DIPLOMARBEIT

Titel der Diplomarbeit

Development of Stimuli to Examine Self-Other Distinction in the Context of Empathy

verfasst von

Angelika S. Längle

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2015

Studienkennzahl: A 298

Studienrichtung: Psychologie

Betreuer: Univ.-Prof. Mag. Dr. Claus Lamm

Abstract

The aim of this diploma thesis was to develop and validate stimuli in order to investigate one prerequisite in empathy research, specifically the self-other distinction. For this purpose videos were produced in which a target-person described an emotional life-event either how it actually happened (congruent condition), or with a commonly not expected emotional reaction (incongruent condition). It was hypothesized that the self-other distinction becomes visible with incongruent videos due to different emotions in the perceivers depending on whether they empathized with the target-person or remained within their own emotions. Emotions were operationalized by an affect rating dial. In the realization of the research, perceivers were asked to empathize with the target-person and to rate the valence of the target's emotions (perceivers' other-ratings) and then to put themselves in the same situation and to rate the valence of their own feelings (perceivers' self-ratings). These ratings constituted the basis for analysing self-other distinction (ANOVA). The results showed that the perceiver's self-other distinction was visible and significant at incongruent videos. Hence they appear to be useful stimuli in order to examine self-other distinction in the context of empathy.

Keywords

self-other distinction, empathy, incongruent video-stimuli, affect rating dial.

Abstract

Das Ziel dieser Diplomarbeit war die Entwicklung und Validierung von Stimuli um eine Voraussetzung der Empathieforschung zu untersuchen, und zwar die Self-Other Distinction. Hierfür wurden Videos entwickelt, in denen eine Target-Person emotionale Lebensereignisse beschreibt: entweder so, wie sie sich tatsächlich zugetragen haben (kongruente Videos), oder mit einer unerwarteten und „unüblichen" emotionalen Reaktion (inkongruente Videos). Es wurde die Hypothese aufgestellt, daß aufgrund unterschiedlicher Gefühle bei Versuchspersonen, je nachdem ob diese/r sich einfühlt oder bei den eigenen Emotionen bleibt, bei inkongruenten Videos Self-Other Distinction sichtbar wird. Emotionen wurden mithilfe eines Affect Rating Dials operationalisiert. In der Umsetzung wurden die Versuchspersonen gebeten sich einerseits in die Target-Person einzufühlen und deren Gefühlsvalenz einzuschätzen (perceivers' other-ratings) und andererseits sich selbst in diese Situation hineinzuversetzen und die eigene Gefühlsvalenz einzustufen (perceivers' self-ratings). Diese Einstufungen dienten als Grundlage der Datenanalyse (Varianzanalyse). Bei den inkongruenten Videos zeigte sich bei den Versuchspersonen eine signifikante Self-Other Distinction und somit können diese Videos als geeignete Stimuli zur Untersuchung von Self-Other Distinction im Kontext der Empathie angesehen werden.

Schlüsselwörter

Self-Other Distinction, Empathie, Inkongruente Video-Stimuli, Affect Rating Dial.

Danksagung

Als erstes möchte ich mich bei meinen Eltern bedanken, die mich stets unterstützt haben, sowohl emotional als auch finanziell und mir dieses Studium überhaupt ermöglicht haben. Ebenso, daß sie mir die Möglichkeit gaben, dieses Studium nicht nur als Wissensgewinn zu betrachten, sondern dabei auch die Zeit zu haben persönlich in vielen Bereichen zu Reifen und die Welt kennen zu lernen.

Großer Dank gilt natürlich meinem Betreuer Univ.-Prof. Mag. Dr. Claus Lamm für die kompetente, fachlich hochwertige und kollegiale Betreuung während der gesamten Diplomarbeitsphase.

Bedanken möchte ich mich auch bei all jenen, die mir in den verschiedensten Phasen meiner Diplomarbeit geholfen haben. Hierbei speziell bei Andreas Gartus für die tollen Rating Dials und die technische Unterstützung. Und Daniela Pfabigan für sovieles zwischendurch.

Ganz besonders möchte ich mich bei meinen Geschwistern bedanken. Regina Längle für die Unterstützung beim Filmen der Videos und das fachliche Wissen bei Photoshop. Und Arthur Längle für die Hilfe mit fachkundigem Wissen durch das Dickicht von R durchzukommen und die LATEX Tricks. Und Armina Längle für die vielen Aufmunterungen.

Ebenso vielen Dank an meine Mutter und meine Freunde Benjamin Weber, Kathrin Knapp (danke Kitzi auch für das Mikrophon) und Veronika Piringer für das Korrekturlesen.

Eine große Hilfe waren mir Annabel Müller und Sophie Müller durch die seelischen Aufbauarbeiten (mit und ohne Rotwein).

Natürlich möchte ich mich auch bei meinen tollen SchauspielerInnen bedanken. Zudem bei den vielen Versuchspersonen.

Zu guter Letzt bedanke ich mich besonders dafür, daß meine Freunde und meine Familie während meiner gesamten Studienzeit immer für mich da waren und sind.

Contents

1 Introduction

Social interaction and deep relationships seem to be essential for mankind. Empathy plays a fundamental part in social interactions and its research is the background of this thesis. Empathy is defined in many ways, one definition given by Rogers leads illustratively in this subject.

> The state of empathy, or being empathic, is to perceive the internal frame of reference of another with accuracy, and with the emotional components and meanings which pertain thereto, as if one were the person, but without ever losing the "as if" condition. Thus it means to sense the hurt or the pleasure of another as he senses it, and to perceive the causes thereof as he perceives them, but without ever losing the recognition that it is *as if* I were hurt or pleased, etc. If this "as if" quality is lost, then the state is one of identification. (Rogers, 1959, pp. 210-211).

One key aspect in empathy is this "as if" quality. In applied research this experiential description is grasped under the theoretical term *self-other distinction*. Self-other distinction refers to the ability to know if an affective state is triggered by another person or by one's own reason (de Vignemont & Singer, 2006; Decety & Jackson, 2004; Decety & Lamm, 2006). Self-other distinction is a prerequisite for empathy as is pointed out in the following definition of empathy. Empathy is "if: (i) one is in an affective state; (ii) this state is isomorphic to another person's affective state; (iii) this state is elicited by the observation or imagination of another person's affective state; (iv) one knows that the other person is the source of one's own affective state." (de Vignemont & Singer, 2006, p. 435).

Empathy is a very complex phenomenon which seems to have two different routes: experience sharing and mentalizing (Zaki & Ochsner, 2012). This differentiation is often observable in the literature concerning empathy. The terms *Affective Empathy* and *Cognitive Empathy* are frequently used to discriminate these two processes (e.g., Soto & Levenson, 2009; Shamay-Tsoory, 2011). These empathic processes have also been described using other terms. Terms to describe the top-down process are for example, beside cognitive empathy, perspective taking (Eslinger, 1998), Theory of Mind (ToM) (Premack & Woodruff, 1978), and mentalizing (Zaki & Ochsner, 2012). Whereas affective empathy, as a bottom-up process, is also meant with shared self-other representations (Zaki, Weber, Bolger, & Ochsner, 2009), or rather the Perception-Action Model (PAM) (Preston & de Waal, 2002), emotional contagion, and experience sharing (Zaki & Ochsner, 2012). Lamm, Nusbaum, Meltzof, and Decety (2007, p.12) discuss the experience of empathy as "a complex interplay between automatic and bottom-up driven and

controlled top-down processes that result in a joint but highly malleable and individual experience." While these terms describe different approaches to get in touch with another person's internal state, it further has to be distinguished between different kinds of relatedness, as for instance emotional contagion which is the tendency to automatically mimic another person and where self-other distinction is not required (Hatfield, Rapson, & Le, 2011), and empathy where self-other distinction is a prerequisite (de Vignemont & Singer, 2006).

Beside these attempts of operationalization and definitions, studies tried to reveal in different ways distinct aspects of empathy. Some studies concentrated on perceiver's *Empathic Accuracy* (Ickes, Stinson, Bissonnette, & Garcia, 1990; Stinson & Ickes, 1992; Simpson, Ickes, & Blackstone, 1995), which is "their ability to accurately infer the specific content of another person's thoughts and feelings" (Ickes, 1993, p. 588). Others concentrated more on self-other distinction as a basic ability for social interactions. Self-other distinction, the ability to separate the self from the other is a central element of empathy as described in the definitions above. Research in this field revealed the ontogenetic development of self-other distinction as a process through the life span (Deci & Ryan, 1990; Blatt, 1974). According to Bischof-Köhler (1994) the first step of this development to distinguish the self from others is expressed during the second year of life with the so called *self-objectivation*, indicated as recognizing oneself in a mirror. At the age of about three to four years the ability to distinguish between belief and reality develops (Sodian, 2002) and hence the so called *Level-2-Perspective Taking* (meaning to have the insight that an object seen by two persons from opposite sides has a different appearance to them) (Flavell, Everett, Croft, & Flavell, 1981).

The ability of an individual to impute mental states to oneself and to others is described by the term *Theory of Mind (ToM)* (Premack & Woodruff, 1978). Theory of Mind (ToM) as the ability to understand that other people have their own thoughts, beliefs, desires, and intentions is crucial for social interactions, and best seen in children with autism disorder who have a lack in ToM and as a result severe problems in daily social interactions (Baron-Cohen, Leslie, & Frith, 1985). Further the development of inhibitory control around the 4th birthday seems to be a crucial factor for the development of ToM (Carlson & Moses, 2001). Theory of Mind is a prerequisite for being able to identify with others and also to distinguish ourselves from others. This describes a duality of being related and autonomous which is crucial for our ability to interact with others (Decety & Sommerville, 2003).

Self-other distinction is studied in different domains. For instance in a spatial and cognitive context it can be examined by perspective taking tasks (Keysar, Barr, Balin, & Brauner, 2000; Santiesteban, Banissy, Catmur, & Bird, 2012). On a motor level, the ability to distinguish

the self from others is also necessary for the inhibition of a motor imitation (Brass, Ruby, & Spengler, 2009). And when empathy is considered as a state in which the empathizing person feels an emotion which is isomorph to that of the other person (de Vignemont & Singer, 2006) self-other distinction enables to assign the source of an emotion. For example, to notice someone else crying and being affected without distinguishing between self and other (i.e., there is a confusion between self and other), this observation could result in personal distress (Singer & Lamm, 2009). *Personal distress* means "an aversive, self-focused emotional reaction to the apprehension or comprehension of another's emotional state or condition" (Decety & Lamm, 2011, p. 199). In this regard to examine self-other distinction is a pivotal part in research concerning empathy and disorders regarding deficits in the ability to distinguish between the self and others. Nevertheless self-other distinction in an emotion and empathy related domain is rarely explored until now.

To examine self-other distinction in an emotional context presents a specific challenge. Although it is assumed that self-other distinction is a prerequisite to be empathic (de Vignemont & Singer, 2006) and therefore always present when feeling with somebody, under normal circumstances it is not observable. Because when someone feels with somebody, the emotions while being empathic have the same valence (are isomorph) as the feelings the target-person has. Furthermore one knows that the other is the source of the emotions which are felt. Hence to achieve that self-other distinction is measurable, it has to become visible. Therefor stimuli are required in which a tension between the self and other through a difference of emotional sensation is elicited. To explain what this means in concrete at first an operational description of self-other distinction will be deduced. According to the above definition of empathy of de Vignemont and Singer (2006) and in respect of self-other distinction this means that we should be able to feel with another person irrespectively if I myself would react emotionally the same way as the target-person does. Hence to be empathic in a given situation does not necessarily mean that I myself would feel the same way as the target-person, but it means that nevertheless I am able to feel with the other person, wherefore the ability to distinguish the other's feelings from my own one's is required. This opens the possibility to test if an empathic person is able to hold self-other distinction in any situation.

Empirically one can proceed in the way that situations are created in which the observer and the target would emotionally react in the same way (*congruent condition*) and other situations where the observer would not react the way the target does (*incongruent condition*) (Lamm, Meltzoff, & Decety, 2009). This means that the target-person describes for example a situation

where commonly negative emotions would suit this situation, but the target-person reports that he/she had positive feelings then. Signifying the emotions described by the target-person differ from the emotions one commonly would have when being in this situation. In contrast thereto in a congruent condition life-events are told as they actually happened with an expected and common emotional reaction of the target-person.

Therefore to examine self-other distinction in the context of empathy the development of new stimuli who are evoking an observable self-other distinction are required.

The operationalization of this study bases on the so called *Naturalistic Empathic Accuracy Paradigm*. There the general idea of how to measure empathic accuracy is "to compare the content of a target person's actual thoughts and feelings with the content of the corresponding inferred thoughts and feelings reported by the perceiver" (Ickes, 1993, p. 591). Empathic accuracy research is a performance-based measure and hence in contrast to self-report measures it "offers the advantage of capturing people's actual ability to 'read' others rather than just their perceived ability to do so" (Ickes, 2011, p. 66). Amongst others, one possibility to investigate empathic accuracy in a social interaction context is to videotape a target-person and compare "lists of thoughts and feelings", like Marangoni, Garcia, Ickes, and Teng (1995) did. In their study the target-persons got videotaped, while discussing an emotional loaded life-event. Immediately afterwards, in a so called *Target Phase*, each target-person reviewed the videotape of him-/herself and with a remote start/pause control stop the tape at each point the target-person remembered having had a special feeling or thought. This special feelings or thoughts were written down and so the target-person created a time-logged listing of *actual thought/feeling entries*. Afterwards in a so called *Perceiver Phase*, perceivers watched these videos and got instructed to infer the content of the client's thoughts and feelings and write them down, so that a time-logged listing of these *inferred thought/feeling entries* results. Subsequently, by comparing these two lists, a measure of *Empathic Accuracy* were calculated.

With the comparison of the target-person's and perceiver's "lists of entries" it is possible to distinguish between a thought and feeling. But a disadvantage of the study of Marangoni et al. (1995) is that writing down these lists needs a lot of time and is not on-line. Levenson and Ruef (1992) solved this problem by using an *Affect Rating Dial* to collect on-line data of the *actual* and *inferred* emotional valence and intensity. This affect rating dial is a box with an index placed upon a 9-point Likert scale and can be continuously adjusted at each point of time (Ruef & Levenson, 2007).

16

In this diploma thesis this naturalistic empathic accuracy paradigm is adapted to develop short video sequences as stimuli to examine not only empathic accuracy as the above mentioned studies did but also self-other distinction, which is achieved by having two conditions, the congruent and the incongruent condition. Further the combination with an affect rating dial enables to reflect the emotional state of participants by rating the emotional valence on-line, meaning at the same time as watching the video.

In these videos a target-person describes an autobiographical life-event in both, a congruent and an incongruent way. While in the congruent description the facts of the event and target's emotional reaction thereto fits and is as "normally expected", in the incongruent condition target's emotional reaction is unexpected and mostly opposite to the "expected emotion". The concrete implementation is described at section 2.3.3 at p. 35. The important achievement of this diploma thesis is that these incongruent videos provide one possibility to make self-other distinction visible.

The mean length of the videos is about 90 seconds and videos show a specific structure. For an overview see Figure 1.
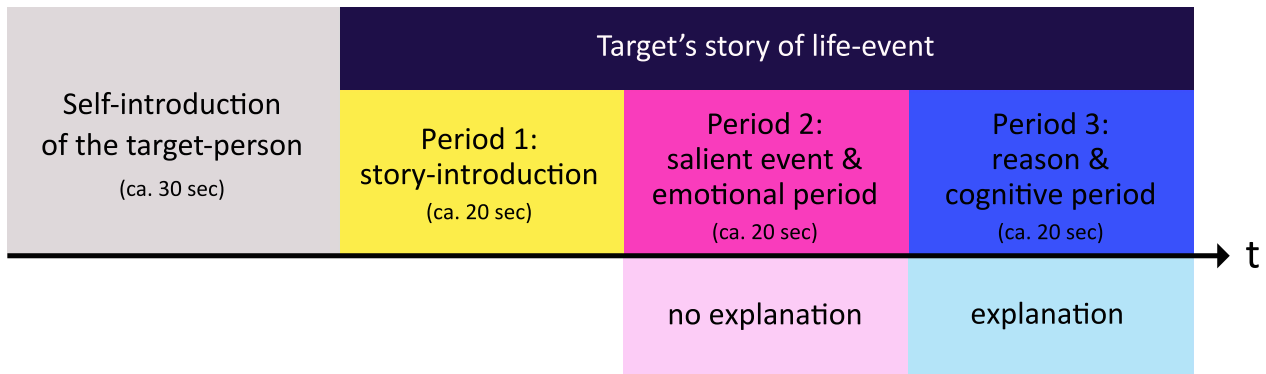

Figure 1: Time-structure of the stimuli for congruent and incongruent condition.

