
14. Conscientiousness	-.07	-.13	-.02	.10	-.03	.01	.02	.09	.11	.20*	.13	-.24**	-.11	.32**									
15. Agreeableness	.17*	.13	.14	-.06	-.15	.25**	.23**	.21*	.37**	.04	.30**	-.03	.27**	.07	.11								
16. Openness	-.46**	-.30**	-.35**	-.20*	-.21*	-.08	.00	.14	.11	.02	.09	.71**	-.30**	-.21*	-.18*	-.03							
17. State Anxiety	.50**	.31**	.34**	.17*	.11	-.06	-.01	-.20*	-.18*	-.14	-.17*	-.36**	.23**	.24**	.13	.02	-.52**						
18. Mood	.10	-.01	.22*	.02	.05	-.14	-.03	-.08	.01	-.03	-.04	-.23*	-.08	.02	.10	.02	-.14	.06					
19. Positive Coping	-.50**	-.12	-.37**	-.17*	-.20*	.04	-.03	.13	.04	.01	.04	.59**	-.08	-.10	-.08	-.03	.45**	-.22*	-.25**				
20. Negative Coping	-.45**	-.12	-.48**	-.26**	-.26**	.07	.00	.15	.05	.01	.07	.39**	-.12	-.18*	-.08	-.01	.43**	-.33**	-.14	.50**			
21. Distress	.13	.10	.28**	-.06	-.06	-.04	.21*	.17*	.13	.12	.20*	.06	.05	.05	.07	.08	-.13	.12	.08	.01	-.06		
22. Satisfaction Grade	3.25	3.15	3.21	24.27	.44	27.55	.42	.40	.55	.36	.43	18.74	27.66	32.62	28.57	33.52	37.59	3.39	14.13	9.96	3.33	4.04	
Mean	.56	.46	.53	5.51	.50	4.22	.08	.06	.08	.06	.05	4.40	5.28	4.87	4.10	5.88	9.02	.72	3.29	3.64	.81	1.04	
SD	1.31	1.31	1.39	1.43	1.43	1.43	1.38	1.38	1.38	1.38	1.38	1.43	1.43	1.43	1.43	1.43	1.43	1.38	1.28	1.28	1.35	1.33	
N																							

^a Composite scores for affective predictions and experiences are used (over items and points of time). High scores indicate positive affect. Experiences of the general other are defined as the mean experience.

^b Positive correlations with gender indicate higher values for males.

^c EI: Emotional Intelligence

* $p < .05$

** $p < .01$, two-tailed

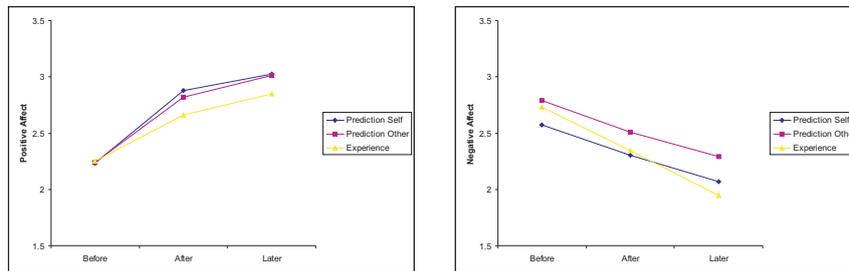


Figure 3.1: Course of Predictions and Experiences ($N = 128 - 141$)

(before, after, later) has a significant main effect. Positive affect increases and negative affect decreases over time for both transitions before vs. after and after vs. later. Only for positive affect, there is a main effect for type (predictions, experiences), i.e. predictions are better than experiences.

The interaction of *time* and *type* is significant for positive and negative affect. This is mainly due to predictions and experiences drifting apart for later predictions. Specifically, participants expect better feelings after an exam than they experience. For positive affect, this is due to the first time transition. It is already overestimated after the test and again one hour later. One's own negative affect is overestimated *before* the test, rather accurately predicted after, and underestimated one hour later. This replicates the findings of Gilbert et al. (1998) that negative affect decreases faster than expected ($\eta = .06$). Also, negative affect predictions for others are generally worse than predictions for the self. They are even worse than the actual experiences.

3.2.3 Accuracy as Bias

As previously stated, accuracy can be conceptualized as *bias* (i.e. Euclidean distance) and *correspondence* (i.e. profile correlation between predictions and experiences). In terms of the accuracy components of Cronbach (1955), bias as Euclidean distance is analogous to *elevation*. After examining the mean differences between forecasts and experiences, individual differences

Table 3.4: Repeated-measures ANOVA and Contrasts with Positive and Negative Affect as Outcome Variables

Source of variance	<i>F</i>	<i>df</i>	η
Positive affect			
Time	120.66***	1.83	.51
Before vs After	104.99***	1	.48
After vs Delayed	18.66***	1	.14
Type	5.41*	1	.05
Time \times Type	6.95**	1.95	.06
Before vs. After \times			
Forecast vs. Experience	10.59**	1	.8
After vs. Delayed \times			
Forecast vs. Experience	.42	1	<.01
Negative affect			
Time	123.39***	1.78	.52
Before vs After	62.71***	1	.35
After vs Delayed	86.53***	1	.43
Type	.25	1	<.01
Time \times Type	9.56***	1.25	.08
Before vs. After \times			
Forecast vs. Experience	2.86	1	.02
After vs. Delayed \times			
Forecast vs. Experience	7.69**	1	.06

Note. Results are corrected following Greenhouse-Geisser. Contrasts for *type* are not reported because it has two levels: predictions and experiences.

* $p < .05$

** $p < .01$

*** $p < .001$, two tailed

in accuracy are explored. For each participant a bias score is calculated as the Euclidean distance, i.e. the sum of squared differences. Correlations are given in Table 3.5. Euclidean distances disregard whether affect is over- or underestimated, and they accentuate larger deviations. In contrast, simple differences are sensitive to over- and underestimation. Higher simple differences indicate overestimation of positive affect, i.e. the predicted affect is better than the experienced one.

Results show that bias is higher for women and for participants with

higher trait anxiety and neuroticism. Also, it is associated with more negative coping and distress. Interestingly, emotionally perceptive participants show a higher bias. This relationship is partly due to shared variance between emotional perception and neuroticism ($r = .19$, see Table 3.3). For simple differences, the correlational pattern is distinct, and the correlations are weaker in general. Higher simple differences, i.e. expecting better affect than is experienced, is associated to low neuroticism, better mood, and less negative coping strategies. Also, older participants tend to expect better affect than they experience.

When different points of time (before, after, later) are distinguished, correlational patterns are similar. The correlations, however, tend to be stronger for later predictions. Also, gender is not associated to bias for the before prediction but to predictions for subsequent points of time ($-.21$ for after, $-.28$ for later, $p < .01$).

To examine incremental validity of bias over and above intelligence and personality, hierarchical regression analyses are conducted with the criteria as outcome variables. In a first step, age, gender, intelligence, and personality were entered in the regression equation. In a second step, bias was added². The inclusion of bias did not improve the prediction of positive coping ($\beta = .02, \Delta R^2 < .01, p = .84$) or satisfaction with the grade ($\beta = -.10, \Delta R^2 < .01, p = .34$). It does show incremental validity in predicting negative coping and stress. Initially, negative coping is best predicted with low neuroticism ($R^2 = .36, p < .001$). Bias improves the model ($\beta = .22, \Delta R^2 = .04, p = .01$) in that low bias is related to less negative coping over and above other abilities and traits. For distress the model improvement is also significant. Low bias is incrementally related to less distress ($\beta = .20, \Delta R^2 = .04, p = .02$). In contrast, the simple difference scores improve none of the models (all $\Delta R^2 < .01, p > .13$).

²Trait anxiety was not included because of high covariation with neuroticism. If anxiety is entered instead of neuroticism, results are similar.

Table 3.5: First-order Correlations of Difference Scores and Criteria

	Bias	Simple Difference ^a
Age	-.10	.19*
Gender ^b	-.27**	-.03
Intelligence	.03	-.09
Emotional Intelligence		
Overall	.25**	-.13
Perception	.28**	-.08
Facilitation	.16	-.11
Understanding	.15	-.08
Management	.14	-.15
Personality		
State Anxiety	.36***	-.15
Neuroticism	.32***	-.19*
Extraversion	-.04	.02
Conscientiousness	-.12	.12
Agreeableness	-.10	.04
Openness	.11	-.07
Mood while Forecasting	-.19*	.21*
Coping		
Positive Strategies	-.17	.05
Negative Strategies	.36***	-.20*
Distress	.30***	-.14
Satisfaction with Grade	-.06	-.16

Note. Bias is log-transformed to reach normality.

^a High simple differences indicate that participants predict better affect than they experience.

^b Positive correlations with gender indicate higher values for males.

* $p < .05$

** $p < .01$

*** $p < .001$, two tailed

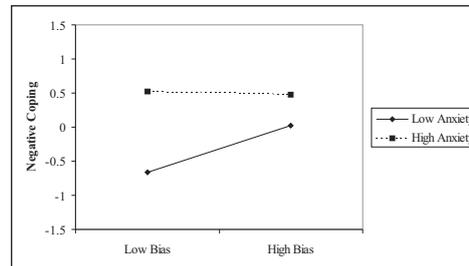


Figure 3.2: Anxiety as a Moderator of the Relationship between Bias and Coping

Anxiety as a moderator of the relationship between bias and criteria is explored by adding an interaction term. For negative affect, the relationship between bias and negative coping is moderated by anxiety ($R^2=.24$, $\Delta R^2=.04$ for the interaction term, $p < .05$). Figure 3.2 clarifies the nature of this interaction. For low anxiety individuals, lower bias is associated with less negative coping. High anxiety participants show negative coping regardless of a high or low bias in their predictions. A similar interaction effect can be found for distress ($R^2=.2$, $\Delta R^2=.03$ for the interaction term, $p < .05$) but not for positive coping and satisfaction (all $\Delta R^2 < .01$, all $p > .39$).

3.2.4 Accuracy as Correspondence

In terms of the accuracy components of Cronbach (1955), correspondence is a composite measure that can be corrected for *stereotype accuracy* and *differential elevation accuracy* in order to obtain *differential accuracy*. Stereotype accuracy captures the accuracy as to the mean answers for certain emotions, e.g. that people are generally more nervous and less calm. Differential Elevation Accuracy describes the accuracy as to the prediction of a response biases for certain targets, e.g. that someone is generally in a worse mood than others. *Differential accuracy* means the accuracy due to unique predictions for certain targets on certain emotions, e.g. that someone will be tired and nervous, as opposed to others). The second study of this thesis allows for the full decomposition of accuracy but the current design limits it. Since

there is only one specific target (the self), differential elevation accuracy can not be examined. Stereotype accuracy, however, can be calculated as the correspondence of predictions with the mean profile of experiences.