An important observation considered in this study is that we are not constantly empathizing with everyone, rather empathy has to be triggered (de Vignemont, 2006). In respect thereof in the first 30 seconds of the video the target-person gives a short self-introduction. This period constitutes a first approach to the target-person where he/she tells about him-/herself what he/she does, where he/she lives, comes from, etc. This should help participants to get familiar with the target's way of describing and facial expression and further it may increase

the willingness and accuracy of the perceiver to capture his/her *frame of reference*[1] and to empathize with the target-person by bringing the target-person closer and reducing distance.

After this self-introduction, three periods of interest (period 1, period 2, and period 3) follow. Period 1, the *period of story-introduction*, starts when the target-person initiates to tell the specific life-event and takes about 20–30 seconds. Because no specific emotions nor thoughts will be discussed, this period could be thought as a comparison-period to the second and third period.

In this complex phenomenon of empathy another duality is the interplay of bottom-up and top-down processes (Lamm et al., 2007). This point is clearly expressed too by de Vignemont (2006, p. 194) when discussing "however, we cannot account for empathy with a purely bottom-up process; we need the top-down input". To investigate the interplay and what role has the bottom-up process for empathizing, the next two periods have a clear cut in the description. Period 2 starts with the salient event and the focus lies on the emotions. In the incongruent condition the salient event displays a sudden change in the story into an unexpected direction and target-persons were instructed to describe only their emotions in this period. Therefore it is also referred as *salient event & emotional period*. After about 20 seconds the describing leads into the last period, period 3. Period 3 initiates when the target gives a reason for the before described emotions and the focus lies on the cognitive explanation. Period 3 is therefore also denoted as *reason & cognitive period*. With this structure is was tried to evoke a focus on either the affective empathy route or the cognitive empathy route in the perceiver's process of empathizing.

The duration of each period is counted from the onset-point on. The onset-point is always 2–3 seconds after the time-point when the target starts to discuss the specific content and this delay reflects the time perceivers need to capture the message. These three time-periods of interest (period 1, period 2, and period 3) serve as a basic material to collect data for the topic of this diploma thesis—to estimate if a video provides a useful stimulus to examine self-other distinction (at incongruent stimuli) and empathic accuracy (at congruent and incongruent stimuli).

As mentioned above an affect rating dial is used to provide data reflecting the valence of the emotional state of participants (the rating dial is described at section 2.1 at p. 31). Participants of the validation study watch each video twice and while watching they are asked to rate in the first run the emotional valence they refer the target-person and in a second run their own

---

[1] "We can achieve real communication [. . .] when we listen with understanding. This means seeing the expressed idea and attitude from the other person's point of view, sensing how it feels to the person, achieving his or her frame of reference about the subject being discussed." (Rogers & Roethlisberger, 1991, p. 106)

emotional valence they would have in this situation. But not only perceivers rating dial data are collected, also target's rate their own emotional valence. Hence following data were obtained: data from the targets when they rate his/her own emotional valence (target's self-rating), from the perceivers when perceivers put themselves in the watched situation and rate their own emotional valence (perceivers' self-ratings), and when perceivers feel with the target-person and rate the emotional valence they ascribe to the target-person (perceivers' other-ratings). With these rating data it is analyzed if an incongruent video achieves the aim of making self-other distinction visible which is operationalized as the difference between perceiver's other-rating (when rating the emotional valence they ascribe to the target-person) and their self-rating (rating their own emotional valence).

Furthermore some questionnaires and tasks who survey parts of empathy and/or self-other distinction were applied (for further information see 2.5.2, p. 41ff.). These questionnaires and tasks are assumed to correlate with perceiver's ability to empathize or to distinguish the self from others.

Emotions embody themselves in multiple forms, amongst others they influence physiological reactions, so changes in the somatic muscles and in the viscera may occur (Bradley & Lang, 2000). For example, when confronting individuals with emotional stimuli, their heart rate and skin conductance change depending on the emotional valence and arousal of the stimulus (Lang, Greenwald, Bradley, & Hamm, 1993). As reported in Levenson and Ruef (1992), observing a person in distress can effect the autonomic nervous system in the observer.

The Electrocardiogram (ECG) measures the cardiovascular system which shows strong reactions relating to psychological processes like emotions, attention, or stress (Gramann & Schandry, 2009). Skin conductance response (SCR) as a measure of Electro-dermal activity (EDA) is a non-suggestible dependent variable of the degree of arousal of the autonomic nervous system (ANS) (Schandry, 1996).

Literature concerning the observation of another person's emotional display and its result in similar emotions in the observer (e.g., Lanzetta & Englis, 1989) lead Levenson and Ruef (1992) to use in their study physiological measurements to investigate the physiological linkage between perceivers and targets. Hence in addition to the rating dial data, psycho-physiological data will be collected to examine the linkage between perceivers and target. Analysis of these psycho-physiological data will be undertaken in another master-thesis.

To summarize, the aim of this diploma thesis is to develop and validate videos as stimuli to examine self-other distinction. Therefore videos will be developed and validated which are told either in a congruent or incongruent way. Further the usage of an affect rating dial enables an on-line, second-to-second adjustment, reflecting the perceiver's or target's emotional valence. With these rating dial data at incongruent videos the extent of the difference in perceivers' self-ratings and perceivers' other-ratings will be calculated to prove the video's ability to make self-other distinction visible. And with target's self-rating and perceiver's other-rating it will be calculated if a video provides a useful stimulus to examine empathic accuracy (at both, incongruent and congruent videos).

Videos were chosen as material for these stimuli, so constantly changing emotions can be presented which provide quite naturalistic events. Further the videos' stories are descriptions of actual life-events and hence these videos present stimuli which are drawn from life.

Therefore important is also a kind of control if participants were able to engage with this kind of stimulus at all. Congruent videos as a stimulus for information and comparison can be used to examine perceiver's ability and commitment to engage with these kind of stimuli (i.e., to engage with an unknown target-person in a video-sequence who describes an event of his/her life and to feel with this person although there is no the possibility to interact and to communicate with this target-person).

So innovative in this diploma thesis is that the naturalistic empathic accuracy paradigm is used in a modified and extended way to examine not only empathic accuracy but also self-other distinction.

The steps for the development and validation of these stimuli can be subdivided into three main parts, which are constituting my diploma thesis:

I. The development of the stimuli. I.e., 1st videotaping of target-persons while describing an emotional life-event in congruent and incongruent ways. And 2nd target's data collection in which the target-persons were watching their own videos and rate their own emotional valence they had while describing their own life-events by using an affect rating dial.

II. A pilot-study to select the most differentiated videos.

III. And finally a validation study in which the selected videos get validated in respect of their ability to provide useful stimuli to examine self-other distinction and empathic accuracy by using perceivers' self-rating and their other-ratings.

1.1 Hypotheses

As written in the introduction, two types of videos were developed, congruent and incongruent ones, and in each video three periods of interest exist (period 1, period 2, and period 3). While at congruent videos the story is normal and reported emotions are as expected, at incongruent videos the emotions are not expected and don't suit the told situation (at period 2 and period 3). Further data from the target-person (target's self-rating) and from perceivers (perceiver's self-rating and perceiver's other-rating) were obtained. Hence the deduced main hypotheses are:

I. An incongruent video would provide a useful stimulus to examine self-other distinction (SOD) if perceivers' self-ratings differ from perceivers' other-ratings (= SOD is visible) at period 2 and period 3.

II. A congruent video would represent a convenient stimulus to examine empathic accuracy (EA) if perceivers' other-ratings match target's self-ratings at all three time-periods of interest. In addition an incongruent video would be an adequate stimulus to examine empathic accuracy (EA) if perceivers' other-ratings and target's self-rating match at all three time-periods.

III. A congruent video would provide an appropriate stimulus for information and comparison to examine participant's commitment to engage with this type of stimulus if perceivers' other-ratings and perceivers' self-ratings match (= $\neg SOD$, meaning SOD is not visible) at all three time-periods of interest.

Analogue to the three main hypotheses, three parts of concrete hypotheses will now be presented. At the beginning of each part a definition of the focused object (SOD, EA, or $\neg SOD$) and its mathematical expression ($|SOD|$ and $|EA|$)^2 will be given. Followed by the concrete hypotheses concerning the incongruent or congruent videos and applied questionnaires.

I. **Incongruent videos as stimuli to examine self-other distinction (SOD).**
    SOD (self-other distinction) is assigned as the difference between perceiver's other-rating-data (when the perceiver feels with the target-person and rates the target-person's emotional valence) and perceiver's self-rating-data (when the perceiver puts him-/herself in

---

^{2}|| stands for their absolute value.

this situation and rates his/her own emotional valence). The distance between these rating-data describes the score of self-other distinction.

$$|SOD| = |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})|$$

$|SOD|$ correlates positively with self-other distinction (SOD).

(a) *SOD (self-other distinction) at incongruent videos.*

SOD is visible when perceivers' self-ratings differ significantly from perceivers' other-ratings ($|SOD|$ is significant). Since the twist into incongruity starts with the second period, the distinction is expected in period 2 (the salient event & emotional period) and period 3 (the reason & cognitive period). No significant difference is expected in period 1 (the story-introduction) because no incongruent content is discussed there. Hence an incongruent video provides a useful stimulus to examine self-other distinction if $|SOD|$ is significant (i.e., SOD is visible) at period 2 and period 3.

$H_0 : |SOD|_{\text{(at period 1)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{sig.}$
$H_1 : |SOD|_{\text{(at period 1)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{n.s.}$

$H_0 : |SOD|_{\text{(at period 2)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{n.s.}$
$H_1 : |SOD|_{\text{(at period 2)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{sig.}$

$H_0 : |SOD|_{\text{(at period 3)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{n.s.}$
$H_1 : |SOD|_{\text{(at period 3)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{sig.}$

Because it is expected that participants distinguish in both, the salient event & emotional period (period 2) and the reason & cognitive period (period 3), between their own (self-rating) and target's feelings (other-rating), following interaction effects regarding Self-Other Distinction (i.e., Type of Rating) x Time are expected: significant interaction effects are expected for SOD x Period 1 vs. 2, and for SOD x Period 1 vs. 3. Because it is not clear how perceivers' SOD at period 2 behaviours to their SOD at period 3, the interaction effect of SOD x Period 2 vs. 3 will be analyzed exploratory.

$H_0 : |SOD|_{\text{(at period 1)}} = |SOD|_{\text{(at period 2)}}$
$H_1 : |SOD|_{\text{(at period 1)}} \neq |SOD|_{\text{(at period 2)}}$

$H_0 : |SOD|_{\text{(at period 1)}} = |SOD|_{\text{(at period 3)}}$

$$H_1 : |SOD|_{(\text{at period 1})} \neq |SOD|_{(\text{at period 3})}$$

$$\text{exploratory} : |SOD|_{(\text{at period 2})} \gtreqless |SOD|_{(\text{at period 3})}$$

(b) *The correlation of SOD (self-other distinction) and the applied questionnaires.*

It is assumed that self-other distinction correlates positively with questionnaires and tasks who survey kinds of self-other distinction, perspective taking, or theory of mind. As written above SOD is assigned as the difference between perceiver's other-rating and perceiver's self-rating. Since $|SOD|$ correlates positively with self-other distinction (SOD), consequently a relation between self-other distinction and questionnaires/tasks is displayed by a positive correlation.

In general this implies that when perceivers' self-ratings differ from perceivers' other-ratings (SOD is visible, SOD is significant), a positive correlation with questionnaires/tasks concerning self-other distinction related abilities is expected. Since it is assumed that at incongruent videos SOD is visible at period 2 and period 3, the hypotheses are as followed:

H_1 at period 1 : $corr(|SOD|, Questionnaires) = $ direction is not specified

H_1 at period 2 : $corr(|SOD|, Questionnaires) = $ positive

H_1 at period 3 : $corr(|SOD|, Questionnaires) = $ positive

For the specific hypotheses concerning every questionnaire, see page 27.

II. **Congruent and incongruent videos as stimuli to examine empathic accuracy (EA).**

Empathic accuracy (EA) is defined as the similarity of perceivers' other-ratings (when perceivers feel with the target and rate the target-person's emotional valence) and target's self-rating (when the target rates his/her own emotional valence they had while describing the story).

$$|EA| = |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})|$$

EA is thought to be observable at both conditions, congruent and incongruent, as the match/overlap of target's self-rating and participant's other-rating.

(a) *EA (empathic accuracy) at congruent and incongruent videos.*

It is hypothesized that a congruent or a incongruent video provide a useful stimulus

to examine empathic accuracy (EA) if perceivers' other-ratings match target's self-rating i.e., EA is not significant at all time-periods of interest.

$H_0 : |EA|_{\text{(at period 1)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})| = \text{sig.}$

$H_1 : |EA|_{\text{(at period 1)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})| = \text{n.s.}$

$H_0 : |EA|_{\text{(at period 2)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})| = \text{sig.}$

$H_1 : |EA|_{\text{(at period 2)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})| = \text{n.s.}$

$H_0 : |EA|_{\text{(at period 3)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})| = \text{sig.}$

$H_1 : |EA|_{\text{(at period 3)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{target's self-rating}})| = \text{n.s.}$

(b) *The correlation of EA (empathic accuracy) and the applied questionnaires.*

It is assumed that at congruent and incongruent stimuli empathic accuracy correlates positively with questionnaires and tasks which survey empathy, parts of empathy, or empathic behaviour (at all time-periods of interest). Due to the assignment of empathic accuracy (EA) as the similarity of perceiver's other-rating and target's self-rating, this implies that the more similar the ratings (represented by $|EA| \to 0$), the greater is perceiver's empathic accuracy. Therefore $|EA|$ correlates negatively with EA (meaning a small value of the ratings' difference means strong empathic accuracy and the larger the $|EA|$-score the poorer EA).

This means that (thematically) questionnaires/tasks are expected to correlate positively with empathic accuracy (EA), since (as numerical value) EA is negatively correlated with $|EA|$, hence the questionnaires/tasks are expected to correlate negatively with $|EA|$.

Hence the hypotheses are as followed (for congruent and incongruent stimuli):

H_1 at period 1 : $corr(|EA|, Questionnaires) = \text{negative}$

H_1 at period 2 : $corr(|EA|, Questionnaires) = \text{negative}$

H_1 at period 3 : $corr(|EA|, Questionnaires) = \text{negative}$

For the specific hypotheses concerning every questionnaire, see page 27.

Ad hoc hypothesis to measure "empathic accuracy".

The definition of $|EA|$ as a measure of empathic accuracy (EA) uses target's self-rating as a basis to calculate perceiver's empathic accuracy (see therefore p. 23). But as seen at

Figure 27 to Figure 38 (p. 155–160) target's self-rating sometimes does not even go into the same direction as the emotional valence the target displayed in the video (this will be discussed in section 4.3.5). Hence target's self-rating doesn't provide useful data as a basis to compare perceivers' other-ratings with it, and to calculate the extent of empathic accuracy. Therefore ad hoc a way to operationalize EA independent of target's self-rating was considered. This alternative possibility uses the "statistical norm" and pleads the plausibility-argument that with a high probability other-ratings of the majority match the actual emotional valence of the other person. To test if this argument is legitimate and firm would be topic of another study. Anyway, "empathic accuracy" is assigned as the match of a single participant's rating when being empathic with the target-person and rating the target's emotional valence (perceiver's other-rating) and all participants' ratings when being empathic with the target-person and rate the target's emotional valence (other-ratings of all participants):

$$|\text{``}EA\text{''}| = |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{mean of all perceivers' other-ratings}})|$$

This determination implies (equally to $|EA|$ at page 23) that $|\text{``}EA\text{''}|$ correlates negatively with EA (small values of differences between the ratings ($|\text{``}EA\text{''}| \to 0$) means strong perceiver's "empathic accuracy"). Hence $|\text{``}EA\text{''}|$ correlates negatively with EA (a small value of ratings' difference means strong "empathic accuracy"). In general this means (equally to $|EA|$) that when perceivers' other-ratings and the mean of all perceivers' other-ratings match ($|\text{``}EA\text{''}| \to 0$), then there is a positive correlation (represented by a negative correlation, because a match is assigned as $|\text{``}EA\text{''}| \to 0$) with questionnaires/tasks which survey empathy-related abilities.

Hypotheses for $|\text{``}EA\text{''}|$ are based on the hypotheses of $|EA|$. Hence it is expected that at congruent and incongruent videos at all three time-periods of interest perceivers' other-ratings and the mean of all perceivers' other-ratings match and therefore a relation (negative correlation) is expected at all time-periods:

$$H_1 \text{ at period } 1 : corr(|\text{``}EA\text{''}|, Questionnaires) = \text{negative}$$

$$H_1 \text{ at period } 2 : corr(|\text{``}EA\text{''}|, Questionnaires) = \text{negative}$$

$$H_1 \text{ at period } 3 : corr(|\text{``}EA\text{''}|, Questionnaires) = \text{negative}$$

Due to the hypotheses of $|\text{``}EA\text{''}|$ concerning the questionnaires have the same direction as hypotheses of $|EA|$, the specific hypotheses concerning the applied questionnaires from $|EA|$ can be adopt one-to-one for $|\text{``}EA\text{''}|$ (for the specific hypotheses see page 27).

III. **Congruent videos as stimuli for information and comparison to examine commitment.**

At congruent videos the story is told as it actually happened with common and expected emotions. Hence is it assumed that similar emotions emerge in perceivers at both, when they put themselves in this situation and rate their own emotional valence (self-rating) as well as when they feel with the target-person and rate his/her emotional valence (other-rating). This similarity would be reflected by a match of perceivers' other-ratings and perceivers' self-ratings ($\neg SOD$).

Due to the definition of SOD as the visible and significant difference between perceivers' self-ratings and perceivers' other-ratings (p. 22), $\neg SOD$ means that perceivers' self-ratings do not differ significantly from perceivers' other-ratings (i.e., perceivers' self-ratings and perceivers' other-ratings match).

It is assumed that a congruent video provides an appropriate stimulus for information and comparison to examine participant's commitment (and ability) to engage with this type of stimulus if self-other distinction is not visible ($\neg SOD$).

(a) *$\neg SOD$ at congruent videos.*

Hence a congruent video would provide a useful stimulus for information and comparison if there is neither a significant main effect between perceivers' self-ratings and perceivers' other-ratings at any time-period

$$H_0 : |SOD|_{\text{(at period 1)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{sig.}$$
$$H_1 : |SOD|_{\text{(at period 1)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{n.s.}$$

$$H_0 : |SOD|_{\text{(at period 2)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{sig.}$$
$$H_1 : |SOD|_{\text{(at period 2)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{n.s.}$$

$$H_0 : |SOD|_{\text{(at period 3)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{sig.}$$
$$H_1 : |SOD|_{\text{(at period 3)}} : |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})| = \text{n.s.}$$

nor any significant interaction effect between Time x Other/Self.

$$H_0 : |SOD|_{\text{(at period 1)}} \neq |SOD|_{\text{(at period 2)}}$$
$$H_1 : |SOD|_{\text{(at period 1)}} = |SOD|_{\text{(at period 2)}}$$

$$H_0 : |SOD|_{\text{(at period 1)}} \neq |SOD|_{\text{(at period 3)}}$$
$$H_1 : |SOD|_{\text{(at period 1)}} = |SOD|_{\text{(at period 3)}}$$

$$H_0 : |SOD|_{\text{(at period 2)}} \neq |SOD|_{\text{(at period 3)}}$$
$$H_1 : |SOD|_{\text{(at period 2)}} = |SOD|_{\text{(at period 3)}}$$

The single hypotheses of SOD and EA with respect to the applied questionnaires.
The applied questionnaires are described more precisely at section 2.5.2 (p. 41ff.). The following
formulations are describing the hypothesized correlations in regard to their theoretical concepts.
It has to be considered that the operationalization into measuring values may lead to inversed
correlations, since $|EA|$ correlates negatively with EA.

1. The *Saarbrücker Persönlichkeitsfragebogen (SPF)* (Paulus, 2009) is the german version
   of the *Interpersonal Reactivity Index (IRI)* (Davis, 1980) and measures four separate
   aspects of empathy: perspective taking (PT), fantasy (FS), empathic concern (EC), and
   personal distress (PD). Perspective taking, fantasy, and empathic concern can be added
   and generate the scale "empathy" which is expected to correlate positively with empathic
   accuracy. PT is hypothesized to correlate positively with empathic accuracy and self-other
   distinction, FS and EC positively with empathic accuracy and negatively with self-other
   distinction, and PD negatively with empathic accuracy but positively with self-other
   distinction.

2. Alexithymia is expected to correlate negatively with empathic accuracy and self-other
   distinction. Because people with high scores in the *Bermond–Vorst Alexithymia Ques-
   tionnaire (BVAQ)* (Müller, Bühner, & Ellgring, 2004) are considered alexithymic, a neg-
   ative correlation between BVAQ (of each sub-scale, both higher-order factors, and of the
   total sum of all sub-scales) and empathic accuracy or self-other distinction is expected.

3. The *Emotional Contagion Scale* (Doherty, 1997) as a measure of the tendency to auto-
   matically mimic and to synchronize with the expressions of others is expected to correlate
   positively with empathic accuracy and negatively with self-other distinction.

4. *Reading the Mind in the Eyes Task* (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb,
   2001) as a measure of mentalising is expected to correlate positively with both empathic
   accuracy and self-other distinction.

5. The *Imitation Inhibition Task* (Brass, Bekkering, Wohlschläger, & Prinz, 2000) and

6. the *Perspective Taking Task* (Santiesteban et al., 2012) are both measures of self-other
   distinction in different contexts and therefore it is expected that both tasks correlate
   positively with self-other distinction.

7. The Post-Video Questionnaire (PVQ) of the Validation Study: the first three questions (concerning target's (1) attractiveness, (2) how interesting, and (3) likeable he/she is) are expected to correlate positively with empathic accuracy. The next two questions ((4) if the perceiver had a want to distance from the story or (5) if he/she divagated/zoned out) are expected to correlate negatively with empathic accuracy but positively with self-other distinction. The last question ((6) if the perceiver was already in a similar situation) is expected to correlate positively with empathic accuracy and concerning self-other distinction no direction will be hypothesized.

Related to the operationalization[3] of the self-other distinction and empathic accuracy the statistical hypotheses will be as followed:

1. Saarbrücker Persönlichkeitsfragebogen (SPF) consists of the subscales perspective taking (PT), fantasy (FS), empathic concern (EC) and personal distress (PD). An empathy-scale can be calculated by adding PT + FS + EC.

    H_1 at period 2/3 : $corr(|SOD|, SPF_{\text{PT}})$ = positive

    H_1 at period 2/3 : $corr(|SOD|, SPF_{\text{FS}})$ = negative

    H_1 at period 2/3 : $corr(|SOD|, SPF_{\text{EC}})$ = negative

    H_1 at period 2/3 : $corr(|SOD|, SPF_{\text{PD}})$ = positive

    H_1 at period 2/3 : $corr(|SOD|, SPF_{\text{Empathy}})$ = not specified

    H_1 at period 1/2/3 : $corr(|EA|, SPF_{\text{PT}})$ = negative

    H_1 at period 1/2/3 : $corr(|EA|, SPF_{\text{FS}})$ = negative

    H_1 at period 1/2/3 : $corr(|EA|, SPF_{\text{EC}})$ = negative

    H_1 at period 1/2/3 : $corr(|EA|, SPF_{\text{PD}})$ = positive

    H_1 at period 1/2/3 : $corr(|EA|, SPF_{\text{Empathy}})$ = negative

2. Bermond–Vorst Alexithymia Questionnaire (version B) (BVAQ-B) consists of the subscales Emotionalizing, Fantasizing, Identifying, Analyzing, and Verbalizing; further of two higher-order factors, the affective factor and cognitive factor, and finally of the total sum of all scores (*allSum*).

---

[3]As written above $|SOD|$ correlates positively with SOD, whereas $|EA|$ correlates negatively with EA.

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{Emotionalizing}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{Fantasizing}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{Identifying}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{Analyzing}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{Verbalizing}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{affective factor}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{cognitive factor}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, BVAQ\text{-}B_{\text{allSum}}) = $ negative

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{Emotionalizing}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{Fantasizing}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{Identifying}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{Analyzing}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{Verbalizing}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{affective factor}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{cognitive factor}}) = $ positive

H_1 at period $1/2/3$: $corr(|EA|, BVAQ\text{-}B_{\text{allSum}}) = $ positive

3. Emotional Contagion Scale (EC) is measured by calculating the total-score of all items. But it is also possible to analyze the items separate for emotional contagion by positive emotions and emotional contagion by negative emotions.

H_1 at period $2/3$: $corr(|SOD|, EC_{\text{positive}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, EC_{\text{negative}}) = $ negative

H_1 at period $2/3$: $corr(|SOD|, EC_{\text{allMean}}) = $ negative

H_1 at period $1/2/3$: $corr(|EA|, EC_{\text{positive}}) = $ negative

H_1 at period $1/2/3$: $corr(|EA|, EC_{\text{negative}}) = $ negative

H_1 at period $1/2/3$: $corr(|EA|, EC_{\text{allMean}}) = $ negative

4. Reading the Mind in the Eyes Task (RM):

H_1 at period $2/3$: $corr(|SOD|, RM) = $ positive

H_1 at period $1/2/3 : corr(|EA|, RM) =$ negative

5. Imitation Inhibition Task / Finger Lifting Task (FingLift):

   H_1 at period $2/3 : corr(|SOD|, FingLift) =$ negative

   H_1 at period $1/2/3 : corr(|EA|, FingLift) =$ negative

6. Perspective Taking Task / Shelve Task (shelveTask):

   H_1 at period $2/3 : corr(|SOD|, shelveTask) =$ positive

   H_1 at period $1/2/3 : corr(|EA|, shelveTask) =$ negative

7. The Post-Video Questionnaire (PVQ) of the Validation Study:

   H_1 at period $2/3 : corr(|SOD|, PVQ_{\text{interesting}}) =$ not specified

   H_1 at period $2/3 : corr(|SOD|, PVQ_{\text{attractive}}) =$ not specified

   H_1 at period $2/3 : corr(|SOD|, PVQ_{\text{likeable}}) =$ not specified

   H_1 at period $2/3 : corr(|SOD|, PVQ_{\text{distancing}}) =$ positive

   H_1 at period $2/3 : corr(|SOD|, PVQ_{\text{divagate}}) =$ positive

   H_1 at period $2/3 : corr(|SOD|, PVQ_{\text{situation}}) =$ not specified

   H_1 at period $1/2/3 : corr(|EA|, PVQ_{\text{interesting}}) =$ negative

   H_1 at period $1/2/3 : corr(|EA|, PVQ_{\text{attractive}}) =$ negative

   H_1 at period $1/2/3 : corr(|EA|, PVQ_{\text{likeable}}) =$ negative

   H_1 at period $1/2/3 : corr(|EA|, PVQ_{\text{distancing}}) =$ positive

   H_1 at period $1/2/3 : corr(|EA|, PVQ_{\text{divagate}}) =$ positive

   H_1 at period $1/2/3 : corr(|EA|, PVQ_{\text{situation}}) =$ negative

2 Methods

2.1 The Affect Rating Dial

To make participant's emotions visible and measurable the use of an affect rating dial (described in Ruef & Levenson, 2007) was chosen, analogue as it was used in the study of Levenson and Ruef (1992). The affect rating scale is infinitely variable because the index is contentiously adjustable between the poles *very negative* and *very positive*. Further participants can rate their emotional valence while watching the videos. Hence a second-by-second attunement of their emotional valence is possible. Before each measurement a rating dial calibration was undertaken. There the dial had to be moved to *very negative* and after three seconds to *very positive*, and after another three seconds to the middle (*neutral*). This helped afterwards to calculate the exact range of the score. For further calculations a continuous score was used. This means that on the basis of the calibration, for each participant a normalized rating was calculated, with the poles of the scale of the rating dial represented by -1 and $+1$. Hence the maximum of the achievable distance between the two poles by the rating dial is 2 points.
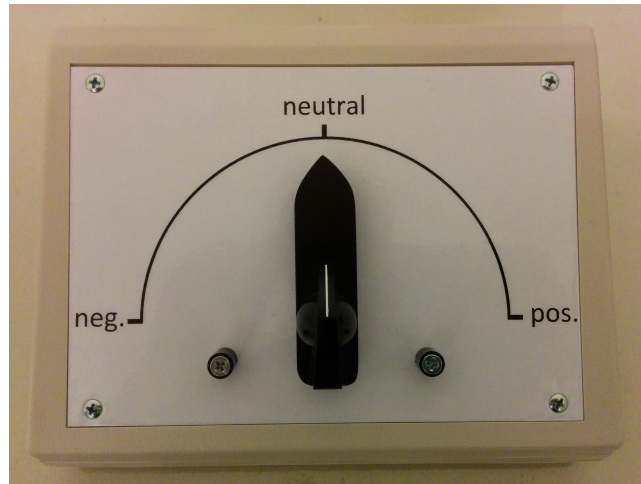
Figure 2: The affect rating dial.

For the construction of the rating dial, the selection of the dial-button was based on the consideration of having a haptic information whereto the index points (see Figure 2). The participant should be able to feel at what position the index is currently. So that it is not necessary to look constantly on the rating dial or to switch permanently the look between the screen and the index. This enables that the concentration may focus on the video and eliminates possible confounder.

The instruction for the usage of the rating dial always pointed out that it should be used

31

continuously and adjusted as often as they sense that, depending on the instruction, their own or the target's emotions change. Hence the pointer should always indicate the emotional valence.

2.2 Psycho-physiological Data

Psycho-physiological data of each target-person (the "actors" and "actresses") and of each participant of the validation study were collected. The collected psycho-physiological data were

- heart rate by means of an Electrocardiogram (ECG) and
- Electro-Dermal Activity (EDA), respectively Skin Conductance Response (SCR).

For EDA-application the electrodes were fixated at the volar middle phalanx of the index and ring-finger of the non-dominant hand (Boucsein et al., 2012). Boucsein et al. (2012) recommends to applicate the electrodes at the distal phalanges but the distal phalanges especially at women were too slim for an appropriate fixating of the EDA-electrodes. Hence for having a uniform measurement, the middle phalanges were chosen for application at all participants of the validation study and the target-persons. For ECG-application one electrode was attached on the right lateral clavicle and the other at the left rib (Jennings et al., 1981). The ground electrode was fixated at the wrist of the non-dominant hand (Boucsein et al., 2012). Before fixating the electrodes the area of the skin where the electrodes were placed was cleaned with ethyl alcohol (70%) (Gramann & Schandry, 2009). And for consumption of stimulants as coffee and tobacco was asked (Jennings et al., 1981). All exclusion criteria for ECG and EDA are listed in Table 1.

Data collection of the physiological measurements took place in the laboratory of the SCAN-Unit of the Department of Psychology of the University of Vienna while watching the videos and rating the own or the other's emotional valence (target's self-rating, perceiver's other-rating, and perceiver's self-rating). All over the time of the experiment target's or participant's heart rate (ECG) and skin conductance response (SCR) was recorded. Because movement or taking a deep breath disturbs the exactness of the measurement (Gramann & Schandry, 2009), participants were asked to seat comfortable but to move as little as possible during the session. Before starting a video, there was always a rest period of 30 seconds. This time-period's psycho-physiological data serves as a baseline of physiological arousal to further calculate the individual's relative variation (Boucsein et al., 2012).