Correspondence is calculated using a multilevel modeling approach (Raudenbush & Bryk, 2002; Kreft & de Leeuw, 1998) and the software HLM 6.0 (Raudenbush, Bryk, Cheong, & Congdon, 2004). For accuracy estimates that are sample independent, forecast were group-mean centered (Kreft, de Leeuw, & Aiken, 1995) and person characteristics were grand mean centered. The parameter estimates are comparable to unstandardized regression coefficients. They are therefore not directly comparable. One of the main problems in interpreting multilevel models, is that error variance is not stable. With the addition of a new parameter, this variance can change. Standardization of variables is therefore not recommended (Nezlek, 2001). For effect size estimates, *proportion reduction of error* is available and comparable to ΔR^2 . Also, the deviance test can be used to evaluate comparative model fit.

Correspondence reflects how profiles of affective states can be accurately predicted. In the HLM model, the data is treated as hierarchically nested. Specifically, measures (level 1) are nested within persons (level 2). The emotional experiences (EXP) are treated as the dependent variable in a series of multilevel random coefficient models. The basic model is referred to as the unrestricted model (see Equation 3.1a for level 1 and Equation 3.1b for level 2).

$$EXP_{ij} = \beta_{0j} + r_{ij} \quad (3.1a)$$

Here, EXP_{ij} is the emotional experience of person i on measurement j . β_{0j} is the random intercept coefficient for person j across the measures and r_{ij} represents the error component of the measures j for person i . Basically, it is assumed that each person has a different mean of emotional experiences

that is estimated with β_{0j} . In multilevel modeling, β_{0j} can vary between individuals. This is modeled as a random error component on level 2 (u_{0j}), i.e. the person level.

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3.1b)$$

In Equation 3.1b, β_{0j} , as the known intercept component of level 1 is predicted by a general mean of experiences (γ_{00}) and a person specific error component (u_{0j}).

As shown in Table 3.6, the mean experience is estimated as 3.43 on a scale from 1 to 5 with high scores indicating positive feelings. The variance in the measures is 1.25 ($SD = 1.12$) and person level variance in emotion ratings is .21 ($SD = .46$). This means that within-measures variance accounted for approximately 83% of the overall variance which corresponds to an intraclass correlation of .17. This indicates considerable variability of emotion ratings but still dependency in the measures due to persons. The average reliability estimate (Raudenbush & Bryk, 2002, p. 49) for the mean personal experience is .75.

Table 3.6: One-way ANOVA with Random Effects. Unrestricted Model

Fixed effect	Coefficient	SE	t Ratio	p Value
Average personal experience, γ_{00}	3.43	.05	66.66	<.001
Random effect	Variance component	df	χ^2	p Value
of personal experience, u_{0j}	.21	111	459.87	<.001
of level 1 (measures), r_{ij}	1.25			

Correspondence between predictions and experiences can be evaluated when predictions (P) are used to estimate experiences:

$$EXP_{ij} = \beta_{0j} + \beta_{1j}P_{ij} + r_{ij} \quad (3.2a)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3.2b)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (3.2c)$$

Here, both the intercept (β_{0j}) and the slope of the predictions (β_{1j}) are modeled as random on level 2 (3.2b, 3.2c), i.e. they are allowed to vary for each person. The results of this model are given in Table 3.7. The fixed effects part of the results shows a similar intercept to the unrestricted model³. The average prediction-experience-slope is .55⁴. This indicates that forecasts which are one point higher on a five point scale are associated with experiences that are increased by .55. Both intercept and slope are significantly different from zero. The random effects in this model represent the variances of mean experiences and prediction-experience-slopes across the participants. There is considerable variability in both the intercepts and the slopes. This is necessary in order to relate this variability to other traits.

The random effect on level 1, or residual variance, drops from 1.25 in the unrestricted model to .84 after the inclusion of predictions. This proportion of reduction of error (PRE) that is due to the inclusion of predictions is approximately 33% which indicates considerable improvement of the model. Reliability estimates are .84 for the intercept and .41 for the slope. Slopes generally show lower reliability and values over .2 are called sufficient by Raudenbush and Bryk (2002).

The next step is to take a look at moderators of accuracy, i.e. whether certain person characteristics lead to a higher correspondence between predictions and experiences. Such characteristics are modeled on the level 2 (person) to estimate person-specific intercepts β_{0j} and slopes β_{1j} . For the example of gender (GEN), following model is estimated:

³The intercept estimate does not change since the predictions are group-centered, i.e. centered for each person.

⁴The slope coefficient can be interpreted as an unstandardized regression coefficient.

Table 3.7: Estimating Experiences with Predictions

Fixed effect	Coefficient	SE	t Ratio	p Value
Mean personal experience, γ_{00}	3.42	.05	66.43	<.001
Mean prediction-experience-slope, γ_{00}	.55	.03	20.55	<.001
Random effect	Variance component	df	χ^2	p Value
Personal experience, u_{0j}	.24	104	666.22	<.001
Forecast-experience-slope, u_{1j}	.03	104	182.65	<.001
level 1 (measures), r_{ij}	.84			

$$EXP_{ij} = \beta_{0j} + \beta_{1j}P_{ij} + r_{ij} \quad (3.3a)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}GEN + u_{0j} \quad (3.3b)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}GEN + u_{1j} \quad (3.3c)$$

Such a model is called intercepts-and-slopes-as-outcomes model because both intercept and slope are predicted with level 2 variables. A significant parameter γ_{01} indicates that gender relates to the mean affective experience, whereas a significant parameter γ_{11} indicates that gender moderates the strength of the relationship between forecast and experience, i.e. the accuracy in terms of correspondence.

Table 3.8 summarizes the results for a series of models, testing one moderator at a time. The significance tests are comparable to those of correlations but the coefficients are not standardized. Only emotional perception is associated with the prediction-experience slope. Beyond that, correspondence does not significantly relate to the available intelligence and personality measures. Also, it does not relate to satisfaction with the achieved grade, coping, or distress in the examination phase.

To evaluate the moderators in context and to further investigate the role of emotional perception, hierarchical intercepts-and-slopes-as-outcomes

Table 3.8: Moderators of Correspondence

Moderator	Fixed Slope Coefficient	% PRE ^a in slope variance
Age	<.01	.2
Gender ^b	-.06	<.1
Emotional Intelligence		
Perception	.93*	14.7*
Facilitation	.62	2.2
Understanding	.33	1.1
Management	.28	.8
Overall	.92	5.9
Intelligence	<.01	<.1
Personality		
State Anxiety	<.01	.2
Neuroticism	<.01	<.1
Extraversion	<.01	.5
Conscientiousness	<.01	.2
Agreeableness	<.01	1.6
Openness	<.01	<.1
Coping		
Positive Strategies	<.01	.9
Negative Strategies	<.01	.5
Mood while forecasting		
General Index (Pos)	-.04	<.1
Satisfaction with grade	<.01	<.1
Distress Exams	.05	4.7

^a Proportion of reduction of error compared to the model without moderator.

^b Positive correlations with gender indicate higher values for males.

* $p < .05$, two tailed

models are conducted. First, age, gender, intelligence, and personality are used as predictors (step 1). Then, emotional perception is added (step 2). Again, state anxiety is omitted because of its large covariation with neuroticism. When the intercepts are outcomes, neuroticism is again a strong predictor of mean negative experience. Furthermore, openness covaries with better experiences, and males tend to report slightly better affect than females. Emotional perception however does not add to that prediction. When the slopes are outcomes (i.e. the correspondence between predictions

and experiences), age, gender, intelligence, and personality are again unrelated. Here, emotional perception as an additional predictor is significant ($b = 1.15, t = 2.920, p < .01$). It reduces the error variance in the slopes considerably (PRE= 17.3 %) and leads to significant model improvement ($\chi^2(2) = 9.523, p = .009$).

Stereotype accuracy refers to the correspondence between predictions and mean experiences. When correspondence is *adjusted* for stereotype accuracy, mean experiences are subtracted from the experiences, and mean predictions are subtracted from the predictions. Then, correspondence is again calculated. When emotional perception is used to predict this adjusted correspondence, the parameter is nonsignificant ($b < .01, t < .01$) and does not lead to model improvement ($\chi^2(1) < .001, p > .5$). This indicates that emotional perception, as measured in the MSCEIT, is associated with higher correspondence due to a match with the stereotypic profile (i.e. the mean profile).

The last step is to evaluate the adaptive value of correspondence as to coping, distress, and satisfaction. Multilevel random coefficient modeling only allows for dependent variables on level 1. Therefore these criteria cannot be formally regressed on. Still, the covariation between correspondence and criteria can be evaluated when criteria are modeled as moderators of the prediction-experience slope. Table 3.9 shows the results. Negative coping, distress, and low satisfaction significantly predict worse experience (see intercepts-as-outcomes), but none of these characteristics is associated with a higher correspondence (see slopes-as-outcomes) Thus, there is no evidence that high correspondence is adaptive.

Again, anxiety is addressed as a moderator for the relationship between criteria and correspondence. An interaction term is implemented on level 2 (cf. Preacher, Curran, & Bauer, 2006), and for the example of distress (DIS) following model is estimated:

Table 3.9: Adaptability of Correspondence

Fixed effects	Coefficient	t ratio
Intercepts-as-Outcomes		
Intercept	3.42***	77.53
Age	-.01	-1.66
Gender ^a	.20*	2.39
Positive Coping	-.00	-.15
Negative Coping	-.04*	-2.42
Distress	-.12*	-2.03
Satisfaction	.12**	2.75
Slopes-as-Outcomes		
Intercept	.55***	20.98
Age	<.01	.14
Gender	-.02	-.41
Positive Coping	-.01	-.93
Negative Coping	-.01	-1.36
Distress	.05	1.27
Satisfaction	.01	.60

^a Positive effects with gender indicate higher values for males.