With these data psycho-physiological linkage between target's and perceivers as well as intra-correlations of perceivers while being empathic or concentrate on the own person will be

Table 1: Exclusion Criteria for ECG and EDA

Exclusion Criteria
For both (ECG & EDA)
Disorder of the central nervous system
Psychiatric Disorder (e.g., Depression, Schizophrenia,...)
Use of neuropsychotropic medications
Big amount of alcohol or other psychotropic substances in the last 24 hours
Dependency of drug, alcohol, or pharmaceutical drugs
Insufficient sleep in the last 24 hours
Metallic objects on the body/in clothing
Additional for ECG
Prior negative reactions to ECG
Heart disease (Heart-Arrhythmia, cardiac pacemaker)
Additional for EDA
Prior negative reactions to EDA
Chronically dermatological condition
Additional will be asked for
the consume of stimulants like caffeine or nicotine
Handedness (EDA application will be on the non-dominant hand)
Participants will be informed about the possible side effects of ECG and EDA, such as itching sensation or skin irritation under the electrodes.
The signing of a written informed consent was requisite.

calculated in another master-thesis.

2.3   I. Development of Stimuli

2.3.1   Sample.

There existed considerations of taking an actor/actress as target person because of their special training in emotional perception (of their own feelings) as well as emotional expression (facial expression and gesture). But because it should be a natural situation and not a play, it was refrained from this idea. Persons with a pleasant voice, an average in physical attractiveness and middle aged seemed to be the best targets to evoke the willingness for empathizing in a sample consisting mostly of students (at the validation study).

By asking my circle of friends and their friends, 11 persons were found who were willing and interested in getting filmed while describing emotional loaded life-events. Five female and six male constituted the sample of target-persons. Their age varied from 22 to 34 ($M = 27.36$, $SD = 3.78$). Two of the target-persons are acting occasionally and one target-person was in a drama school (anyway, from this target-person no video was validated). Five target-persons

were students, the other six employee. All target-persons hat to fulfill the requirements for ECG and EDA application and candidates were scanned in advance therefor.

2.3.2 Measuring instrument and statistical analysis & software.

Participants for the development of stimuli were selected and scanned by their handedness (only right-handed persons were allowed) and fitting for ECG and EDA. Further two expressivity questionnaires were applied to measure target's expressivity.

- ***Edinburgh Händigkeits-Inventar*** (the German version of *The Edinburgh Inventory*) (Oldfield, 1971) was used to check participant's handedness. A total of 10 daily motor activities are listed: writing; drawing; throwing; holding scissors, toothbrush, knife, spoon, or broom; lighting a match; and opening a box. The subjects are asked to indicate which hand (left or right) they prefer to perform each activity.

- The ***Exclusion Criteria for ECG and EDA*** are listed in Table 1 at p. 33. Participants were informed about the technical equipment (ECG and EDA), the process of the study, exclusion criteria, and possible side-effects of ECG and EDA (like itching). After this they were asked if they had any questions thereto and to sign the exclusion criteria for ECG and EDA (the informed consent is attached at page 150f.).

- Some ***socio-demographical data*** as age, sex, highest level of education, current employment, and prior experience with acting were collected (see page 149). These data were collected to encompass possible confounder.

- One of the two applied expressivity questionnaires was the ***Berkeley Expressivity Questionnaire (BEQ)*** (Gross & John, 1997). The BEQ measures the individual's emotional expressivitiy which is divided into 3 facets: Negative Expressivity, Positive Expressivity, and Impulse Strength. These facets can be summated to calculate the General Expressivity. The BEQ consists of 16 items, each item is to be answered on a 7-point Likert-scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The BEQ was adapted into German by Mohiyeddini, John, and Gross (2008) and their validation showed Cronbach's α for the subscale Positive Expressivity of .86, Negative Expressivity of .84, and Impulse Strength of .81.

- To measure target's expressivity more differentiated, the ***Fragebogen der Facetten emotionaler Expressivität (FFEE)*** (Tausch, 2006) was applied. The FFEE is a self-rating questionnaire and each item is to be answered on a 4-point scale. The 70

questions measure emotional expressivity on 3 ranges of keys: I. core-expressivity with the dimensions *Impulsstärke* and *Expressivität*, II. social expressivity with the dimensions *Darstellung* and *Verstellung* and a further differentiation by the subscales between the ability and on the other hand the tendency of each dimension, and III. cognitive expressivity with the dimensions *Gefühlsklarheit* and *Ambivalenz*. Subscales' Cronbach's α vary between .66 and .88.

- After video taping, targets were asked to permit the use of the videos for further investigations and experiments.

For the development of the stimuli, also a self-rating of the target-person's were required (target's self-rating). This means that every target-person came to the laboratory of the SCAN-Unit and watched their own videos and while watching they were requested to rate continuously their own emotional valence they had while describing (not during the event itself) those life-events by using the affect rating dial. Furthermore while watching and rating their own videos, ECG and EDA data were collected. Therefore at this 2nd session

- for the ***consumption of stimulants*** was asked (how much coffee and/or tobacco was consumed before coming to the laboratory and if this amount is consistent to the normal/daily quantity), because of their influence on cardiovascular functions, which further effects ECG (Jennings et al., 1981).

Equipment for filming. The filming took place in a studio of a professional photographer in Vienna. To ensure similar conditions of light, a portable floodlight and LED-lights were employed. For filming a Canon 5D Mark II camera with a Canon 24-70mm; 1:2,8f objective was used. An external microphone was installed to record good voice quality.

For target's self-rating, as well as for the validation study, tasks were programmed by using E-Prime (2012). Therefor at first Adobe Photoshop (Photoshop, 2013) was used to render the videos. Then by using Super (2013) the videos were converted into MSmpeg4vs2. These video-files were included in the task which was presented with E-Prime (2012).

2.3.3 Experimental design.

Each target-person came to the studio of a professional photographer in Vienna. At first the target-person got informed about the exact aim and that their part at this study consisted of two sessions. The first session consisted of filming the videos and the second session took

place in the laboratory of the SCAN-Unit of the Department of Psychology at the University of Vienna for ECG and EDA application and target's self-rating. When they agreed to this process, the first part, the filming started. Therefor target-persons were asked to sign the informed consent concerning the exclusion criteria for ECG and EDA and to fill out the above explained questionnaires. Then the target-person was requested to list and entitle the 5 most positive and 5 most negative autobiographical events they were willing to describe in front of a camera, similar to the study of Zaki et al. (2009). When listed about 4–5 positive and negative life-events, they were asked to rate each event for emotional valence and intensity by using a 9-point Likert scale. Only events with ratings of their emotional intensity above the scale's midpoint were used further. For each target-person 2–3 negative and 2–3 positive events were selected to be described. The setting for recording the videos was for each target-person the same. Target-persons sat in front of a white wall and only the face and upper part of the body was seen in the video. This setting was chosen to enable perceivers to see target's mimic as well as their gestures. And the homogeneous background was selected to minimize the possibility to get distracted or influenced by the surroundings. The first idea of seeing more a dyadic setting was replaced because so the target's facial expression can be observed better by the perceiver.

The aim was to develop videos with two different conditions, congruent videos and incongruent videos. The conditions are described at p. 15 and the exact time-structure of the videos at p. 17. First the development of the videos for the congruent condition will be described and afterwards the development of incongruent condition's videos. And finally the session in the laboratory, target's self-rating, will be explained.

Congruent Condition. The target person was seated in front of the camera. As just mentioned, the frame of the video showed the target-person from the chest up and directly faced. Before starting the videotaping the target-person was instructed about the time-structure of the video (see p. 17). Then the target-person was asked to take a minute for re-empathizing and evoking the emotions they had at the time of the event which was chosen for describing. When they thought they were "captured" in this event, the describing and filming was initiated. Each video started with an initial "neutral introduction" by the target-person where he/she introduced him-/herself. This self-introduction was thought to help the perceiver to get familiar with the target-person's way of expression and mimic and lasted about 30 seconds. After this initial introduction of themselves, targets described both, the details of the life-event and the emotions they felt during the event. The average length of the described autobiographical event was approximately one minute. So in total, with the self-introduction, videos last about

90 seconds. Each life-event was described and filmed that often until the target-person and I were content.

Incongruent Condition.   Quite similar to the congruent condition's videos, the development of videos for the incongruent condition was proceeded. The difference between the congruent condition and the incongruent condition was the unexpected emotional reaction of the target-person in the latter condition. Hence the target-person and I considered of possible unexpected, mostly opposite emotions to the actual life-event and a reason to explain those emotions. For example one life-event treats of the story that the best friend became a father. The unexpected emotion thereto was that the target-person described that he was not happy. And the reason therefor was that he felt that his best friend got everything and he has reached nothing.

When we came to an agreement about the content and what emotions were described in the incongruent condition, targets were instructed to introduce him-/herself similarly to the congruent condition. After this initial self-introduction the story-introduction (period 1) was initiated. Subsequently to the story-introduction, the turn into incongruity started. Targets described the unexpected emotions to the chosen life-event without giving a reason for about 20–30 seconds (period 2: the salient event & emotional period). After this period targets explained their emotions by giving a reason (period 3: reason & cognitive period). The total lengths of the incongruent videos last as well about 90 seconds. Videotaping was repeated as often until we were satisfied with the result. The narrating was tried to be as authentic as possible, especially in the incongruent condition.

When all life-events which the target-person was willing to describe in a congruent and incongruent way were recorded, the target-person was asked to sign a letter of agreement to permit the further use of these videos.

Target's self-rating.   The second part, target's self-rating took place in the laboratory of the SCAN-Unit of the Department of Psychology of the University of Vienna. Targets were requested to watch their own videos of their described autobiographical events and while watching to re-feel and rate the emotional valence they had while describing by using the affect rating dial (the affect rating dial is described in section 2.1 at p. 31). The instruction pointed out that the rating should reflect the emotional valence they had while describing and not the emotional valence they had at the event itself. The index of the rating dial should always point to the emotional valence they had while describing. The aimed result was a continuous rating of the target's emotional valence—target's self rating. It was very important that targets' self-ratings

was as exact as possible. Therefore targets could repeat their ratings as often as they wished until they found their ratings reflected the emotional valence they had while describing.

While watching target's ECG and EDA data were collected. ECG and EDA application was described precisely at section 2.2 (p. 32).

Out of these total set of 101 recorded videos, those videos which seemed to be most proper to elicit an emotional reaction in the perceiver and where the narration felt authentic were preselected for further evaluation in the pilot study. 27 congruent and 28 incongruent videos (in total 55 videos) remained and constituted the pool for the pilot study.

2.4   II. Pilot Study

The pilot study was an important step for the selection of videos which were on the one hand emotionally differentiated i.e., that the feelings of the target person are comprehensible and could easily be recognized. And on the other hand that the video's story was believable and appears true. 55 videos, 27 congruent and 28 incongruent videos, provided the pool of videos for the pilot study.

The randomized presentation and the pilot study video questionnaire were presented by using E-Prime (2012).

2.4.1   Sample.

Twenty healthy volunteers participated in the pilot study (4 males, 16 females). Participants were aged between 21 to 39 years ($M = 26.5$, $SD = 1.16$). The percentage of their highest completed education was 45% AHS-Matura, 5% Berufsbildende mittlere Schule, 10% BHS, 15% Bachelor-Studium, 25% Master/Diplomstudium. Their current jobs were 50% students, 10% self-employed, 35% employee, and 5% other.

Data collection took place in the laboratory 2 of the SCAN-Unit of the Department of Psychology of the University of Vienna.

2.4.2   Measuring instrument and statistical analysis & software.

The applied measuring instruments in the pilot study were a Pilot Study Video-Questionnaire and a semi-structured interview. The ***Pilot Study Video-Questionnaire*** consisted of following questions (to be answered on a 9-point Likert scale or dual choice).

- Kennen Sie die Person? / Do you know this person? (yes/no question)

- Wie attraktiv finden Sie die Person? / How attractive do you find the target-person? (Ickes et al. (1990) demonstrated in a study that partner's physical attractiveness was positively correlated with content accuracy.)

- Wie sympathisch ist Ihnen die Person? / How likeable do you find the person you saw?

- Wie leicht war es, sich in die Person einzufühlen bzw. mitzufühlen? / How easy was it to feel with the target?

- Wie leicht war es, die Gefühle der dargestellten Person zu verstehen bzw. nachzuvollziehen? / How easy was it to understand the target's emotions?

- Waren die Gefühle der dargestellten Person eindeutig oder verwirrend? / Were the target person's emotions clear or confusing? (Because the affect rating dial has only two poles at one dimension, the emotions have to be clear in this dimension.)

The ***semi-structured interview*** was initiated by a very open question if there is anything they would like to mention to the study or the watched videos. Then the questions got a bit more focused and participants were asked if there was maybe something striking or conspicuous (and if yes, what exactly) or if they found the stories of the four videos normal. Further if anything attracted their attention e.g., the people in the videos, the content, etc. Finally they were directly asked how believable they think the described situations were ("Wie glaubwürdig empfandest du die dargestellten Situationen?").

For statistical analysis of the pilot study video-questionnaire Microsoft Excel (2003) was used. The answers of the semi-structured interview built the basis to exclude those videos which were easily recognized as fictitious.

2.4.3 Experimental design.

At the beginning some socio-demographical data of each participant were collected. This included age, sex, highest level of education, current employment and their field of study/working (see page 149) and is supposed to encompass possible confounder.

Then participants were seated in front of a monitor and watched a randomized and counterbalanced (of congruent and incongruent) selection of 20 or 21 videos out of the pool of 55 pre-selected videos. Participants were instructed to feel with the person in the video and after each video they were asked to fill out the pilot study video questionnaire answering questions to the just watched video (presented on the same computer-screen on which the participants watch the videos).

After watching all videos, participants were shortly interviewed to the videos. At the very end participants got informed about the aim of the study and that half of the videos were

incongruent and therefore not true. The clarification should also protect the target-persons against possible damage, like misunderstandings, if somehow a participant meets a target-person by chance in the future.

The pool of possible incongruent and congruent videos for the validation study is presented in Table 24 and Table 23 (pages 162 and 161). Out of them 12 videos, equally incongruent and congruent, were selected. Table 2 presents the final pool of videos for the validation study.

Table 2: Final pool of videos for the validation study.

Video[a]	Valence[b]	Title	Displayed Emotion (formulated by the targets)	Duration (in sec.)
congruent				
Sora 5con	pos	1. Freund	joy, happiness, to be in love with so., floating, agravic	85
Rudi 3con	pos	Kilias Kindsfeier	joy, happiness	82
Lesly 4con	pos	Aufnahmeprüfung FH	joy, pleased; relief	87
Sora 6con	neg	Chefin kritisiert mich vor Kollegin	be treated unfair, attacked, to be betrayed; to be ashamed, embarrassed (in front of the colleague); hurt, offend; shame	90
Michi 1con	neg	verschimmelter Gouda	disgust	82
Michi 3con	neg	Papa trinkt Apfelsaft weg	anger, disappointment, another person's power, to be inferior	86
incongruent				
Rudi 3inc	pos	Kilias Kindsfeier	joy yielded, became sad and thoughtful/musing; be inferior to him ("I don't have anything and we are the same age—that wore me down, somehow")	101
Sora 2inc	pos	Tante werden	be stressed, not joyful ("it was hard for me telling her that I'm pleased"); unfair ("I am 34 years too")	76
Stephan 3inc	neg	Bananenschnitte	tasting good (everyone else would find it disgusting)	92
Rudi 5inc	neg	Ex-Freundin macht Schluß	joy, feeling free; be certain to get together again/to reconvene	85
Michi 1inc	neg	verschimmelter Gouda	tasting good	83
Olcay 5inc	neg	Tot vom Opa	joy	88

[a] The names of the target-persons are those how they presented themselves while videotaped.
[b] Valence: The valence is always from the perceiver's perspective.

2.5 III. Validation Study

2.5.1 Sample.

The sample of the validation study consisted of 63 participants. Two participants had to be excluded (see therefore section 4.3), hence 61 participants constituted the final sample of the validation study. The age varied between 18–32 years ($M = 20.75$, $SD = 3.09$). 46 participants were women (75.4%) and 15 men (24.6%). Highest completed education was 75.4% AHS-Matura/Reifeprüung, 1.6% Lehre/Berufsschule, 11.5% Berufsbildende höhere Schule (BHS), 6.6% Bachelor-Studium, and 4.9% Master/Diplomstudium. And at current occupation most specified to be students (91.8% students, 4.9% employee, 3.3% other).

Each participant was randomly assigned to one of three groups (G1, G2, and G3; this will be described later), hence each group's sample was 21 participants. But because of the exclusion of two participants, one group (G3) consisted only of 19 participants.

2.5.2 Measuring instrument and statistical analysis & software.

Some questionnaires and tasks were applied. They were empathy and self-other distinction related. The used questionnaires (German versions) and a short explanation what they measure will be given.

For the validation study right-handedness and fitting for ECG & EDA were preconditions for participating. Hence participants were informed about the technical equipment (ECG and EDA), the process of the study, exclusion criteria, and possible side-effects of ECG and EDA (like itching). After this they were asked if they have any questions thereto and to sign the

- **Exclusion Criteria for ECG and EDA** (exclusion criteria are listed in Table 1 at page 33; the submitted information sheet is presented at p. 150) and
- the **consumption of stimulants** were checked. Because of their influence on cardio-vascular functions, which further effects ECG (Jennings et al., 1981), participants were asked how much coffee and/or tobacco they consumed before coming to the laboratory and if this amount is consistent to the normal/daily quantity.
- With the **Edinburgh Händigkeits-Inventar** (the German version of *The Edinburgh Inventory*) (Oldfield, 1971) participant's handedness was checked.
- Further some **socio-demographical data** were collected, like age, sex, highest level of education, current employment, and field of study/working to encompass possible confounder (see also p. 149).

- The **Saarbrücker Persönlichkeitsfragebogen (SPF)** (Paulus, 2009) is the into Ger-

man adopted version of the **Interpersonal Reactivity Index (IRI)** developed by Davis (1980). The test consists of 4 sub-scales (subscales' Cronbach's α): Empathie (empathy) (.74), Fantasie (fantasy) (.78), Perspektivenübernahme (perspective taking) (.79), and Bekümmerung (distress) (.76) (Paulus, 2009). According to Paulus (2009) perspective taking refers to the cognitive aspect of empathy, whereas the other three sub-scales correspond to emotional empathy factors. Each sub-scale contains 4 items and is to be answered by the respondent on a five-point scale, indicating how well the item describes the respondent, anchored by 1 (*does not describe me well*) and 5 (*describes me very well*) (Paulus, 2009).

- Another domain in social interaction refers the *primitive emotional contagion* which is defined as "the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person and, consequently, to converge emotionally." (Hatfield, Cacioppo, & Rapson, 1994, p. 5). The **Emotional Contagion (EC) Scale** (Doherty, 1997) is a unidimensional self-report questionnaire measuring the susceptibility to catch other persons' feelings by means of the tendency to automatically mimic and synchronize with the emotions of others. Cronbach's α for the total test is .90 (Doherty, 1997). Although EC is unidimensional, a positive subscale (love and happiness) and a negative (fear, anger, and sadness) subscale can be analyzed too (Cronbach's α = .82 and .80) (Doherty, 1997).

- The term *alexithymia* primarily refers to the disability to perceive and to name appropriately one's own or other people's emotions ("[die] Unfähigkeit, bei sich und anderen Gefühle in angemessener Weise wahrzunehmen und zu benennen") (Starke-Perschke, 2001, p.23). The **Bermond–Vorst Alexithymia Questionnaire Form B (BVAQ-B)** (Müller et al., 2004) is a self-report questionnaire measuring five dimensions of alexithymia: (1) Emotionalizing, (2) Fantasizing, (3) Identifying, (4) Analyzing, and (5) Verbalizing. After Müller et al. (2004) also two higher order factors can be calculated: (I) an *affective factor* consisting of the subscales Emotionalizing and Fantasizing, and (II) a *cognitive factor* as a result of the subscales Identifying, Analyzing, and Verbalizing. The BVAQ-B is one of the two parallel forms of the BVAQ and consists of 20 items. Each item is rated on a 5-point Likert scale and total score ranges from 20 to 100 points "with high scores indicating high proneness to alexithymia" (Müller et al., 2004, p. 376). Deborde et al. (2008) proposed a cutoff score for the BVAQ-B (sum of all items) of people scoring 43 or less are considered nonalexithymic, whereas those scoring 53 or

above are considered alexithymic, and individuals scoring between 44 and 52 are considered intermediate-alexithymic. Cronbach's α for the BVAQ-B is .73 and for its subscales Cronbach's α vary between .39 and .61 (Müller et al., 2004).

- The **Reading the Mind in the Eyes** (Baron-Cohen et al., 2001) is a measure of "mentalising" in adults and is inversely correlated with the Autism Spectrum Quotient.

- **Imitation Inhibition Task** (Brass et al., 2000), in this thesis also denoted as *Finger Lifting Task*. On a motor level, observing an action (*perception*) automatically activates the same motor representations as when internally generated (*acting*) (*shared representations* of perception and action) (Brass et al., 2009). But to understand the fact that we are not imitating all behaviour, needs to take into account mechanisms that allow us to distinguish if the motor representation is self-generated or externally triggered—a self-other distinction on a motor level (Brass, Derrfuss, & Von Cramon, 2005). These mechanisms enable us to control these shared representations and to avoid automatic imitative behaviour (Brass et al., 2009). Therefore the control of automatic imitative response tendencies—measured with the Imitation-Inhibition Task—can be used to index the control of shared representations (Brass et al., 2009). In the Imitation-Inhibition Task participants are instructed to lift their index or middle finger in response to a number (1: index finger; 2: middle finger). In addition to the digit a video-taped hand is presented. In the baseline condition, the hand on the computer screen is unmoved. In the congruent condition, a congruent finger movement accompanies the number (1: index finger; 2: middle finger). And in the incongruent condition, the hand executes an incongruent finger-lifting movement (2: index finger; 1: middle finger) (Brass et al., 2000). The response to the congruent condition does not require to distinguish between their own intended action from the observed one and can be considered as a "quasi-imitative" reaction. Whereas in the incongruent condition the instructed movement differs from the observed one. So participants have to inhibit their imitation tendency and to focus on their own movement intention, which leads to an *interference effect* with a prolonged reaction time (or more errors) in these trials (Brass et al., 2000, 2009). The interference effect can be calculated by subtracting the congruent condition's reaction time (or errors) from the incongruent condition's reaction time (or errors) (Brass et al., 2009) (*interference effect* $= RT_{\mathrm{inc}} - RT_{\mathrm{con}}$) and indicates the self-other distinction in this paradigm.

In this investigation 20 baseline trials (10 index finger; 10 middle finger), 20 congruent trials (10 index finger; 10 middle finger), and 20 incongruent trials (10 index finger; 10

middle finger) were used. The training consits of 6 training-items (from each condition one item). The 1st and 4th frames lasted 1500ms; the 2nd and 3rd frames each lasted 34ms. The instruction was "lift your index or middle finger as fast as possible".[4]

- In the **Perspective-Taking Task** (Keysar et al., 2000; Santiesteban et al., 2012), also named as *Shelve Task*, participants see on a computer screen a shelf and a "director" who stands behind this shelf and has therefore an opposite view on it. This director "instructs" the participant to move an object in this shelf. In experimental trials, the participant's perspective differs from that of the director, because some of the grids' back sides are closed and the object inside is hidden for the director's perspective. In this "conflict situation" the participant needs to enhance the other's perspective and to inhibit the own perspective—this requires a cognitive self-other distinction. The task includes also two control conditions to ensure that the results can not be explained by other factors, like stress or other cognitive processes. The instruction is to "be as fast and precise as possible" and reaction time and error rate are used to index self-other distinction.

- **Post-Video Questionnaire**. After having seen all videos, to each video six questions had to be answered on a 9-point Likert scale. To help the participant to remember the target-person and the content of the event, the first few seconds of each video were replayed.

  - Finden Sie die dargestellte Person interessant? / Do you find the person in the video interesting? [5]
  - Wie attraktiv finden Sie die dargestellte Person? / How attractive do you find the person in the video? [6]
  - Wie sympathisch finden Sie die dargestellte Person? / How likeable do you find the person?
  - Hatten Sie das Bedürfnis sich von der erzählten Situation distanzieren zu müssen oder zu wollen? / Did you have the want to distance from the describes situation? [7]
  - Hatten Sie das Gefühl nicht "bei der Sache" gewesen zu sein / abgeschweift zu sein? / Did you feel not to have been "in the matter" / to have zoned out?

---

[4]Analysis included only data of correct trials. According to Rauchbauer, Majdandžić, Hummer, Windischberger, and Lamm (2015) a winsorising procedure on the mean response times (RT; in ms) of every participant was performed to account for potential outliers in the behavioural data (Wilcox, 2011).

[5]Ickes et al. (1990) showed that content accuracy was influenced by perceiver's interest in the partner.

[6]As well partner's physical attractiveness is positive correlated with content accuracy (Ickes et al., 1990).

[7]This questions was asked to estimate perceiver's defense.

– Waren Sie in Ihrem Leben schon einmal in einer ähnlichen Situation? / Have you been once in your life in a similar situation? (At this question only four answers were listed ("nein, gar nicht", "annähernd ähnlich", "ähnlich", and "ja, genau gleich").)

- **Rating dial practice-tasks.** To become acquainted with the handling of the rating dial and the instructions, four rating dial practice tasks were undertaken. At first participants were instructed verbally about the handling of the rating dial. Then to get used to it, they exercised the handling by doing the practice tasks. Three tasks consisted of rating the emotional valence of another person (other-rating). One task focused on the own emotional valence (self-rating). It was possible to repeat each task as often as requested.

1. Other-rating (words). Different words of emotional valence (negative–positive) were selected to provide a first acquaintance with the handling of the affect rating dial. Following words were used (only the German expressions were used in this study): (1) unglaublich traurig, zum heulen / incredibly sad, it makes me want to cry; (2) traurig / sad; (3) eher traurig / rather sad; (4) verstimmt / displeased; (5) neutral, egal, unberührt / neutral, unaffected; (6) angenehm gestimmt / enjoyable tempered; (7) gut drauf / to be in a good mood; (8) freudig, froh / glad; (9) unglaublich glücklich / incredibly happy. The instruction was "Es werden jetzt am Bildschirm unterschiedliche Wörter für je 4 Sekunden präsentiert. Bitte versuche mittels Rating-Dials diese einzustufen. Es gibt hierbei kein 'richtig' oder 'falsch'!" ("Different words will be presented on the screen. Each lasts for 4 seconds. Please try to rate them by using the rating dial. There is no 'right' or 'wrong'!"). The order of the words was more fluent without abrupt changes (the exact order of the words was: 6-4-3-2-3-2-1-2-4-5-6-5-6-7-8-9-7) and between every word a fixation-cross was presented for one second. Because this task was the first time they handled the rating dial, it was explicitly asked if the position of the rating dial was comfortable and handy so that the handling was easygoing and possible without looking at the rating dial too often.

2. Other-rating (photos). Photos of faces with different emotional expressions ranging from very happy to very sad (taken from the data bank of the SCAN-Unit) were presented. Each picture lasted 4 seconds and the order of the different emotions displayed by the faces was, as in the task before, fluent and without abrupt changes. Participants were asked to rate the valence of the emotions of the faces (other-rating) by using the affect rating dial.

3. Self-rating (vignettes). While the first two tasks were to practice the other-rating, in this task the self-rating was exercised. Therefor participants were instructed to put themselves in the situation but to sense the own feelings and to rate the own emotional valence they would have in this situation by using the rating dial. Because the rating dial has only one dimension—the valence of the emotion—participants were briefed that all displeasing emotions are localized at the "negative half" of the scale of the rating dial and all pleasant emotions on the "positive half". The exact instruction was following:

> Je nach Instruktion, geht es aber auch darum, daß Du Dich zwar in die Situation eindenkst und einfühlst, aber dabei gleichzeitig ganz bei Dir bleibst. Daß Du also Deine Gefühle, die Du in dieser Situation hättest, mittels Rating-Dial einstufst. Mit dem Rating-Dial kann man allerdings nur die Valenz der Gefühle, also "positiv" bis "negativ" einstellen. Alle für Dich unangenehmen Gefühle sind also auf der "negativen Hälfte" und alle Dir angenehmen Gefühle auf der "positiven Hälfte" lokalisiert. Es gibt hierbei kein "richtig" oder "falsch"! Es geht nur darum, daß Du einstufst, wie Du Dich fühlen würdest.

Nine different situations were presented. The first six vignettes described individual situations with different degrees of involvements like "Imagine you would win in the lotteries. What feeling would you have?", "Imagine the neighbors would tell you, as you met them by chance, that they are getting married. What feeling would you have?", or "Your beloved dog would die. What feeling would you have?" One vignette was incongruent: "Your best friend told you that he got 5000,- Euros but he wasn't happy about that. How would you (you, with your own personal biography and personality) feel in this situation?" And the last two vignettes had a sudden change in their valence. For example "Imagine, your favorite pasta is a special offer today: two packages at the price of one! How would you feel? And now imagine that you forgot the package in the supermarket after buying it. How would you feel now?"

4. Other-rating (video). In the last rating dial practice-task participants watched one congruent video which was developed during this study but not selected for the validation study. The instruction was exactly the same as the instruction for perceiver's other-rating as written above.

As mentioned each practice-task could be repeated as often as requested. I noticed that all four rating dial practice-tasks were important and necessary to get participants acquainted to both, the challenging instructions and the handling of the rating dial.

Statistical analyses. To make emotions visible and to measure them an affect rating dial was used (described in section 2.1, p. 31). The computer program MATLAB (2015) was used to list the raw data and the normalized data of the rating dial of each second of each video from each participant. Normalized data were in regard of the rating dial's scale, so that the two poles (*negative* and *positive*) were represented by −1 and +1. This lists were imported into R-Studio (R Development Core Team, 2015) to plot the rating dial data as graphics. Finally SPSS (2010) was used to calculate one-way repeated-measures ANOVAs, two-way repeated-measures ANOVAs, and correlations (all correlations are two-tailed).

2.5.3 Experimental design.

Allowing for the fact that concentration is limited it was considered that a duration of about 15 minutes of concentrated watching, feeling, and affect rating is reasonable. Each video had to be watched twice and each video lasted about 90 seconds. Hence watching and rating four videos took about 12 minutes. Further some targets were represented more than once in the selection of videos. Videos were arranged into three groups (G1, G2, G3), each group consisted of four videos. Due to this arrangement it was possible that each target was watched by a participant only once. The arrangement into the three groups (G1, G2, G3) of the selected videos is seen in Table 3.

Table 3: Validation Study: Videos' Arrangement to Groups

Group	Name of Video	con/inc	Valence (for perceiver)	Duration (in sec.)
Group 1	Lesly 4con	congruent	positive	87
	Sora 6con	congruent	negative	90
	Rudi 3inc	incongruent	positive	101
	Michi 1inc	incongruent	negative	83
Group 2	Rudi 3con	congruent	positive	82
	Michi 1con	congruent	negative	82
	Sora 2 inc	incongruent	positive	76
	Olcay 5inc	incongruent	negative	88
Group 3	Sora 5con	congruent	positive	85
	Michi 3con	congruent	negative	86
	Rudi 5inc	incongruent	negative	85
	Stephan 3inc	incongruent	negative[*]	92

[*] Due to an error in reasoning group 3 contains three negative videos, instead of two negative and two positive ones.

47

In each group two videos were incongruent and the other two congruent. And it was intended to counterbalance the videos, so that one of the congruent and one of the incongruent videos had a positive valence and the others a negative. But due to an error in reasoning in group 3 three of the four videos have negative valence. The valence in incongruent videos (as well as in congruent videos) is always out of the perspective of the perceiver. So irrespective of how the target's feelings were directed to that specific event, the *valence* describes what kind of emotional valence is "normally" expected in this situation (the content of each video will be described in section 2.5.4).