* $p < .05$

** $p < .01$

*** $p < .001$, two tailed

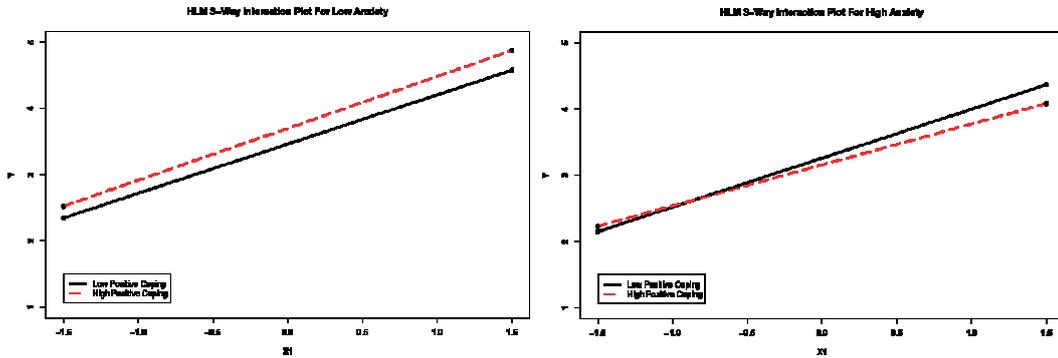


Figure 3.3: Relationship between Emotional Predictions (x) and Experiences (y) for High and Low Anxiety, and High and Low Positive Coping

$$EXP_{ij} = \beta_{0j} + \beta_{1j}P_{ij} + r_{ij} \quad (3.4a)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}ANX + \gamma_{02}DIS + \gamma_{03}ANXxDIS + u_{0j} \quad (3.4b)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}ANX + \gamma_{12}DIS + \gamma_{13}ANXxDIS + u_{1j} \quad (3.4c)$$

The term γ_{13} is tested for the interactions between anxiety and each of the criteria (positive coping, negative coping, distress, and satisfaction). Results show a significant interaction for anxiety and positive coping ($b < -.01$, $t = -2.32$, $p = .02$, see Figure 3.3). For high anxiety individuals, less correspondence relates to higher positive coping, again due to correspondence with the stereotypic experience. In other words, when trait anxiety is high, not predicting the course of emotions that everyone else does is beneficial. For low trait anxiety, in contrast, high positive coping is associated with better experiences (y) but unrelated to correspondence. For the other criteria, interactions with anxiety are not significant (negative coping: $b < .01$, $t = .31$, $p = .76$, distress: $b < .01$, $t = .73$, $p = .47$, satisfaction with the grade: $b < .01$, $t = -1.04$, $p = .30$). Thus, anxious participants with higher correspondence do not show more or less negative coping, distress, or satisfaction with the grade than non-anxious and accurate participants.

3.3 Discussion

The main research question is the nature (3.3.1) and adaptive value (3.3.2) of accuracy of emotional predictions in terms of a low bias and high correspondence. Then, the findings will be connected to the affective forecasting paradigm (3.3.3).

3.3.1 Bias, Correspondence, and other Abilities and Traits

Predictions and experiences, as the basis of accuracy, are strongly related to mood and personality and rather unrelated to intelligence or emotional intelligence. Not surprisingly, anxious participants and participants in a bad mood predicted worse affective responses to the test. Mood guides the retrieval from memory and activates aspects of the self that are mood congruent (f.e. Loewenstein & Lerner, 2003). The imagination of the upcoming test itself could not have influenced the mood since mood was measured before the emotional predictions were introduced.

Accuracy is distinguished as to bias and correspondence. Lower bias is unrelated to intelligence, relates to personality in terms of lower trait anxiety and neuroticism. It was not related to emotional understanding but to emotional perception. Moreover, higher emotional perception was related to a higher bias or larger deviations between predictions and experiences. Common variance between emotional perception and neuroticism/anxiety⁵ can explain this relationship (see Table 3.3). Still, this undermines the assumption of low bias as a proxy measure of emotional intelligence in general, or emotional understanding in particular. There is, however, another study that links emotional perception to stress for social service workers (Ciarrochi, Dean, & Anderson, 2002). If emotional perception can be maladaptive, it should be in a context that elicits negative responses. The simple differ-

⁵This relationship is specific to the sample at hand. A current meta-analysis reports a mean correlation between emotional perception and emotional stability of .12 (Joseph & Newman, 2010).

ences show better discriminant validity to personality but no substantial associations to EI either.

Correspondence shows slightly better construct validity. It was unrelated to intelligence, as well, but also unrelated to personality traits and positively associated with emotional perception. Specifically, stereotype accuracy, i.e. the correspondence of predictions to mean affective profiles is higher for emotionally perceptive individuals. This indicates that a focus on correspondence might be better suited to reflect an ability as to emotions. However, the relationship to emotional understanding is also not significant. In sum, emotionally perceptive individuals accentuate the intensity compared to the sample mean (which leads to higher bias), but they are better in predicting the constellation of different affective dimensions and different points of time.

3.3.2 The Adaptability of Bias and Correspondence

Low bias is in fact adaptive and predicts lower distress and less negative coping, even incrementally to intelligence and personality. Also, it is a stronger predictor of criteria than the simple difference score. It is not just about an optimistic view but about accurate knowledge of intensity, also over time. However, low bias as to an anxiety-eliciting situation does not relate to criteria of positive affective value, like improved satisfaction with the grade or more positive coping.

Trait anxiety moderates the relationship between bias and negative coping. Participants with high anxiety show negative coping anyway. For low anxiety individuals, though, lower bias is beneficial in terms of e.g. less rumination and resignation. This supports the conclusion that the consequences of accuracy are person-specific. In contrast, simple difference scores are not incrementally valid to personality. Their relations to criteria can be explained with shared variance. The notion to overestimate negative affect

in part reflects anxiety.

Correspondence is not generally adaptive but anxiety serves as a moderator again. High anxiety individuals with lower correspondence show more positive coping. Thus, low accuracy in terms of correspondence (or more specifically low stereotype accuracy) is beneficial for this subgroup. High correspondence means to know about the profile of emotions that is likely to arise. Here, this also involves knowing how they develop over time. Basically, emotions are rank ordered as to their intensity. Accurately predicting the course of the emotions seems to be maladaptive for anxious participants. This supports the hypothesis of Gilbert et al. (2002) that errors can be useful and that our view on the temporal progression of emotions should be distorted. Under certain conditions, this seems to be that case.

3.3.3 Connection to the Affective Forecasting Paradigm

When forecasts and experiences are examined on a group level, there is clear evidence for immune neglect (Gilbert et al., 1998). Obviously, participants underestimate mood repair processes that occur after the exam. In addition, positive affect after the event was overestimated. This can also enhance motivation and help to achieve one's goals.

On a group level, there is evidence for a general bias. A focus on interindividual differences, though, shows that the bias should not be as high as possible. Lower bias was related to less distress and less negative coping, even more so for the *after* and *later* predictions. Also, accurate predictions are again preferable to over- or underestimation as reflected in the simple difference scores. Dunn et al. (2007) already suggested that higher affective forecasting accuracy relates to ability EI, especially emotional management. They, however, used the simple differences on a nine-point scale as a measure of accuracy. Given the results of the current study, it seems possible that these correlations are at least partly due to personality and not incre-

mentally valid.

Correspondence in the anxiety study captured how well the predicted emotional transitions for different affective dimensions matched the experienced ones. Results suggest that high anxiety individuals cope better with less knowledge about the transitions. This can be explained with motivated distortions (Gilbert et al., 2002). Playing down the negative impact of the test situation leads to a smaller drop of negative affect after the test which reduces correspondence.

In this study, the focus was on the self as the target of prediction. Given the impact of anxiety on the accuracy-criterion relationship, the examination of other targets seems crucial. Affective forecasting research emphasizes biases in the prediction of one's own future emotional states. Emotional intelligence, however, claims that emotional knowledge should be employed for self and others and the measurement of emotional understanding includes a test that uses vignettes for hypothetical others. The following study will focus on the social aspect of emotional predictions.

Chapter 4

Study 2: Work-Group Study

4.1 Methods

4.1.1 Sample

The study was conducted with psychology freshman at Humboldt-University Berlin. The subjects were participants of an empirical observation course in two consecutive years. The aim of the course was to gain experience in empirical research. Students were expected to plan, conduct, and analyze observational data in study groups of 3-6 participants. Data was collected over the course of the semester, i.e. approximately over 6 months. There was sample attrition due to discontinuation of the course or of the undergraduate studies. 207 students in 43 groups completed the course, 102 in 21 groups in the first year and 105 in 22 groups in the second year. 78.2 % were female with a mean age of 22.9 (SD = 5.55, range = 18 - 50). Participation was rewarded with detailed feedback on the conducted tests and credits for partaking in an empirical study. The compliance with the study requirements was good in general, although there was a considerable amount of missing data due to nonattendance in certain waves of the data collection. The N for different measures varied from 158 to 180 in 43 groups. Peer

ratings were available for 195 students.

4.1.2 Design and Procedure

Students were randomly assigned to work groups in order to prevent self-selection based on mutual affection. The random procedure also promoted zero acquaintance (cf. Kenny, 1994) in the work groups to avoid advantages as to peer predictions. Accordingly, none of the group members knew another member longer than for two weeks. First, participants took standardized psychometric tests under the supervision of an investigator. During the semester, the students completed questionnaires that asked for emotional predictions regarding relationship variables and their collaboration in the group. Participants were allowed to complete the questionnaires at home and return them to a mailbox at university. The ratings followed a round robin design, meaning that everyone answered each question for every other member of the work group. This way, different specific targets of predictions can be distinguished. Also, teacher ratings for performance and the grade of the term paper were available.

4.1.3 Material

This study was part of a larger research cooperation that also analyzed personality development. In the second year of the study, some questionnaires had to be shortened due to time constraints in the course. For the current thesis, the following tests and questionnaires are relevant.

Psychometric tests. *Intelligence* was measured with the short version of the Berlin Structure of Intelligence Test (BIS-K IV, Jäger, Süß, & Beauducel, 1997) that lasts approximately 45 minutes. Each item belongs to a certain content domain (verbal, numerical, and spatial) and to a certain operational domain (capacity, memory, speed, and creativity). A test score for general intelligence and for intellectual capacity can be calculated. *General knowl-*

edge was measured with a shortened version of the knowledge test available in the Intelligence-Structure-Test 2000R (IST-2000, Amthauer, Brocke, Liepmann, & Beauducel, 2001). For the shortened version, the 18 items with the highest part-whole-correlation were selected. Every correct item was awarded with one point. A maximum of 18 points was attainable.

Emotional intelligence was measured with the MSCEIT V2.0, an instrument that represents the ability approach of emotional intelligence. Here, four branches of emotional intelligence are distinguished, namely emotional perception, emotional facilitation, emotional understanding and emotional management. The test was consensus-scored based on a German sample with approximately 400 participants.

Personality was measured using the BIG5 inventory (BFI, Lang et al., 2001). It contains 42 items to measure neuroticism, extraversion, conscientiousness, agreeableness, and openness. It takes approximately 5-10 minutes to administer.

Emotional predictions. Emotional predictions regarding mutual feelings at the end of the term were conducted twice. First, the questionnaire was completed with roughly a month of acquaintance and during the planning stages of the student project. It was repeated after another eight weeks and with more experience of collaboration. Retest reliability can be calculated. Here, emotional predictions focus on the relationship development in the work groups. Participants predicted affection, satisfaction with the collaboration, anger, and fun that every group member was going to experience with every other member at the end of the term (when presentations for the project were due). Anger and fun were rated on 5-point scales (1: not at all to 5: absolutely), affection and satisfaction with the collaboration on 6-point bipolar scales (dislikable to likable, and unsatisfied to satisfied).