Participants were randomly assigned to either one of these groups (G1, G2, or G3). Videos within groups were also assigned in random order.

The experimental design of the validation study consists of two parts. In the first part participants were asked to fill out some questionnaires via Internet in advance (all questionnaires are described in section 2.5.2). This shortened the time that participants had to spend in the laboratory at the second part, because the session in the laboratory lasted about 1 to 1.5 hours. In advance participants were asked to fill out questions concerning their sociodemographic data and to read the exclusion criteria for the application of an electrocardiogram (ECG) and electro-dermal activity (EDA) (which they will be asked to sign at the second session at the laboratory). Further following questionnaires were applied: Bermond-Vorst Alexithymia Questionnaire (version B) (BVAQ-B), Saarbrückener Persönlichkeitsfragebogen (SPF) (that is the German version of the IRI), Emotional Contagion Scale (ECS), and Reading the Mind in the Eyes Task.

The second part took place at the laboratory (lab 2 of the SCAN-Unit of the Department for Psychology of the University of Vienna). At first participants were asked to sign the exclusion criteria for ECG and EDA and it was checked if the questionnaires of part one were filled out completely. The Edinburgh Händigkeits-Inventar was applied to check participants right-handedness. Then participants were asked to leave all electronic devices (mobile phone, laptop, mp3-player, etc.) as well as other possible interference factors (like keys, piercings, coins, etc.) outside and to go into a little room with a Faraday shield for the conduction of the physiological parameter. At first the electrodes for the ECG and EDA application were attached. The attachment of the electrodes was absolutely equal to that of the target-person's. At first the skin of the area where the electrodes were placed was cleaned with alcohol (70%). The ground electrode was attached at the wrist of the non-dominant hand (all participants were right-handed, hence the non-dominant hand was always the left hand). The electrodes for EDA-application were placed at the middle phalanx of the inner face of the index and ring-finger

at the non-dominant (left) hand. One electrode for the ECG-application was attached on the right lateral clavicle and the other at the left rib. Although the ECG- and EDA-application was started later, participants got used to the sensation of the electrodes. Further because of the ECG and EDA application participants were asked to absolutely not move their left hand and body during the videos. Hence (automatic) movements of the left hand and the body could be observed and (if necessary) corrected while participants were asked to do some rating dial practice-tasks. Four different practice-tasks were applied to help participants on the one hand to get used to the handling of the affect rating dial and on the other hand to become acquainted with the instructions for other-rating and self-rating (i.e., to feel with another person and to rate their emotional valence or to sense the own emotions and to rate their valence) (the practice-tasks are described more precisely at p. 45). Attention was paid to a comfortable position of the rating dial for participants.

Before the experiment started, both instructions (other-rating and self-rating) were explained verbally to the participants. To check if the meaning of the instructions was captured, participants were asked to shortly reproduce the instructions in their own words.

After the practice tasks and the affirmation of understanding the instructions, ECG and EDA conductions were initiated and the actual experiment started. At the beginning a rating dial calibration was undertaken. Therefor participants were instructed to point the index of the rating dial for 3 seconds to *very negative* (the minimum), *very positive* (the maximum), and *neutral* (the middle) of the scale. On the one hand the calibration helps (with data analyzing) to check how exactly a participant points to the middle of the scale for indexing *neutral*. And on the other hand it servers to inspect the rating dial's extension (and possible changes) and to calculate z-scores for each participant individually.

As written above, participants were randomly assigned to one group (G1, G2, or G3). According to Table 3 (p. 47), four videos were presented in random order. Each video was consecutively presented twice. Instructions were always in the same order. First participants were instructed to empathize with the displayed target-person and to rate with the affect rating dial the emotional valence the target-person had while describing the story (perceivers' other-ratings). In the second run participants were instructed to put themselves into the described situation and to rate their own emotional valence which they would have (as the person they are with their own biography and personality) in that situation (perceivers' self-ratings). Before the video started the verbatim instructions were presented on screen and were as followed:

For perceiver's other-rating:

> Bitte versuche Dich auf die Person, die gleich zu sehen ist, einzulassen. Versuche Dich in die Person und ihre Emotionen einzufühlen und stufe diese mittels des Rating-Dials ein. Bitte stelle das Rating-Dial auf "neutral" und drücke die "Leertaste", wenn Du bereit bist.

For perceiver's self-rating:

> Du siehst jetzt das gleiche Video nocheinmal. Aber diesmal versetze Dich bitte in die dargestellte Situation. Wie würdest Du selbst (Du mit Deiner ganz persönlichen Biographie und Persönlichkeit) dich in dieser Situation fühlen? Stufe bitte mittels Rating-Dial das Gefühl ein, welches Du in dieser Situation hättest. Bitte stelle das Rating-Dial auf "neutral" und drücke die "Leertaste", wenn Du bereit bist.

The exact process for each video of a group (G1, G2, or G3) was:

– 30 seconds fixation cross (which serves as resting ECG),
– instruction for perceiver's other-rating,
– one of the four videos,
– 30 seconds fixation cross,
– instruction for perceiver's self-rating,
– the same video as before.

This process was repeated in total four times, one time for each of the four videos per group (G1, G2, or G3).

After the last of the four videos had been watched and rated, ECG and EDA electrodes were removed. Immediately afterwards participants were asked to fill out a Post-Video Questionnaire concerning different aspects of each video. The first three questions concerned the target-person, if they perceived him/her as interesting, attractive, and likeable. The next two questions asked about participant's wish to distance from the described situation and if they zoned out. In the last question participants were asked if they had been in a similar situation once in their life. For the exact questions see the Post-Video Questionnaire at p. 152.

When finished the Post-Video Questionnaire, participants were asked to do two more tasks, the Finger-Lifting Task (Imitation Inhibition Task) and the Shelve Task (Perspective Taking Task).

At the end participants were asked if they found all videos common or if there was anything strange? Or if they want to comment anything? With these questions I tried to find out if

participants believed all videos. And to check further if the plausibility influenced empathic accuracy or self-other distinction.

Before farewell the participants, they were informed about the aim of this study and possible questions were answered. Up to this point of time when participants asked if the stories were true, I always answered that they are autobiographical life-events. This is a true fact, just what I didn't mention was that half of them were incongruent and not experienced that way. Therefore now it was clarified that half of the videos they had watched weren't true. This is also thought as a protection for the target-persons. In case of participants meeting this target-person by chance, no misunderstanding will emerge and the target-persons do not suffer any damage.

The collected data for each condition (congruent and incongruent) are on the one hand physiological data of the perceivers (which will be analyzed in another master-thesis) and on the other hand the affect-data (by using the affect rating dial) while rating first target's emotional valence (perceivers' other-ratings) and in a second time perceivers' own emotional valence (perceivers' self-ratings). With these data the video's validity to examine empathic accuracy and self-other distinction will be calculated.

2.5.4 Content of the selected videos for the validation study.

Each Video consists of a self-introduction of the target-person and then three periods of interest: *Period 1:* the story-introduction, *Period 2:* the salient event & emotional period, and *Period 3:* the reason & cognitive period (for detailed information to the structure of the videos see p. 17). Each of the selected videos will be shortly characterized. For an overview of the selected videos see Table 3 (p. 47).

As described in section 2.5.3, there are three groups of participants in the validation study. Each video was assigned to one group and for identifying easily to which group (G) it belongs, each video got a code (G1, G2, and G3) (see also Table 3 at page 47). The names of the target-persons are those how they presented themselves while videotaped.

G1 Lesly 4con – Entrance Exam / Aufnahmeprüfung FH (congruent, positive): In the self-introduction Lesly tells that she is 24 years old, her family is from India and she has a big family with one older sister, one younger brother, and one younger sister. Differently to her siblings she is more interested in the technical context and currently works in the building industry. *Period 1:* The story of Lesly deals that she started studying construction engineering at the Technical University Vienna (TU). She tells that she had only little success there, maybe

she was not that motivated or however it was difficult for her. She was already quite down because it didn't proceed much and she had big doubts about herself. After all she thought that maybe an advanced technical college would be the better option for her. So after two years studying at the TU she applied at an advanced technical college and felt like dangling on a string, because there was an entrance examination. *Period 2:* The emotional peak consists in the feelings of being very, very happy as she passed it and got accepted. *Period 3:* The reason therefor is that she had a glimmer of hope concerning her future that she could go on now and was not standing in front of a wall anymore. She was very, very glad about the positive result and the admission.

G1 Sora 6con – Boss Criticizes Me in Front of a Colleague / Chefin kritisiert mich vor Kollegin (congruent, negative): In the self-introduction Sora tells that she is 34 years old, born and living in Vienna. She has a younger brother. Her mother comes from Korea, her father from Vienna, but because born and grown up in Vienna she feels being Austrian. She sees herself as a competent woman, especially in her job. *Period 1:* In the part of the story-introduction Sora describes that there was once a situation where her boss called her into her office and criticized her very much in front of a colleague. Furthermore the boss didn't give her a chance to justify herself or to explain her point of view. *Period 2:* In the emotional period Sora points out that this event was very embarrassing for her. *Period 3:* And in the cognitive period Sora underlines that it was embarrassing especially because Sora had a good relation to her boss before and was open to her; so the boss knew about Sora things that she normally would not talk about at work. Afterwards Sora wondered why she was that wrong concerning the relation to her boss and it was very, very displeasing for her.

G1 Rudi 3inc – Celebration of a Newborn / Kilias Kindsfeier (incongruent, positive): The video starts with Rudis self-introduction telling that he is 30 years old and from Upper Austria. He has an older brother and sister. Rudi is quite a family-guy, so he tries to drive home every weekend and to look if there is something to help. He is very connected with both siblings. *Period 1:* In the period of the story-introduction Rudi narrates the event when his best friend became a father. At first he was very joyful for him and immediately took a train to see him. *Period 2:* The salient event deals of the sudden change of Rudi's joy into something like sadness and being thoughtful. *Period 3:* The reason therefore is that Rudi and his best friend are the same age, grew up together and he has already a house, a wife, a child, is working already for some years, and Rudi only just finished his studies. His best friend has

already everything and Rudi has nothing, but they are the same age. Although Rudi hasn't endeavoured it, it wore him down.

G1 Michi 1inc – Mouldy Gouda / Verschimmelter Gouda (incongruent, negative): Introducing himself Michi tells that he is 30 years old, works in Vienna as a software developer and lives together with his girlfriend since seven years. To initiate the story Michi tells that while working as a software developer it happens sometimes that he comes home late and then he just goes to the fridge and takes what he catches first. *Period 1:* The story-introduction deals of an event that happened last week. He gripped a package of gouda (cheese), sat down in front of the the PC, didn't look and just took a bite of the cheese. It tasted incredibly tart comparing to usual. And he thought tart but tasty. The introduction into the salient event consists of the narration as he looked the first time at the cheese he saw that on one side the whole cheese was covered by green mould. *Period 2:* The period of the salient event deals with Michi thinking that therefore (for the mould) the cheese tastes so good and that he ate the whole rest of the cheese. He found it really tasty. *Period 3:* The reason is that he always wanted to bite into a piece of mould.

G2 Rudi 3con – Celebration of a Newborn / Kilias Kindsfeier (congruent, positive): In the self-introduction Rudi tells that he is 30 years old and from Upper Austria. He has an older brother and sister to whom he is very close. He drives home nearly every weekend to help them. He just finished his studies and shifted his center of life to Vienna. The change into the story is about Rudi still feels very connected to Upper Austria and all his best friends are from there. *Period 1:* The story-introduction of Rudi is about when his best friend became a father. He was the first of his clique who became a father and it was a big ballyhoo. Everyone drove to him the day after the child's birth. They mounted a stork and took the father and got him drunk, celebrating and eating. *Period 2:* In the emotional period Rudi underlines that it just was such a joy. They were only the guys like they grew up and the event was such a pleasure, just brilliant. *Period 3:* In the period of the reason Rudi explains that the father was very happy and transmitted this to him and the friends and it was a huge event. The first child in the circle of friends and everyone was there, all together. It just was easygoing.

G2 Michi 1con – Mouldy Gouda / Verschimmelter Gouda (congruent, negative):
The video starts with Michis self-introduction that he is 30 years old, working as a software developer in Vienna, and he lives together with his girlfriend since about three years. The con-

nection to the story consists that Michi tells that as a software developer it happens often that he comes home late, starving a lot, and just takes the next he grips out of the fridge. *Period 1:* The story of Michi is that last week a disagreeable event happened. He took a package of Gouda out of the fridge, he opened it, and took a bite. *Period 2:* The salient event is that the cheese tasted more tart than usual. But he took some more bites. At the other end it was a bit better. *Period 3:* In the period of the reason Michi explains that when he first took a look at the cheese he saw that on one side it was fully covered by mould. Immediately he spat out what he had in his mouth. He didn't like it and it was quite disgusting. But he had eaten most of the cheese already. He thought he will tolerate it but won't do it anymore. He couldn't understand it because the cheese was only a week old. But it was not pleasant.

G2 Sora 2inc – Become an Aunt / Tante werden (incongruent, positive): To introduce herself Sora tells that she is 34 years old, lives and is born in Vienna, has a younger brother, her mother is from South Korea and her father from Vienna. To initiate the story she discusses that her mother has a friend in Vienna and her daughter belongs quasi to the family. *Period 1:* The story-introduction deals of the life-event that Sora and this daughter phoned last week. Although they like to speak to each other, it was not a pleasure for Sora to hear the actual positive event that her friend becomes a mother with 38 years. Of course the friend was thrilled. *Period 2:* The salient event starts when Sora tells that during the call she couldn't be happy for her. It stressed Sora. Although it is nice for her friend, (*Period 3:*) Sora explains that for herself this is not amazing because she is also 34 years old and finds it kind of unfair that she becomes a mother and not Sora herself. Anyhow she had to tell her that she is happy for her but this was quite hard for Sora.

G2 Olcay 5inc – Death of Grandfather / Tod vom Opa (incongruent, negative): In the self-introduction Olcay tells that she is 26 years old and is from Vorarlberg. She has 4 older sisters, lives in Vienna since 2006, studied law and currently she is at the Vienna University of Economics and does the doctorate. Olcay works in a office and also honorary with children—she teaches mathematics to asylum seekers, she sees the children twice a week. *Period 1:* Olcay's story is about that two years ago her grandfather died. With him she undertook a lot, he taught Olcay how to ride a bicycle, how to grave, how to play chess and she loved him very much,—*Period 2:* the salient event—but nevertheless she was very happy. At the moment as she heard about his death she was happy although she should have been sad, but it was liberating, totally alright and she didn't shed a tear. *Period 3:* The reason for Olcay's feelings

is that Olcay was mentioned in the testament and inherited 10.000 Euro. And the amazing thing is that with 26 years as a student you can't afford a journey. And with 10.000 Euro she thought that is great, immediately she could do a small journey round the world. And who doesn't like that. Her friends can't do that and she can do a big journey because she quasi won in the lotto.

G3 Sora 5con – 1st Boyfriend / 1. Freund (congruent, positive): For introducing herself Sora tells that she is 34 years old, lives and was born in Vienna. She has a younger brother. Sora's mother is from Korea and her father is from Vienna, therefore she is half-Asian but educated in Austrian style and feels completely as Austrian. As a connection to the story Sora tells she still isn't married. *Period 1:* Her story: She can remember very good her first boyfriend. It was at the age of 18 which is very late for her notion. Sora remembers that good the details, after they came together she went by tram, it was spring, the sun was shining and—*Period 2:* the emotional period starts with—she was so enamoured, just like how someone would imagine it. She was so happy, sitting in the tram with a smile next to all the grumpy Viennese and grinned at all. Sora was just doing great and it was amazing fine, agravic, and carefree. *Period 3:* To assign a reason Sora explains that finally she got the first boyfriend, after all her friends had already one. It was the greatest for Sora and she thought from now on it will go on only positively.