Emotional experiences. At the end of the term, after the presentations, experiences were measured analogously. Every group member reported on

their actual feelings towards every other member of the group as to affection, satisfaction with the collaboration, anger, and fun.

Relationship measures. In the beginning and at the end of the semester, participants were asked to rate their relationship to the other group members in detail. In the current analysis, the ratings at the end of the semester are used as criteria. The items were administered in the same questionnaire as the emotional experiences. Measures of task and relationship conflict were adopted from Jehn (1995) and slightly rephrased in order to address individuals instead of groups. The original item asked whether there were personal difficulties in the team. The adapted version asked for personal difficulties with a specific person. Each type of conflict was measured with four items on a 5-point scale (*none to a lot*). Items for relationship conflict covered the amount of friction, personal conflict, tension, and emotional conflict. Task conflict contained disagreements, conflict of ideas, conflict about work and differences of opinion. An oblimin factor analysis confirmed two factors that account for 71.6 % of the variance (56.3 for relationship conflict and an additional 15.3 for task conflict)¹. The factor intercorrelation is $r = .45$.

Four additional items were administered, asking for *friendship* (s/he is my friend, I spend spare time with him/her) and *competency* (s/he did good work, s/he is competent as to what we do). Items were rated on a 5-point scale (*no to absolutely*). In an oblimin factor analysis, the scales are well distinguished with a factor intercorrelation of .02. Both scales account for 83.14 % of the variance, with 57.45 % for friendship and 25.68 % for competence.

For each team member a peer rating of task conflict, relationship conflict, friendship, and competence was calculated as the mean rating of all fellow work group members.

Teacher ratings of the term paper. Several ratings were available to as-

¹The analyses were performed on an aggregated level, i.e. for 43 groups in order to avoid dependency in the data

sess the quality of the term paper. Teachers rated the paper for quality of depiction, style, completeness and thoughtfulness on a 6-point scale with higher scores indicating better performance. Also, papers were graded on a six point scale. Oblimin principal axis analysis on the group level ($N = 43$) strongly supported one factor that accounts for 83.72 % of the variance. A composite score was calculated, and higher ratings indicate better performance.

4.2 Results

Initially, first order correlations, descriptive statistics, and retest reliabilities of emotional prediction accuracy will be addressed. Due to the round robin design of the measurements, a complete decomposition of accuracy proposed by Cronbach (1955) will be possible. This accuracy will be related to other abilities and traits. The adaptive value of accuracy will be examined as to peer ratings of friendship and competence, conflict, and quality of the term paper.

The current design differs considerably from the anxiety study. Here, participants are nested within groups. This dependency of data potentially distorts the results when individuals are the unit of analysis² (Kashy & Kenny, 2000). Three strategies were applied to deal with group level variance. First, analyses were conservatively conducted on a group level (N=43). A similar approach was taken in context of retest reliability. Since aggregation on a group level is likely to lead to overestimation of stability, random subsamples of independent group members were drawn. Then, the range of retest reliabilities was reported. Finally, multilevel modeling and variance decomposition as to the *social relation model* (SRM, Kenny, 1994) were used. SRM analyses are based on (Kenny & West, 2006) and were adapted to the conditions at hand.

4.2.1 The Group as Unit of Analysis

4.2.1.1 Descriptive Statistics

Table 4.1 presents intercorrelations and descriptive statistics for the main variables of this study except accuracy measures. Means and standard de-

²The degree of dependency in the data can be estimated using the intraclass correlation (ICC, Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). As a rule of thumb, intraclass correlations over .07 are critical. In this sample, intelligence and personality are, as to be expected, fairly unrelated (ICC below .07). However, the criteria differ considerably between groups (ICC up to .22).

viations are based on the complete sample, intercorrelations are reported on a group level ($N=43$). This controls the dependency in the data, but it reduces the power of significance tests. Therefore, significant trends are reported as well ($p < .10$).

Relationship conflict was rated very low with a mean of 1.58 on a 5 point scale. Emotional understanding and emotional management, as measured with the MSCEIT, relate to the general knowledge test, which supports the assumption of EI as an ability with emotional knowledge being the prerequisite of emotional management. Work groups with high mean emotional understanding and management also show high mean knowledge. Personality traits are rather unrelated to EI, with the exception of neuroticism, which is again associated with perception and, moreover, with facilitation.

When predicting the criteria of work group success on a group level, emotional understanding predicts the term paper rating. Emotional management is related to higher peer ratings of friendship, and facilitation to higher peer ratings of competence. Personality traits are less successful in predicting criteria. Still, in more conscientious groups, peer ratings of competence are higher. In groups with high relationship conflict, mean mutually perceived competence is lower ($r=-.77$, $p<.001$), and in groups with high task conflict, there is also relationship conflict ($r=.64$, $p<.001$). The term paper rating is rather independent from the peer ratings.

4.2.1.2 Bias and Correspondence

Table 4.2 summarizes retest reliabilities, descriptive statistics, and intercorrelations for emotional experiences, predictions, bias, simple differences and correspondence as profile correlations. The target of prediction (self vs other) and the point of time (T1/T2) are differentiated. To avoid overestimation of stability, retest reliabilities were not calculated on a group level. Instead, 4 random subsamples were drawn and the range of the resulting

Table 4.1: Intercorrelations and Descriptive Statistics (Group Level, N=43)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. Age																				
2. Gender ^a	-.08																			
3. Intelligence	-.06	-.06																		
4. Knowledge	-.24	-.31*	.36*																	
5. EI Overall	.01	.13	.29 [†]	.45**																
6. EI Perception	.05	-.07	-.15	.05	.63**															
7. EI Facilitation	.04	.03	.15	.10	.73**	.43**														
8. EI Understanding	.02	.16*	.43**	.47**	.70**	.18	.34*													
9. EI Management	-.06	.12	.33 [†]	.57**	.65**	.08	.39 [†]	.29 [†]												
10. Neuroticism	-.04	-.13	.11	.15	.35*	.32*	.31*	.17	.21											
11. Extraversion	.23	.11	-.18	-.16	-.07	-.07	.01	-.17	.04	-.31*										
12. Conscientiousness	-.18	-.09	-.16	.07	.22	.08	.27	-.01	.31*	-.05	.17									
13. Agreeableness	.09	-.21	-.01	-.25	-.02	.00	-.06	.04	-.03	-.19	-.08	.03								
14. Openness	.14	.03	.02	-.12	-.04	-.12	.13	-.05	-.02	.39**	.21	-.08								
15. Task conflict	.16	-.09	.11	.19	.04	-.10	-.16	.14	.19	.13	.19	.00	-.16	.15						
16. Relationship conflict	.12	-.08	-.04	-.16	-.24	-.10	-.23	-.09	-.24	.13	.23	-.17	.01	.10	.64**					
17. Friendship	-.11	-.07	.08	.10	.18	.06	.25	-.04	.29 [†]	.15	.15	.19	.01	.08	.00	-.33*				
18. Competence	-.21	.07	-.05	.07	.25	.05	.41**	.06	.23	-.03	-.08	.32*	-.12	.04	-.44**	-.77**	.37*			
19. Term paper ^b	-.20	.16	.34 [†]	.21	.22	-.09	.15	.30*	.23	.08	-.03	-.01	.01	-.04	-.18	-.16	-.14	.21		
Mean	22.90	.78	99.92	13.58	.44	.43	.41	.57	.37	2.89	3.64	3.58	3.61	3.94	1.93	1.58	2.01	3.72	3.97	
SD	5.55	.41	5.60	2.62	.04	.06	.05	.07	.06	.69	.63	.66	.53	.57	.61	.58	.63	.87	1.08	
N	170	170	180	158	174	174	174	174	174	175	175	175	175	175	195	195	195	195	195	195

^a Positive intercorrelations indicate higher values for males.

^b Higher term paper ratings indicate better performance on a scale from 1-6.

+ :p < .10

* :p < .05

** :p < .01, two-tailed

reliabilities is reported. Bias and simple differences show good retest reliability. Correspondence on the other hand is fairly stable for self predictions and rather unstable for correspondence.

Intercorrelations are calculated on a group level which complicates the interpretation of the effects, especially for forecasts, experiences, and simple differences. The main focus is on the last four columns, namely accuracy as bias and correspondence. Generally, accuracy of self-predictions is higher in terms of both a low bias (2.13/1.51 as opposed to 2.73/2.68) and high correspondence (.24/.27 as opposed to .44/.68). Also, only self-predictions improve over time. In groups with high knowledge, *later* (not earlier) predictions for the self are more accurate. At least partly, the higher later accuracy for the self should be due to the fact that feelings towards others stabilize in the early stages of the relationship. Hence, accuracy of self predictions is improved because actual feelings at the time of the prediction are closer to the target feelings.

EI scales, especially management and facilitation are mainly related to the accuracy for the self which points to the need to further study EI in an interpersonal context. The MSCEIT originally aims to capture an ability in regard to the self and others. Extraversion goes along with less accuracy for others, i.e. a higher bias and lower correspondence. Mean openness in a group leads to a higher bias, and mean conscientiousness leads to more accurate predictions for the self.

On a group level, low bias is related to positive peer ratings, with the exception of friendship and task conflict. This is not limited to self predictions. Low bias and higher correspondence for others is also adaptive (e.g. to know who will like whom or who will be angry with whom). This is remarkable given that correspondence for others can be unreliable. Accuracy for the self is also related to better term paper ratings, accuracy for others is not.