G3 Michi 3con – Dad Drank Away my Apple-Juice / Papa trinkt Apfelsaft weg (congruent, negative): Michi introduces himself as being 30 years old and living in Vienna where he also was born. In the past, in Michi's childhood, his parents, brother, and himself drove to Burgenland from time to time. The parents had a house there and from there they undertook trips into the surroundings. *Period 1:* Michi's story is about one trip to Güssing. There they viewed the castle. It were only his father and him and Michi was about 4 to 5 years old. After the excursion they were very thirsty and sat in a café, everyone with a glass of apple juice. Michi was very sparing with his juice while his father has finished it very quickly. After a short time his father took Michi's half-full glass and finished it as well. *Period 2:* In the emotional period Michi describes that he was of course totally outraged and said, "This is mine. Why do you drink it away?" And his father answered, "I have paid for it." This answer—Michi was not only disappointed but also very, very angry. And furthermore somehow hurt. *Period 3:* The reason therefor is that his father simple utilizes his authority and took the drink away from him. Because Michi saw it like a gift—there is a beverage and this is yours. Michi felt this was his own and his father simply took it away. He gave it and took it away again.

G3 Rudi 5inc – Girlfriend Dumped Me / Freundin macht Schluß (incongruent, negative): As self-introduction Rudi tells that he is 30 years old and is from Upper Austria. He has two older siblings, a brother and a sister and Rudi tries to drive home as often as possible to spend time with the family. He is quite a family-guy and enjoys being in Upper Austria. *Period 1:* Rudis story is about that in the past he was even more in Upper Austria because his first girlfriend was from there and they were together for many years. She was the love of his life. Rudi could imagine everything with her—building a house, getting children. And she dropped him from one day to the other. *Period 2:* In the salient event & emotional period Rudi describes that even while she broke up, Rudi was doing well, he thought, "yes, alright." Actually Rudi oriented himself towards her like finishing the studies quickly, getting children, and building a house and she dumped him without an actual reason. And Rudi didn't matter; not only that he didn't care he thought, "yes, great! Totally cool!" *Period 3:* The reason therefore is that for Rudi it was obvious that at the end they will be a couple again. In half a year they will reconvene and then they will have the best make-up sex ever. And in the meanwhile, the half year, he is in Vienna, can let himself go, go out, doesn't have to justify himself, enjoy life and then they will come back together.

G3 Stephan 3inc – Banana Cake / Bananenschnitte (incongruent, negative): To introduce himself Stephan tells that he is 27 years old, from Burgenland, and works as a construction engineer what means that he designs and plans houses. His whole family is from Burgenland, from the little town called Neufeld at Neufelder-lake. Three month ago Stephan had his birthday. *Period 1:* The story of Stephan is about that the most beautiful thing at his birthday is that he gets always a banana cake made by his mother. He looks forward to this most and it's the best birthday gift she can give him. Because this is a very special banana cake. His mother always does 3–4 cm sponge, then apricot jam, banana slices, mascarpone, (*Period 2:*) between the bananas mayonnaise, and on the top chocolate. And it is so special because beside the mayonnaise, on the top above the chocolate there are slices of pickled gherkin and to Stephan it tastes especially good, there are some spots of ketchup and ketchup to dip it in—this is a very special flavour-experience, this sweet-sour-feeling. *Period 3:* The reason therefor is that it always remembers him of his childhood. Stephan always got it that way, and it is absolutely unique and tastes extremely good, this sweet-sour-feeling in the mouth is outstanding.

3   Results

3.1   I. Development of Stimuli

To control for target's expressivity, their emotional expressivity is measured by using the Berkeley Expressivity Questionnaire (BEQ) (Gross & John, 1997) and the Fragebogen der Facetten emotionaler Expressivität (FFEE) (Tausch, 2006). Results in percentageages of the BEQ are presented in Table 4 (for the scores see Table 21, p. 153). The result of all subscales of the FFEE are listed in Table 22 (p. 154) and results of selected scores which were relevant for this study are presented in Table 5. In Table 4 and Table 5 percentages are presented. These percentages are based on the maximal score of a subscale and hence represent the achieved percentage of the maximum of a subscale's score. Table 4 revealed that all participants' score at BEQ's scale general expressivity was at least 50%. At the FFEE's subscale expressivity all participants scored between 56% and 88% (see Table 5). The ability to play were at two target-persons below 50%, the other targets scored above 50%.

Table 4: Target's Expressivity Measured by the BEQ.

Target	BEQ			
	Negative Expressivity	Positive Expressivity	Impulse Strength	General Expressivity
Daniel	53%	88%	56%	63%
Michi	36%	71%	50%	50%
Sora	64%	88%	47%	64%
Stephan	58%	83%	56%	64%
Anna	47%	83%	75%	67%
Olcay	50%	46%	72%	57%
Lesly	36%	67%	53%	50%
Rudolf	61%	75%	56%	63%
Konstanze	33%	88%	53%	54%
Matthias	56%	75%	64%	64%
Rudi	31%	71%	67%	54%

Note. Percentage are rounded.

In total 101 videos, equally congruent and incongruent, were developed. Out of these I selected 55 videos, 28 incongruent and 27 congruent videos for the pilot study. The selection was in respect of the story's plausibility or rather if it was not obviously fictitious and further if the target's emotional expression was proper.

Table 5: Selected Subscales from the FFEE as a Measure for Target's Expressivity.

Target	FFEE				
	Expr.[1]	Dar.[2]	Ver.[3]	Gef.[4]	Erl.[5]
Daniel	65%	52%	60%	81%	52%
Michi	65%	81%	67%	100%	23%
Sora	73%	52%	80%	100%	5%
Stephan	69%	76%	47%	76%	19%
Anna	88%	71%	87%	81%	14%
Olcay	72%	33%	53%	43%	29%
Lesly	56%	57%	40%	85%	25%
Rudolf	67%	52%	60%	81%	29%
Konstanze	61%	67%	47%	62%	33%
Matthias	69%	39%	67%	95%	23%
Rudi	63%	67%	67%	62%	38%

[1] Expr. = Expressivität / expressivity
[2] Dar. = Darstellungsfähigkeit / ability to play
[3] Ver. = Verstellungsfähigkeit / ability to dissimulate
[4] Gef. = Gefühlsklarheit / emotional clarity
[5] Erl. = Erlebte Unfähgkeit, Gefühle zu vermitteln / experienced inability to communicate emotions
Note. Percentage are rounded.

3.2   II. Pilot Study

The pilot study served primarily to select those videos which were believable, were good to empathize, understandable, and did not distract the perceiver's attention through something conspicuous. Therefor participants of the pilot study were asked to fill out a Pilot Study Video Questionnaire (see p. 38) for each video they watched. Additionally they were interviewed about the videos they saw.

Especially the interview helped to detect those videos or targets which were not convenient for further use. In the interview it turned out that two target persons (and all their videos) had to be excluded because of the distracting effect of their t-shirt imprints. Another target person evoked very different feelings in participants—half of the participants just "loved" the target, the other half couldn't stand her at all. The videos of that target-person were excluded as well. Further the videos of two more target-persons were excluded because their mimic and/or gesture was found to be inappropriate. From the remaining target-persons, all videos which were found to be not true were excluded, albeit if their narration was true or fictitious.

17 congruent and 12 incongruent videos remained after this first selection. Further the

results of the Pilot Study Video Questionnaire were analyzed. The most important items of this questionnaire were *(4.) how easy was it to understand the target's emotion?* and *(5.) how easy was it to feel with the target?*, because they ask for two fundamental parts of empathy—the cognitive understanding and on the emotional level if it is possible to feel with. From each video the scores of these two items were added and according to the sum of this score they were listed. Table 23 (p. 161) presents this list of congruent videos, and Table 24 (p. 162) the list of incongruent videos. Out of this pool videos were selected primarily on the basis of the sum of the two questions above, but also on how well the story was structured.

12 videos, 6 incongruent and 6 congruent, were chosen for the validation study.

Some citations of participants of the pilot study are listed in Table 25 (p. 163) (only to the 12 selected videos). To sum it up it was said that participants believed the stories. But some outstanding statements were pointed out. These will be discussed in section 4.2.

3.3   III. Validation Study

Each video contains three periods of interest (explained at p. 17). In period 1 the target begins to tell the story (story-introduction). In period 2 of congruent and incongruent videos a salient event happens and the target presents the story on an emotional level (salient event & emotional period). The last period, period 3 (reason & cognitive period), starts when the target-person tells the reason. All periods were counted from the onset-point on. *Onset-point* signifies the second when the message receives the participant and is 2–3 seconds after the time-point when the target started describing the specific content of this period. This delay reflects the time perceivers need to catch the message.

First data were explored graphically. For each video target's self-rating as well as perceiver's other-rating and perceiver's self-rating were illustrated graphically. Figures at pages 167–229 display the other-ratings and self-ratings of each single participant (also the graphs of the two excluded participants, TP311 and TP330). Figure 27 to Figure 38 (p. 155–160) present target's self-rating embedded in perceivers' other-ratings and perceivers' self-ratings. Graphs revealed that targets' self-ratings, especially at incongruent videos, were not working as expected (discussed at section 4.3.5, p. 101). Because this way of measuring empathic accuracy didn't work, empathic accuracy was further defined as "EA" (see p. 25).

Self-other distinction was operationalized as the difference between perceiver's other-rating and their self-rating (see page 22): $|SOD| = |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{perceiver's self-rating}})|$. And empathic accuracy as $|\text{"}EA\text{"}| = |\mu_{\text{perceiver's other-rating}} - (\mu_{\text{mean of all perceivers' other-rating}})|$ (see p. 25). For each video at each time-period of interest one-way repeated-measures ANOVAs

59

were run to calculate SOD as the differences of perceivers' other-ratings compared to their self-ratings (as the measure for self-other distinction). And two-way repeated-measures ANOVAs were conducted to analyze interaction effects of Time x Other/Self. Pairwise comparisons were Bonferroni corrected. Finally correlations (two-tailed) of SOD and "EA" with the applied questionnaires and tasks were calculated.

Data analysis was based on the recommendations of Field (2009). The classification of effect sizes were used according to Cohen (1992).

The sample for group G1 and G2 consisted of 21 participants. Because of the exclusion of two participants (see therefore 4.3), only 19 participants constituted the sample of G3.

3.3.1 G1 Lesly 4con – Entrance Exam / Aufnahmeprüfung FH (congruent, positive)

As shown in Figure 3 the three periods of interest are seconds 33–67 (1. story-introduction), 69–70 (2. salient event & emotional period), and 73–87 (3. reason & cognitive period).
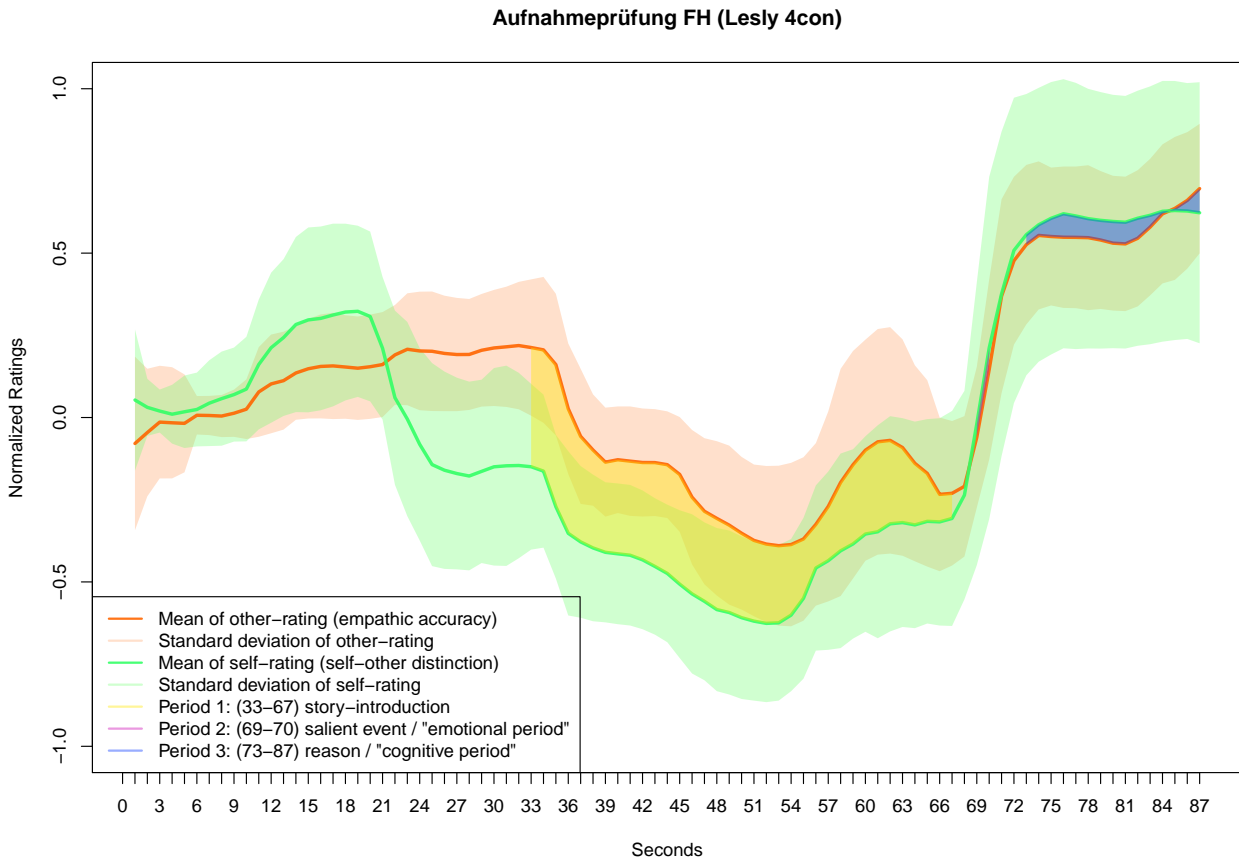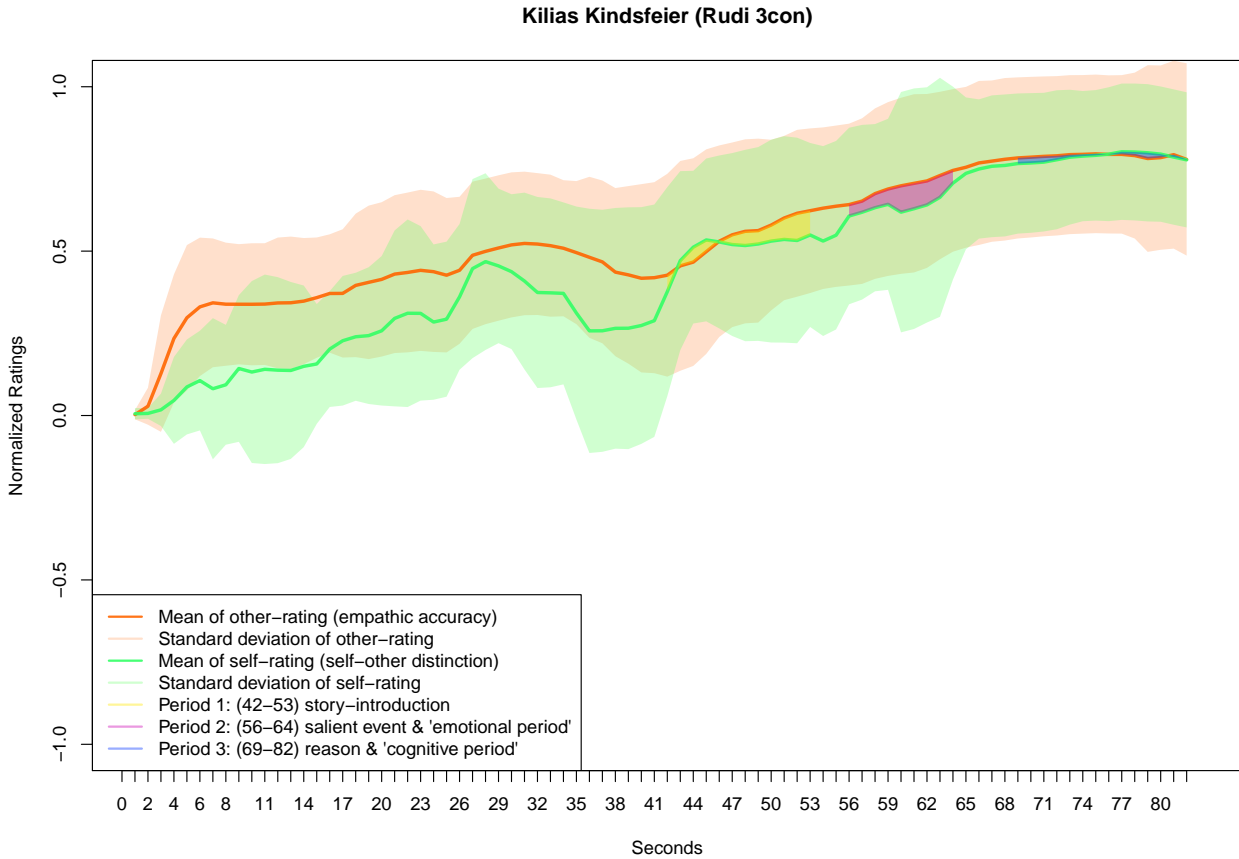

Figure 3: Perceivers' affect-ratings at the video G1 Lesly 4con (Entrance Exam / Aufnahmeprüfung FH).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA indicated that perceivers' other-ratings differed significantly from their self-ratings in the first period, $F(1, 20) = 39.64, p < .01, r = .82$, but not in the second period, $F(1, 20) = .27, p \geq .05, r = .12$, and third period $F(1, 20) = .23, p \geq .05, r = .11$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at

each period are displayed in Table 6.