Table 4.2: Retest Reliability and Validity of Accuracy Measures (Group Level, N=43)

	Experience	Prediction T1/T2		Simple difference T1/T2		Bias T1/T2		Correspondence T1/T2	
		Self	Other	Self	Other	Self	Other	Self	Other
Retest reliability	-	.24 - .74	.26 - .72	.37 - .55	.53 - .76	.38 - .51	.60-.74	.21 - .49	-.16 - .58
Age	-.03	-.07/-12	-.15/-08	-.07/-17	-.04/-00	.03/10	.07/12	.10/.04	-.04/.07
Gender ^a	-.13	-.19/.00	-.02/-13	-.10/.22	.13/.02	.24/.24	.18/-05	-.01/-15	-.37*/.31*
Intelligence	-.17	-.16/-14	-.20/-05	.12/.12	.10/.15	.01/-17	.18/.08	-.18/.08	-.12/-13
Knowledge	-.12	-.12/-01	-.12/-15	.09/.20	.13/.09	-.24/-42**	-.03/-09	.02/.29+	.03/.05
EI Overall	.23	.16/.16	.17/.03	-.09/-18	-.13/.31+	-.24/-25	-.02/.03	.31*/40**	.16/.20
EI Perception	.18	-.17/.09	.23/-01	-.08/-22	-.10/-25	-.09/.09	-.15/-01	.17/.19	.04/.27+
EI Facilitation	.42**	.43**/.28+	.41**/.30+	-.10/-33**	-.18/-33*	-.17/-07	.13/.04	.39**/.29+	.13/.22
EI Understanding	-.06	-.08/-05	-.08/-10	-.01/-02	-.02/-10	-.01/-17	-.05/-01	.02/.11	.28+/-07
EI Management	.17	.04/.18	.02/.00	-.07/.01	-.11/-26	-.39*/-50**	.09/.11	.34*/.51**	-.05/.21
Neuroticism	.07	-.06/.03	.02/-09	-.17/.02	-.07/-23	-.08/-02	-.25/.09	.05/.18	-.00/-25
Extraversion	-.10	-.13/-04	-.10/-09	.04/-08	.06/.05	.08/.16	.43**/.23	.26+/-07	-.34*/.16
Conscientiousness	.27+	.18/.27+	.25/.18	-.17/-11	-.09/-09	-.23/-17	.13/-05	.33*/.19	-.10/.33*
Agreeableness	-.03	.13/.06	.14/.18	.19/.10	.15/.27+	.04/.05	.13/.07	-.10/-09	-.13/-07
Openness	-.07	.19/-07	.23/.14	.26+/-08	.21/.19	.29+/.11	.43**/.30+	-.02/-05	-.11/.06
Task conflict	-.48**	-.37*/-52**	-.38*/-56**	.24/-14	.33*/.07	.16/-10	.21/.49**	-.01/.30*	-.12/-43**
Relationship conflict	-.77**	-.45**/-68**	-.27+/-60**	.44**/.00	.64**/.27+	.51**/.42**	.45**/.68**	-.12/-12	-.32*/-38**
Friendship	.54**	.31*/.53**	.11/.42**	-.17/.01	-.50**/-22	-.23/-27	-.01/-10	.07/.32*	-.15/-15
Competence	.84**	.64**/.65**	.53**/.64**	-.31**/.19	-.54**/-38*	-.46**/-32*	-.31**/-57**	.22/.06	.31**/.28+
Term paper ^b	.06	-.18/.02	-.09/-06	-.26+/.04	-.13/-22	-.36*/-27+	-.26/-26	.34*/-03	.20/.18
Mean	4.49	4.48/4.53	4.43/4.55	.01/-01	-.05/.03	2.13/1.51	2.73/2.68	.44/.61	.24/.27
SD	.69	.69/.83	.49/.59	.70/.59	.64/.63	1.42/1.27	1.16/1.31	.32/.28	.23/.25

Note. High experiences and predictions indicate more affection, more satisfaction with the collaboration, less anger and more fun. High simple differences indicate overestimation, i.e. predictions are higher than experiences. Predictions T1 were conducted with one month of acquaintance, T2 with three months.

^a Positive intercorrelations indicate higher values for males.

^b Higher term paper ratings indicate better performance.

+ :p<.10

* :p<.05

** :p<.01, two-tailed

The first three columns, experiences and predictions of group-related affect are fairly unrelated to intelligence, knowledge and personality. There is little association to conscientiousness, meaning that groups higher in conscientiousness have better average experiences and predictions. Since conscientiousness was useful for goal attainment in the work groups, this association seems plausible. Also, groups with higher facilitation, i.e. individuals who can use or feel emotions, have higher predicted and experienced feelings.

The relationships of experiences and predictions to the criteria are rather trivial. Groups with higher mean experience (in terms of affection, satisfaction with the collaboration, fun, and absence of anger) report less relationship conflict and perceive each other as friends and as competent. Generally, experiences were positive with a mean score of 4.49 on a scale from 1 to 5.

4.2.1.3 Accuracy Decomposition

Profile correlations in a round robin design can be divided into four components that are presented in Table 4.3. Results for elevation (i.e. bias) can be found in Table 4.2. Stereotype accuracy means that the predictions reflect how affective dimensions are generally rated. Differential elevation indicates that the predictions accurately reflect differences between targets, f.e. that a certain member in the group is accurately expected to cause worse emotions than others. Differential accuracy is the accuracy adjusted for elevation, stereotype accuracy and differential elevation. It reflects the accuracy of predicting that certain people feel in a certain way, even if others would not (Kenny & Winqvist, 2001). A distinction as to the self and others (as in Table 4.2 is not possible since different targets are needed for the decomposition. Relationships to tests and criteria, as well as retest reliabilities are reported. Again, correlations are calculated on a group level, and reliability is reported for subsamples.

Reliabilities vary considerably, more so than for bias and correspondence.

Table 4.3: Validity of Accuracy Components (Group Level, N=43)

	Stereotype accuracy	Differential elevation	Differential accuracy
Retest reliability	.11 - .57	.05 - .45	-.17 - .53
Age	.11/-.16	-.05/-.12	.04/-.07
Gender	-.11/.12	-.14/.05	-.38*/.11
Intelligence	-.16/.02	-.24/-.14	.15/.15
Knowledge	.02/.23	.14/.26	-.20/.04
EI Overall	.33*/.30 ⁺	.26/.29	.03/-.15
EI Perception	.43**/.27 ⁺	.24/.17	-.16/-.06
EI Facilitation	.14/.14	.13/.21	.20/-.14
EI Understanding	.24/.14	.21/.06	.10/-.09
EI Management	.05/.25	.10/.35*	.00/-.15
Neuroticism	.12/-.08	.06/-.14	.02/-.15
Extraversion	-.35*/-.11	-.31*/-.05	.17/.04
Conscientiousness	-.15/.20	.01/.37*	.06/-.08
Agreeableness	-.09/-.10	-.21/-.09	.22/.11
Openness	-.20/-.25	-.22/-.10	.30 ⁺ /.30 ⁺
Task conflict	-.31*/-.32*	-.20/-.14	.18/.06
Relationship conflict	-.44**/-.41**	-.47**/-.45**	.31*/.22
Friendship	-.09/-.33*	-.06/-.04	-.08/-.29 ⁺
Competence	.30*/.18	.44**/.38*	-.23/-.28 ⁺
Term paper	.13/.26	.15/.09	.15/.24
Mean	.32/.32	.27/.33	.12/.16
SD	.18/.18	.14/.15	.10/.12

^a Positive intercorrelations indicate higher values for males.

^b Higher term paper ratings indicate better performance.

+ :p < .10

* :p < .05

** :p < .01, two-tailed

Especially the stability of differential accuracy is alarming, and it does not seem suitable as a measure of individual differences. Mean differential accuracy and variance is considerably lower than for the other components. Stereotype accuracy and differential elevation show better stability and are positively related to conflict measures and peer ratings. It can be assumed that a rough prediction, i.e. matching the mean profiles, is adaptive. Also, stereotype accuracy and differential elevation are related to emotional intelligence, i.e. emotional perception, especially for earlier ratings.

4.2.2 The Person as Unit of Analysis

4.2.2.1 Perceiver and Target Effects

Following the social relation model, a round robin design allows to differentiate actor and partner effects in groups (Kenny, 1994). Typically, an actor effect of e.g. affection indicates that certain participants like others to a higher degree. Partner effects indicate that some are better liked than others. In context of interpersonal perception, another dimension is added and perceiver and target effects are distinguished. When people are asked how much affection they will feel for each other, perceiver effects indicate that some people are good in predicting affection. Target effects indicate that affection can be better predicted as to certain targets. In other words, perceiver effects capture that some participants are better judges, target effects indicate that some participants are easier to read.

In the current study, another dimension is added. Participants not only predict their feelings for others, but they are asked to predict the others will feel towards the others, e.g. how group member A will feel toward member B. This adds a partner effect of the second degree, specifically variance due to member B. Data analysis followed the recommendations of Kenny and West (2006) and was conducted with SPSS 17. The recommended mixed model approach was slightly extended to add target effects of the second

degree. Table 4.4 shows the results for predictions and bias. Correspondence can not be analyzed since profile correlations already combine two sets of measurements.

Both for emotional predictions alone and for bias, a lot of variance is unaccounted for. Perceiver variance is the largest component in emotional predictions. This means that participants vary in their tendency to predict higher or lower affect. With more acquaintance (T2) the target of the second degree explains more variance. This suggests that predictions vary in terms of how much anger or fun someone elicits, rather than how much fun is someone likely to feel. Bias on the other hand has a smaller perceiver effect. This is bad news for the search of a good judge of emotions. Accuracy scores in terms of bias mainly vary with the target of the second degree. In other words, the feelings toward some participants are easier to predict, presumably because of special behavior like free riding or diligence. This demonstrates the context-dependency of this measure, and it does not support accuracy as a valid measure of interindividual differences regarding an ability. Perhaps people do focus on certain individuals and disregard others. Which could be adaptive in terms of emotional intelligence, especially when a rough scan indicates the absence of a threat. Still, an independent measure of emotional knowledge (that does not vary regarding different targets) is necessary (see section 5.3).

4.2.2.2 Bias and Correspondence

First, the relationship between accuracy and intelligence, emotional intelligence, and personality is examined on an individual level³. Dependency in

³The analyses are conducted with accuracy for T1. Since a causal effect of accuracy on adaptive criteria is theoretically assumed, the temporal precedence of the cause is necessary (Cook, Campbell, & Peracchio, 1990). This way, the latent skill underlying early predictions can be influential in the interaction with fellow group members. Even if construal level theory suggest that later predictions should be more elaborate because of less temporal distance (Liberman et al., 2002), later predictions are also potentially too close to the actual experiences for the self, or emotional perceptions of others.

Table 4.4: Effects for Perceivers, Targets, and Targets of the Second Degree

	Variance component	SD	Variance explanation
	T1/T2	T1/T2	T1/T2
Emotional predictions			
Perceiver	.15 / .21	.02 / .03	11 / 12%
Target	.03 / .04	.01 / .01	2 / 3%
Target 2	.04 / .12	.01 / .02	3 / 7%
Bias			
Perceiver	.31 / .48	.07 / .10	2 / 3%
Target	.17 / .25	.15 / .14	1 / 2%
Target 2	.99 / .89	.26 / .22	6 / 6%

Note. When person A predicts the anger that person B will feel toward person C, A is the perceiver, B is the target, and C is the target 2.

the data will be controlled with multilevel modeling in a random intercept model (Raudenbush & Bryk, 2002).

In contrast to the anxiety study, individuals are modeled on level one, and groups are modeled on level 2. Abilities, traits, and criteria will be predicted with the accuracy components following Cronbach (1955). For the example of intelligence (I) as the independent variable and elevation (bias), stereotype accuracy (SA), differential elevation (DE), and differential accuracy (DA) as predictors, the model is as follows:

$$I_{ij} = \beta_{0j} + \beta_{1j}Bias_{ij} + \beta_{2j}SA_{ij} + \beta_{3j}DE_{ij} + \beta_{4j}DA_{ij} + r_{ij} \quad (4.1a)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (4.1b)$$

$$\beta_{1j} = \gamma_{10} \quad (4.1c)$$

$$\beta_{2j} = \gamma_{20} \quad (4.1d)$$

$$\beta_{3j} = \gamma_{30} \quad (4.1e)$$

$$\beta_{4j} = \gamma_{40} \quad (4.1f)$$

All accuracy scores are unrelated to intelligence ($p > .38$), gender ($p > .23$), age ($p > .19$), agreeableness ($p > .48$), conscientiousness ($p < .17$), emotional facilitation ($p < .27$), emotional understanding ($p < .44$), and emotional management ($p < .68$). There is an association to knowledge though. Knowledge is related to lower bias (especially for others, $\beta = -.34$, $t = -1.728$, $p = .086$) and lower differential accuracy ($\beta = -3.04$, $t = -2.699$, $p = .008$). Extraverts tend to have a higher bias (both for self and others, $\beta = .16$, $t = 3.382$, $p = .001$) and lower stereotype accuracy ($\beta = -.40$, $t = 1.895$, $p = .059$). There is a trend for neuroticism to be associated to a lower bias (especially for others, $\beta = -.82$, $t = -1.884$, $P = .061$). Openness relates to higher bias (both for self and others, $\beta = .16$, $t = 3.657$, $p = .001$), lower stereotype accuracy ($\beta = -.36$, $t = -2.24$, $p = .027$) and higher differential accuracy ($\beta = .40$, $t = 2.011$, $p = .046$). Differential Elevation is related to emotional perception ($\beta = .08$, $t = 2.59$, $p = .011$).

Next, adaptability of accuracy is analyzed with the individual as the unit of analysis. First term paper ratings are explored. It is not possible to model term paper ratings as a dependent variable in multilevel modeling since it only varies on the group level. Instead, accuracy can be modeled as the outcome variable, and the strength of association can be estimated when the term paper rating (TPR) serves as a group level predictor. For the example of bias, the following model is estimated:

$$Bias_{ij} = \beta_{0j} + r_{ij} \quad (4.2a)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}TPR + u_{0j} \quad (4.2b)$$

Sequential analyses are conducted for all accuracy measures. Better term paper ratings are related to lower bias for the self ($\beta = -.30$, $t = -3.083$, $p = .004$) and unrelated to other accuracy measures. This indicates adaptability of a low bias. Whether this relationship is incremental to knowledge

can not be decided with multilevel modeling. Instead, incremental validity is tested on a group level using hierarchical linear regression. When elevation is added to a model with age, gender, intelligence and the big 5, bias leads to a substantial model improvement, i.e. it shows incremental validity ($\beta = -.59, \Delta R^2 = .27, p < .001$). This model includes 9 predictors which is a lot given that there are 43 groups. Results are replicated for a more parsimonious model, and bias also adds to age, gender, and intelligence ($\beta = -.35, \Delta R^2 = .17, p = .04$). When elevation for self and others are differentiated, elevation for the self leads to the model improvement.

For all other criteria, multilevel models can be used. Table 4.5 shows the results for peer ratings of task and relationship conflict, Table 4.6 those for competence and friendship. In the Tables, the first step uses accuracy to predict the criteria. Then, intelligence, knowledge, personality, and demographic data is included. Typically, these steps are reversed to assess incremental validity, but the current sequence allows to examine the relationship of accuracy to criteria alone. The PRE, as a measure comparable to ΔR^2 is reported for the typical sequence, i.e. it captures the incremental validity of accuracy over intelligence and personality. The MSCEIT will not be entered in the model since the validity of accuracy is not required to be incremental to emotional intelligence.

Results show that none of the accuracy measures relate to task conflict. Participants that are female, intelligent, open, and conscientious, though, are more likely to evoke peer ratings of task conflict. Intelligence, knowledge, personality, and demographic data accounts for 11% of the variance compared to the null model. Accuracy adds less than .1%. For relationship conflict, bias shows incremental validity over personality and intelligence. Other abilities and traits lead to a PRE of 21%, and bias adds an additional 6%. When biases for self vs. others are distinguished, both are significant ($p < .05$).

Table 4.5: Hierarchical Random Intercept Model for Task and Relationship Conflict

	Task conflict	Relationship conflict
	Coefficient (t-ratio)	Coefficient (t-ratio)
Step one		
Intercept	1.83*** (8.06)	1.21*** (6.70)
Elevation	.09 (1.40)	.16** (3.28)
Stereotype Accuracy	-.28 (-1.45)	.05 (-.24)
Differential Elevation	.02 (.07)	-.38 (.25)
Differential Accuracy	-.11 (-.53)	.23 (1.04)
Step two		
Intercept	-1.43 (-1.59)	-.29 (-.44)
Age	.02 ⁺ (1.67)	.01 (1.50)
Gender	-.24* (-2.29)	-.07 (-.65)
Intelligence	.02** (3.09)	.02** (2.99)
Knowledge	-.04 ⁺ (-1.73)	-.06* (2.00)
Neuroticism	.09 (1.04)	-.01 (-.06)
Extraversion	.08 (1.03)	.03 (.52)
Conscientiousness	.13** (2.87)	-.01 (-.25)
Agreeableness	-.08 (-1.09)	-.07 (-1.80)
Openness	.22** (3.65)	-.16** (3.30)
Elevation	.06 (.75)	.13* (2.54)
Stereotype Accuracy	-.32 (-1.37)	-.09 (-.40)
Differential Elevation	.13 (.34)	-.37 (-1.18)
Differential Accuracy	-.33 (-1.27)	.12 (.64)

Note. Gender: positive values indicate that male persons show higher values.

Table 4.6: Hierarchical Random Intercept Model for Competence and Friendship

Fixed effects	Competence	Friendship
	Coefficient(t-ratio)	Coefficient (t-ratio)
Step one		
Intercept	4.47*** (15.85)	2.20*** (10.53)
Elevation	-.19** (-3.05)	.06 (-.95)
Stereotype Accuracy	-.85** (-3.03)	-.20 (-1.04)
Differential Elevation	.41 (.94)	.03 (.25)
Differential Accuracy	-.44 (-1.31)	-.02 (-.07)
Step two		
Intercept	2.49 (1.36)	.09 (.07)
Age	-.02 (-1.02)	-.01 (-.22)
Gender	-.03 (-.16)	-.02 (-.12)
Intelligence	.01 (.59)	<.01 (.71)
Knowledge	.08* (2.43)	.02 (.53)
Neuroticism	.06 (.51)	.11 (1.25)
Extraversion	.10 (.84)	.10 (1.17)
Conscientiousness	<.01 (<.01)	-.04 (-.38)
Agreeableness	-.11 (-.84)	.17 (1.79 ⁺)
Openness	.04 (.30)	-.05 (-.58)
Elevation	-.19* (-2.42)	-.04 (-.60)
Stereotype Accuracy	-.71* (-2.19)	-.24 (-1.07)
Differential Elevation	.45 (.92)	.30 (.90)
Differential Accuracy	-.27 (-.68)	-.09 (-.24)

Note. Gender: positive values indicate that male persons show higher values.

Knowledge is associated with higher peer ratings of competence and, again, lower bias improves this model. Surprisingly, higher stereotype accuracy is related to lower ratings of competence. This accuracy does not seem to be adaptive. On a group level, though, higher stereotype accuracy was associated with higher average competence ratings (see Table 4.3). In sum, the measurement quality of the decomposed accuracy scores has to be questioned. Intelligence, personality, and demographics show a PRE of 25%. Accuracy scores add another 11%.

Agreeable participants are slightly more likely to be perceived as a friend. Apart from that, intelligence, personality, and demographic data do not explain friendship ratings. The PRE is 7 %. Accuracy measures do not add to the model.

4.3 Discussion

Again, results will be discussed with regard to the research questions. First, results on the stability of accuracy are summarized (4.3.1), then the relationship of accuracy to other abilities and traits is revised (4.3.2), and finally the adaptability of accuracy is addressed (4.3.3).

4.3.1 Stability of Bias and Correspondence

Retest reliabilities were good for simple difference scores and bias, but problematic for correspondence. Especially correspondence for others and correspondence after variance decomposition show reliabilities that do not meet common standards. At least results for differential accuracy should be treated with caution. This is the most interesting accuracy type as to Cronbach (1955), but Kenny and Albright (1987) already mentioned that the decomposition can lead to a lack of stability for this component. Hence, they recommended group level analysis. Indeed, on a group level, plausi-

ble associations to stereotype accuracy and differential elevation are found. With the individual as the unit of analysis, correspondence is less successful in predicting the criteria than bias. Future research should try to reduce unsystematic error variance in order to further improve the measurement quality, especially of correspondence (see section 5.3).

Only for self-predictions, bias and correspondence improve over time. This may be due to the fact that the experienced emotions during late predictions were closer to the target emotions at the end of the term. For others, there seems to be a lack of accurate feedback as to their feelings. Generally, predicting the feelings of others is a more difficult task.

In addition, social relations analysis shows that a substantial proportion of variance is due to targets, so that accuracy of predictions is not stable over different targets. Again, context has to be taken into account when examining emotional perception accuracy.

4.3.2 Bias, Correspondence, and other Abilities and Traits

On a group level, accuracy for later self predictions is associated with knowledge. On an individual level, this relationship for correspondence can not be found, but lower bias was again associated with knowledge, especially for others. This is plausible since important errors of emotional predictions, motivated distortions (Wilson & Gilbert, 2003), can not occur when predictions are made for others. This supports the assumption that accuracy captures an aspect of knowledge. There are no relations to intelligence.

On a group level, correspondence for others is associated with emotional understanding which provides evidence that something similar is measured here. Again, relations to correspondence can only be found on a group level. Given the conceptual closeness of the emotional prediction accuracy to emotional understanding, few relations between these concepts are unexpected. It would be cause for more concern, though, if there was a 'gold standard'

for the measurement of EI which currently, there is not (Matthews et al., 2012, p. 120).

Regarding other branches of EI, accuracy for self predictions is associated with emotional management on a group level, and there are some relations to emotional facilitation. Also, emotional perception is related to higher stereotype accuracy on a group level and higher differential elevation on an individual level. Accordingly, some closeness of the accuracy measures to the MSCEIT can be demonstrated.

There are substantial relationships to personality. On a group level, openness is related to a higher bias and extraversion is associated to lower accuracy for others. On an individual level, extraversion and openness again relate to a higher bias with openness also being related to lower stereotype and higher differential accuracy. Neuroticism, on the other hand, goes along with less bias. These relationships seem to reflect personal styles and motivations captured with personality measures.

4.3.3 The Adaptability of Bias and Correspondence

The results regarding adaptability are encouraging. Lower bias is substantially associated with better term paper ratings and better peer ratings, even incrementally to intelligence and personality. Specifically, participants with lower bias were rated as more competent by their peers and there was less relationship conflict reported with them. This supports the assumption that accurate knowledge can provide control over a situation and help to achieve personal and interactional goals. Furthermore, accuracy relates to strong criteria that are independent from self-report. There are no relations to less task conflict, but indeed task conflict can be beneficial and promote group success (De Dreu & Weingart, 2003). Also, no relations to peer ratings of friendship are found. The intention to become friends with specific others should be a moderating influence here.

Chapter 5

General Discussion

5.1 Theoretical Considerations

The focus of this thesis was to evaluate the validity and adaptability of accuracy of emotional predictions. In the anxiety study, this was done in a setting similar to affective forecasting research, and with a clear focus on test anxiety in a specific situation. The work-group study broadened the scope to a complex social context. It focused on the prediction of interpersonal feelings in newly emerging work-groups with a common goal.

It is important to note that EI, as measured in the branches of the MSCEIT, and accuracy, as measured here, are functions not processes (cf. Matthews, Zeidner, & Roberts, 2007). Inferences about the feelings of others require the use of contextual information and a combination of information derived from situation- and person-specific cues (cf. Gnepp, 1989; Robinson & Clore, 2002a). Thus, the context has to be taken into account.

It can be questioned that emotional intelligence and emotional knowledge generalize over different domains (Matthews et al., 2007). Emotional knowledge should guide the selection of coping strategies, or more generally emotion management tactics (cf. Matthews & Zeidner, 2000), but this does not guarantee universally positive outcomes. The employed strategy

might only make a slight difference in certain situations but might still be advantageous to try. Strategies may only apply to certain situations or be maladaptive as to certain criteria (Matthews et al., 2012). Future research should aim to further clarify the nature of emotional knowledge and of the situational specifics that promote its adaptability.

Given the relationship to intelligence, emotional intelligence, and personality, it is obvious that accuracy of emotional predictions can not serve as a proxy measure of emotional understanding. Still, there are plausible relations to the MSCEIT and knowledge, so that there is hope for the concept. In any case, though, measurement quality has to be improved.

Results on the adaptability of bias are encouraging, results on correspondence are mixed. There is evidence that low bias leads to less distress, less negative coping, better term paper ratings, and better peer ratings. In addition, there is indication for incremental validity.

In the anxiety study, there was more control over the situational circumstances. Also, social cognition research provided a theoretical framework for specific aspects of emotional knowledge, i.e. knowledge about mood repair processes and the temporal progression of emotions (Gilbert et al., 1998). The duration bias was found on a group level, but an examination of individual differences suggests that a lower bias is preferable, especially when trait anxiety is low.

The environment in the work-group study was highly complex, and there was a lot of variance unaccounted for. Also, participants reported little relationship conflict. Nevertheless, bias was adaptive as to hard criteria. For correspondence, retest reliability is in part problematic, especially after variance decomposition and especially for differential accuracy. Correspondence is useful to predict criteria on a group level but the measurement quality needs to be improved for analyses on an individual level. However, future research has to explore whether there is use for differential accuracy of emo-

tional predictions, even if it was reliably measured. Being able to predict someone's emotions really well might not be adaptive. It can suffice to have a rough estimate in terms of stereotype accuracy and differential elevation (as general rules about traits and targets) and to be open to changes in the situations. The reasoning about predictions should probably not be overly extensive but flexible.

5.2 Strengths and Limitations

The studies in this thesis focus on emotionally relevant situations with high personal and practical importance. Strong and emotionally relevant criteria were used and incremental validity could be assessed since all participants conducted extensive diagnostic batteries. The work-group-study additionally focused on the social dimension of emotional knowledge and allowed for the distinction of several specific targets. Thus, it offers the opportunity to examine the prediction of emotions in relationships and to decompose accuracy as to certain targets and dimensions. The accuracy measures were more ambitious than in studies that focus on the interindividual differences in affective forecasting accuracy, and that rely on simple difference scores (Dunn et al., 2007; Hoerger, 2012; Hoerger et al., 2012).

The main shortcoming, though, is the lack of control, especially in the work-group study. Participants were to mentally simulate an upcoming event and apply general and target-specific rules to predict emotional reactions. Chances were high that the situation in question (the end of the semester) was misconstrued (cf. Wilson & Gilbert, 2003) since participants had to deal with a lot of uncertainty. Inaccuracy is not only due to a lack of emotional knowledge but presumably due to changes in the situation that could not have been anticipated. For example, someone could have suffered a personal loss and missed out on a lot of work, or someone is won over for the empirical method and engages more than expected. Nonetheless,

it could be argued that the current scenario of uncertainty is realistic and of great practical relevance. Still, given the unsystematic error variance, knowledge should be measured more directly as proposed in section 5.3.

Given that the emotional predictions are designed to measure ability EI, correlations to EI and intelligence are in part disappointing. One can argue that low correlations are also due to the fact that the accuracy of emotional predictions is basically measured with one item, i.e. one emotionally challenging situation.

Differentiated knowledge about the development of emotions over time, and about the way that certain emotions are rank ordered is better reflected in correspondence measures. These measures are more sensitive to situational and personal details but they are also far less robust than bias, as an aggregated composite measure. This is reflected in low retest reliability or at least a broad range of reliabilities. In spite of low measurement quality, some relationships are found, and there is evidence for the theoretical value of correspondence scores. The accuracy decomposition and the focus on various targets are powerful tools to further explore emotional knowledge in a social context.

5.3 Implications for Future Research

Current literature on EI emphasizes the need to focus on specific abilities and to understand these abilities in specific contexts with a broad theoretical basis (Matthews et al., 2012; Mayer et al., 2008). The current findings support this notion. The relationship between emotional knowledge and criteria is complex and moderating influences need to be examined.

It is recommendable to continue researching emotional knowledge in regard to social relations. Motivated distortions or useful biases that might complicate self-predictions (e.g. Wilson & Gilbert, 2003) do not apply to the predictions for others. This warrants the general assumption that ac-

curate knowledge about others is adaptive. The results of the work-group study subtly hint to this possibility. On a group level, emotional understanding is mildly associated with correspondence for others, and it predicts improved peer ratings. Here, accuracy decomposition seems useful but the measurement quality has to be improved.

This can be done in two different ways. The first approach operates within the current framework of predicted and experienced emotional reactions. The goal is to gain control over the situation and to reduce unsystematic variance. To avoid misconstrual and to better capture the affective theories (cf. Wilson & Gilbert, 2003), confederates can help to set a specific context for emotional predictions. This context can be highly comparable among participants, and instructions could additionally set comparable interactional goals. For example, participants get to know the confederate and are instructed to handle their complaints. Predicted associated emotions can then be related to experienced ones. This would also reduce target variance, i.e. that some people are easier to read than others. Target variance was substantial in the current study.

The second approach allows to continue to examine emotional knowledge in a complex setting but it involves taking a different measurement approach to emotional knowledge. People have to deal with complexity and uncertainty in real life and apply situation- and person-specific beliefs as they see fit (Robinson & Clore, 2002a). To better understand the reasoning involved and how individuals anticipate different outcomes, methods of the Berlin wisdom paradigm can be applied (Baltes & Smith, 2008). Think-aloud protocols can be used to capture 'rich factual and procedural knowledge' (p.58), knowledge about certain *contexts*, and knowledge about *uncertainties*. This can be useful to further examine the basis of predictions. Applied rules can be differentiated and rated as to quantity and quality. Also, personal motivations and goals should be elaborated on.

The empathic accuracy paradigm also relies on open answers and adds a qualitative dimension to the measurement of an emotional skill (Ickes, 1993, 2003). In this paradigm, participants watch videos of different scenarios and infer thoughts and feelings of others. In its current form, the paradigm measures emotional perception but it could be incorporated for emotional knowledge, too. An emotion could be elicited in a lab, f.e. anxiety in context of a public speaking task. Some participants could observe others and elaborate on their emotional predictions for them. Other participants partake in the public speaking task and elaborate on their own feelings.

A partly qualitative approach to the measurement of emotional knowledge can add theoretical clarity, help to understand emotional expertise, and guide the construction of future measurement tools. One can identify how elaborate emotional rules are and how flexibly they are applied.

Another open question is the causal relationship between accuracy and criteria. Manipulating accuracy in an experimental design is hardly possible. One could provide accurate and inaccurate emotional expectations in a within-subjects design, but the conceptual closeness to actual knowledge is unclear. Perhaps, after further clarifying the concept of emotional knowledge, studies can compare trained and untrained individuals as to goal attainment in emotionally challenging situations.

Joseph and Newman (2010) suggested emotional labor as a central moderator of the EI-performance relationship. Emotional labor refers to the extent that the regulation and display of emotions is part of the job. This is believed to be particularly stressful when it involves surface acting, i.e. the display of emotions without actually feeling them (Hochschild, 1983; Grandey, 2000). Emotional labor should not have been a general problem in the student work groups, but one could have measured the perceived need to suppress feelings, f.e. with the Emotion Regulation Questionnaire by Gross and John (2003). Following emotion regulation theories, antecedent focused

emotion regulation can be distinguished from response-focused techniques. The basic assumption is that regulating the true feeling is less exhausting than faking a particular feeling (Gross & John, 2003).

Finally, romantic or family relationships are interesting to investigate. As important relationships, they should be associated with rich emotional knowledge: knowing how your partner handles performance situations, how they react to criticism, or what promotes their relaxation after a day of work.

5.4 Conclusion

There was strong support for the assumption that accuracy of emotional predictions is adaptive. Especially a low bias shows substantial correlations to strong criteria and proves to be incrementally valid. It is not about a generally rosy view (i.e. simple difference scores) but about accuracy. However, under certain conditions (f.e. high trait anxiety), low accuracy seems to be advantageous. This is evidence that emotional intelligence studies can profit from studies in other emotion-related areas, f.e. social cognition or coping research.

The support for accuracy as a cognitive ability close to EI is moderate or low. It is argued that emotional knowledge should also be measured independently from actual experiences. With open-ended questions and partly qualitative measures, knowledge underlying emotional predictions can be better captured. The scoring of right and wrong can be more flexible than it was in the current studies, and it can take situational changes into account. Future research should continue to explore the social dimension of emotional knowledge and accuracy decomposition can be used to distinguish emotion- and person-specific beliefs.

References

- Ackerman, P. L. (1994). Intelligence, attention, and learning: Maximal and typical performance. In D. K. Detterman (Ed.), *Current topics in human intelligence: Theories of intelligence* (Vol. 4, p. 1-27). Norwood NJ: Ablex.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology*, *108*, 441-485.
- Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *Intelligenz-Struktur-Test 2000 R [Intelligence-Structure-Test]*. Göttingen: Hogrefe.
- Asendorpf, J. (2002). Emotionale Intelligenz nein, emotionale Kompetenzen ja [Emotional intelligence - no, emotional skills - yes]. *Zeitschrift für Personalpsychologie*(1), 180-181.
- Baltes, P., & Smith, J. (2008). The fascination of wisdom. Its nature, ontogeny, and function. *Perspectives on Psychological Science*, *3*(1), 56-64.
- Bar-On, R. (1997). *The Emotional Quotient Inventory (EQ-I): A test of emotional intelligence*. Toronto Canada: Multi-Health Systems.
- Bernieri, F. J. (2001). Toward a taxonomy of interpersonal sensitivity. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (p. 3-20). Mahwah: Lawrence Erlbaum Associates.
- Borman, W. C. (1977). Consistency of rating accuracy and rating errors

- in the judgment of human performance. *Organizational Behavior and Human Performance*, 20, 238-252.
- Brody, N. (2004). What cognitive intelligence is and what emotional intelligence is not. *Psychological Inquiry*, 15(3), 234-238.
- Cattell, R. B., & Weiß, R. H. (1971). *Grundintelligenztest Skala 3 - CFT 3 [Basic Intelligence Scale 3 (CFT)]*. Braunschweig: Georg Westermann Verlag.
- Christiansen, N. D., Wolcott-Burnam, S., Janovics, J. E., Burns, G. N., & Quirk, S. W. (2005). The good judge revisited: Individual differences in the accuracy of personality judgments. *Human Performance*, 18(2), 123-149.
- Ciarrochi, J., Dean, F. P., & Anderson, S. (2002). Emotional intelligence moderates the relationship between stress and mental health. *Personality and Individual Differences*, 32(2), 197-209.
- Colvin, C. R., & Bundick, M. J. (2001). In search of the good judge of personality: Some methodological and theoretical concerns. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (p. 47-65). Mahwah: Lawrence Erlbaum Associates.
- Cook, T. D., Campbell, D. T., & Peracchio, L. (1990). Quasi experimentation. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology, Vol. 1* (p. 491-576). Palo Alto: Consulting Psychologists Press.
- Cronbach, L. J. (1955). Processes affecting scores on 'understanding of others' and 'assumed similarity'. *Psychological Bulletin*, 52, 177-193.
- Cronbach, L. J. (1984). *Essentials of psychological testing* (Vol. 4). New York: Harper and Row.
- Day, A. L., & Carroll, S. A. (2008). Faking emotional intelligence (EI): Comparing response distortion on ability and trait-based EI measures. *Journal of Organizational Behavior*(29), 761 -784.

-
- De Dreu, C. K. W., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology, 88*(4), 741-749.
- Denham, S. A., & Kochanoff, A. (2002). Why is she crying?: Children's understanding of emotion from preschool to preadolescence. In L. F. Barrett & P. Salovey (Eds.), *Wisdom in feeling: Psychological processes in emotional intelligence* (p. 239-270). New York: Guilford Press.
- Dunn, E. W., Brackett, M. A., Ashton-James, C. E., Schneiderman, E., & Salovey, P. (2007). On emotionally intelligent time travel: Individual differences in affective forecasting ability. *Personality and Social Psychology Bulletin, 33*, 85-93.
- Dunning, D., & Story, A. L. (1991). Depression, realism, and the overconfidence effect: Are the sadder wiser when predicting future actions and events? *Journal of Personality and Social Psychology, 61*(4), 521-532.
- Fiori, M., & Antonakis, J. (2011). The ability model of emotional intelligence: Searching for valid measures. *Personality and Individual Differences, 50*(3), 329-334.
- Gilbert, D. T. (2006). *Stumbling on happiness*. New York: Knopf.
- Gilbert, D. T., & Ebert, J. E. J. (2002). Decisions and revisions: The affective forecasting of changeable outcomes. *Journal of Personality and Social Psychology, 82*(4), 503-514.
- Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 75*(3), 617-638.
- Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (2002). Durability bias in affective forecasting. In T. Gilovich & D. Griffin (Eds.), *Heuristics and biases: The psychology of intuitive*

- judgment* (p. 292-312). Cambridge: Cambridge University Press.
- Gnepp, J. (1989). Children's use of personal information to understand other people's feelings. In C. Saarni & P. L. Harris (Eds.), *Children's understanding of emotion* (p. 151-177). Cambridge: University Press.
- Goleman, D. (1995). *Emotional intelligence. Why it can matter more than IQ*. New York: Bantam.
- Grandey, A. A. (2000). Emotional regulation in the workplace: A new way to conceptualize emotional labor. *Journal of Occupational Health Psychology, 5*(1), 95-110.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348-362.
- Hochschild, A. R. (1983). *The managed heart. Commercialization of human feeling*. London, Berkeley: University of California Press.
- Hoerger, M. (2012). Coping strategies and immune neglect in affective forecasting: Direct evidence and key moderators. *Judgment and Decision Making, 7*, 86-96.
- Hoerger, M., Chapman, B. P., Epstein, R. M., & Duberstein, P. R. (2012). Emotional intelligence: A theoretical framework for individual differences in affective forecasting. *Emotion, 12*, 716-725.
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality, 61*(4), 587-610.
- Ickes, W. (2003). *Everyday mind reading: Understanding what other people think and feel*. Amherst: Prometheus Books.
- Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science, 12*(1), 18-23.
- Janke, W., & Erdmann, G. (2002). *SVF 78. Eine Kurzform des Stressver-*

-
- arbeitsbogens SVF 120 [SVF 78. Short version of the questionnaire for coping with stress (SVF 120)].* Göttingen: Hogrefe.
- Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, *40*(2), 256-282.
- Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Test für das Berliner Intelligenz-Strukturmodell (BIS 4)*. Göttingen: Hogrefe.
- Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. *Journal of Applied Psychology*, *95*(1), 54-78.
- Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (p. 451-477). New York: Cambridge University Press.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: The Guilford Press.
- Kenny, D. A., & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology*, *80*(3), 439-448.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, *102*(3), 390-402.
- Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology*, *83*(1), 126-137.
- Kenny, D. A., & West, T. (2006). Estimation of the SRM not using specialized software. <http://davidakenny.net/doc/srmsoftware.doc>(last access: December, 2012).
- Kenny, D. A., & Winqvist, L. (2001). The measurement of interpersonal sensitivity: Consideration of design, components, and unit of analy-

- sis. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (p. 265-302). Mahwah: Lawrence Erlbaum Associates.
- Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Thousand Oaks: Sage.
- Kreft, I., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, *30*(1), 1-21.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121-1134.
- Lang, F. R., Luedtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen [Test characteristics and psychometric equivalence of a German version of the Big Five Inventory (BFI) for young, middleaged and old adolescents]. *Diagnostica*, *47*(3), 111-121.
- Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *Das State-Trait-Angstinventar (STAI) [The State-Trait-Anxiety-Inventory (STAI)]*. Weinheim: Beltz.
- Liberman, N., Sagristano, M., & Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology*, *38*, 523-534.
- Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (p. 619-642). New York: Oxford University Press.
- Matthews, G., Campbell, S. E., Falconer, S., Joyner, L., Huggins, J., Gilliland, K., et al. (2002). Fundamental dimensions of subjective

- state in performance settings: Task engagement, distress, and worry. *Emotion*, 2(4), 315-340.
- Matthews, G., & Zeidner, M. (2000). Emotional intelligence, adaptation to stressful encounters, and health outcomes. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (p. 459-489). San Francisco: Jossey-Bass.
- Matthews, G., Zeidner, M., & Roberts, R. (2007). Emotional intelligence: Consensus, controversies, and questions. In G. Matthews, M. Zeidner, & R. Roberts (Eds.), *The science of emotional intelligence. Knowns and unknowns* (p. 3-46). Oxford: University Press.
- Matthews, G., Zeidner, M., & Roberts, R. (2012). Emotional intelligence: A promise unfulfilled? *Japanese Psychological Research*, 54, 105-127.
- Matthews, G., Zeidner, M., & Roberts, R. D. (2003). *Emotional intelligence. Science and myth*. Cambridge, MA: MIT Press.
- Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, 59, 507-536.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence* (p. 3-31). New York: Basic Books.
- Mayer, J. D., Salovey, P., & Caruso, D. (2001). *The Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT)*. Toronto: Multi-Health Systems.
- Mayer, J. D., Salovey, P., Caruso, D., & Cherkasskiy, L. (2011). What is emotional intelligence and why does it matter? In R. J. Sternberg & J. Kaufman (Eds.), *The handbook of intelligence* (Vol. 3, p. 528-549). New York: Cambridge University Press.
- Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V 2.0. *Emotion*, 3(1), 97-105.

-
- Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997). Temporal adjustments in the evaluations of events. The rosy view. *Journal of Experimental Social Psychology, 33*, 421–448.
- Murphy, K. R., & Balzer, W. K. (1989). Rater errors and rating accuracy. *Journal of Applied Psychology, 74*(4), 619-624.
- Nezlek, J. B. (2001). Multilevel random coefficient analyses of event- and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin*(27), 771-785.
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*, 437-448.
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks: Sage.
- Raudenbush, S., Bryk, A., Cheong, Y., & Congdon, R. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Scientific Software International.
- Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion, 1*(3), 196-231.
- Robinson, M. D., & Clore, G. L. (2002a). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128*(6), 934-960.
- Robinson, M. D., & Clore, G. L. (2002b). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology, 83*(1), 198-215.
- Salovey, R., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality, 9*, 185-211.
- Scholl, W. (2007). Plädoyer für eine sozialere, interdisziplinäre und an-

- wendbarere Sozialpsychologie [Pleading for a more social, more interdisciplinary, and more applicable social psychology]. *Zeitschrift für Sozialpsychologie*, 38, 273-284.
- Schütz, A., Hertel, J., & Schröder, M. (2002). *Test zur emotionalen Intelligenz. Deutsche Version des MSCEIT von Mayer, Salovey und Caruso* [Test of emotional intelligence. German version of the MSCEIT by Mayer, Salovey, and Caruso]. Chemnitz University of Technology.
- Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)* [The Multidimensional State Questionnaire (MBDF)]. Göttingen: Hogrefe.
- Sulsky, L. M., & Balzer, W. K. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. *Journal of Applied Psychology*, 77, 497-506.
- Van Rooy, D. L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior*, 65(1), 71-95.
- Van Rooy, D. L., Viswesvaran, C., & Pluta, P. (2005). A meta-analytic evaluation of construct validity: What is this thing called emotional intelligence? *Human Performance*(18), 445-462.
- Wilson, T. D., & Gilbert, D. T. (2003). Affective forecasting. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (p. 345-411). New York: Elsevier.
- Wilson, T. D., Wheatley, T. P., Meyers, J. M., Gilbert, D. T., & Axson, D. (2000). Focalism: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, 78(5), 821-836.
- Zeidner, M., Shani-Zivotich, I., Matthews, G., & Roberts, R. (2005). Assessing emotional intelligence in gifted and non-gifted high school students: Outcomes depend on the measure. *Intelligence*(33), 369-391.