Table 6: G1 Lesly 4con: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	−0.17 (0.16)	−0.43 (0.21)
period 2	0.04 (0.23)	0.10 (0.47)
period 3	0.57 (0.20)	0.61 (0.39)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

To reveal interaction effects of Time x Other/Self two-way repeated-measures ANOVAs were conducted.
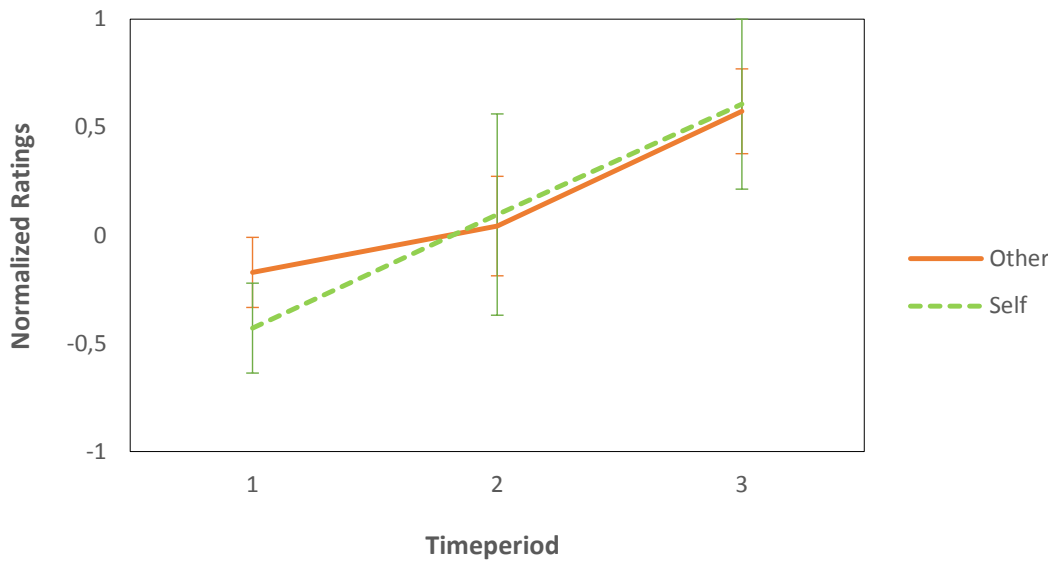
Figure 4: Interaction graph of G1 Lesly 4 con. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the interaction effect of Time x Other/Self, $\chi^2(2) = .04, p < .05$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .78$).

The significant main effect of time, $F(2, 40) = 93.93, p < .05$, and Figure 4 indicated that the emotional valence was rated more positive with increasing time. Contrasts revealed that

each period differed significantly to the other ones, period 2 vs. 1, $F(1, 20) = 34.73, r = .80$, 3 vs. 1, $F(1, 20) = 151.74, r = .94$, and 2 vs. 3, $F(1, 20) = 74.86, r = .89$, all at $p < .01$.

There was no significant main effect of the type of rating (other vs. self), $F(1, 20) = 1.12, p \geq .05, r = .23$.

There was a significant interaction effect between the type of rating (other vs. self) and time-period, $F(1.56, 31.23) = 7.32, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = 7.65, p < .05, r = .53$, and for period 3 compared to period 1, $F(1, 20) = 14.66, p < .01, r = .65$. But results didn't show a significant interaction effect, when comparing other-ratings to self-ratings for period 2 compared to period 3, $F(1, 20) = .06, p \geq .05, r = .06$. Looking at the interaction graph (Figure 4), these effects reflect that self-ratings increased from period 1 to period 2 significantly more than other-ratings. But other-ratings and self-ratings increased with equal ratio from period 2 to period 3.

Correlations of questionnaires/tasks with $|SOD|$ and $|\text{``}EA\text{''}|$. (N=21)

Following significant correlations with $|SOD|$ were found. BVAQ-B fantasizing and period 2, $r = .45, p < .05$. BVAQ-B identifying and period 1, $r = -.45, p < .05$. SPF perspective taking and period 1, $r = .47, p < .05$. SPF perspective taking and period 2, $r = .57, p < .05$. SPF empathy and period 2, $r = .47, p < .05$. Finger lifting task and period 1, $r = -.48, p < .05$. Finger lifting task and period 3, $r = -.45, p < .05$. Correlations with the post video questionnaire (PVQ) revealed the following significant correlations. PVQ-1 (interesting) and period 2, $r = .45, p < .05$. PVQ-2 (attractive) and period 2, $r = .45, p < .05$. PVQ-3 (likeable) and period 2, $r = .46, p < .05$. PVQ-5 (zone out) and period 1, $r = -.51, p < .05$.

Correlations with $|\text{``}EA\text{''}|$ revealed only one significant effect. Reading the Mind in the Eyes correlated with period 3 significantly, $r = -.50, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{``}EA\text{''}|$ with all videos.

3.3.2   G1 Sora 6con – Boss Criticizes Me in front of a Colleague / Chefin kritisiert mich vor Kollegin (congruent, negative)

As shown in Figure 5 the three periods of interest are seconds 45–52 (1. story-introduction), 56–59 (2. salient event & emotional period), and 65–90 (3. reason & cognitive period).
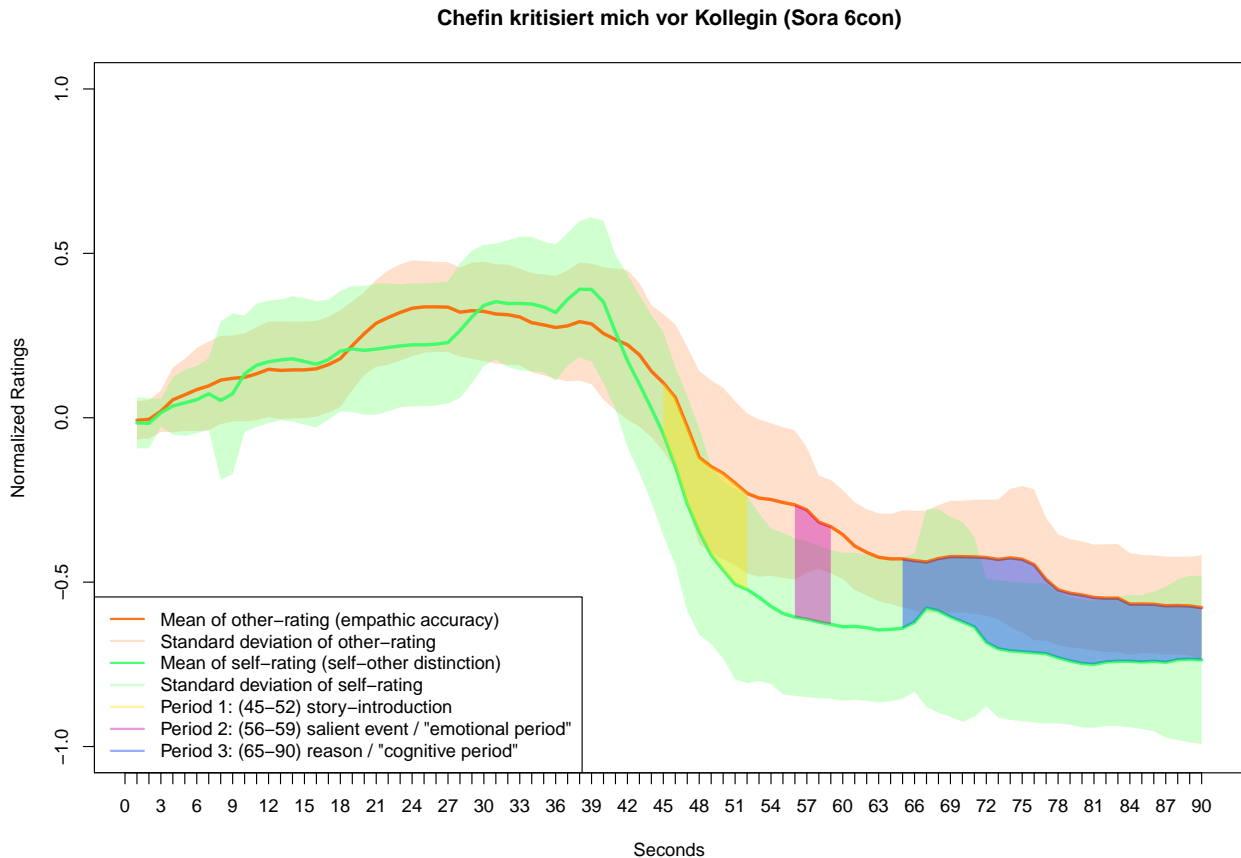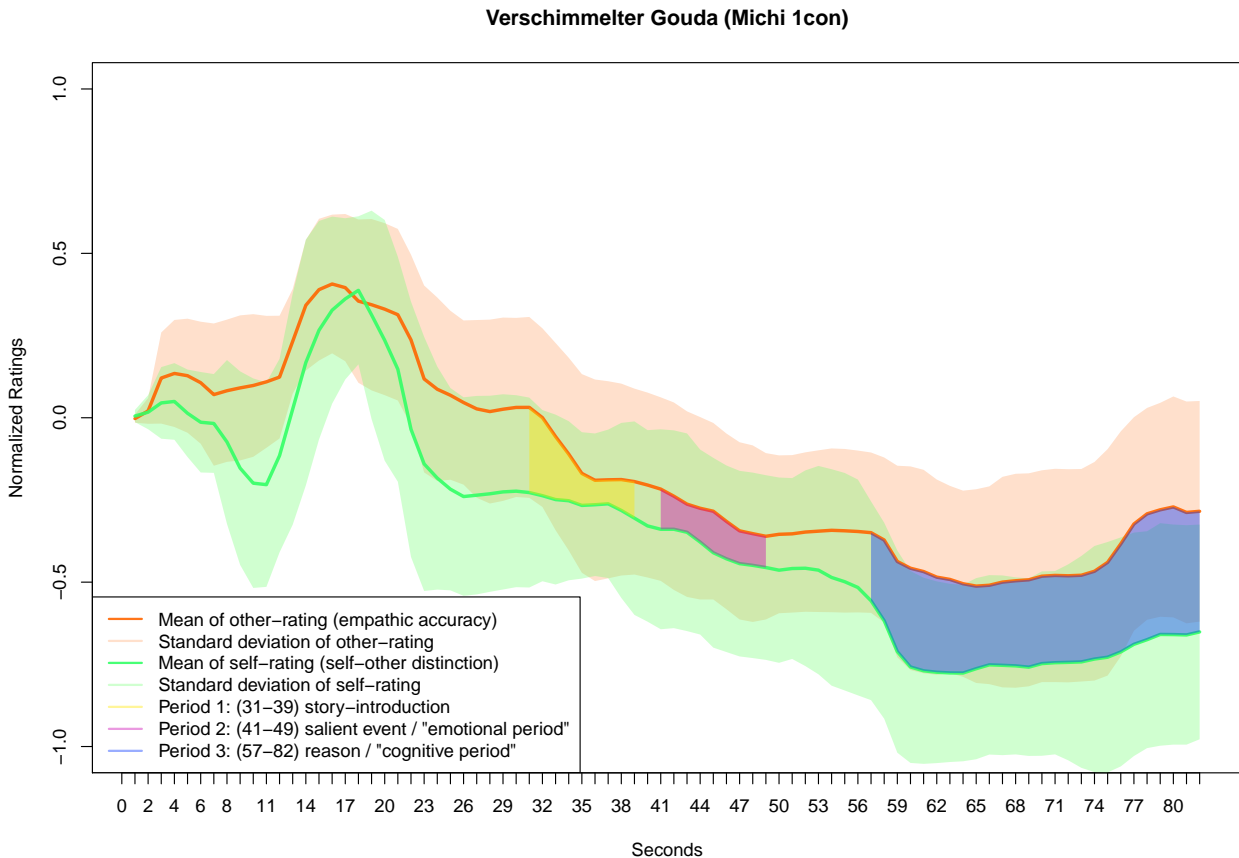

Figure 5: Perceivers' affect-ratings at the video G1 Sora 6con (Boss Criticizes Me in front of a Colleague / Chefin kritisiert mich vor Kollegin).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA revealed that perceivers' other-ratings differed significantly from their self-ratings in all three time-periods of interest; at period 1, $F(1, 20) = 11.88, r = .61$, period 2, $F(1, 20) = 23.23, r = .73$, and period 3 $F(1, 20) = 12.74, r = .62$, all at $p < .01$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 7.

Table 7: G1 Sora 6con: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	mean (sd)	
	other-ratings	*self-ratings*
period 1	−0.09 (0.22)	−0.34 (0.25)
period 2	−0.30 (0.16)	−0.62 (0.23)
period 3	−0.50 (0.15)	−0.70 (0.19)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

To reveal interaction effects of Time x Other/Self two-way repeated-measures ANOVAs were conducted.
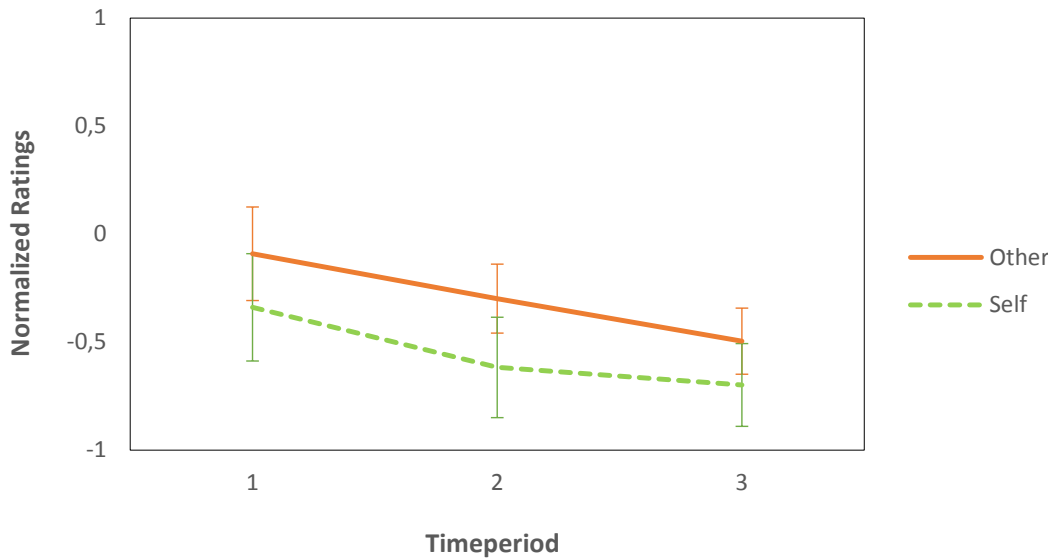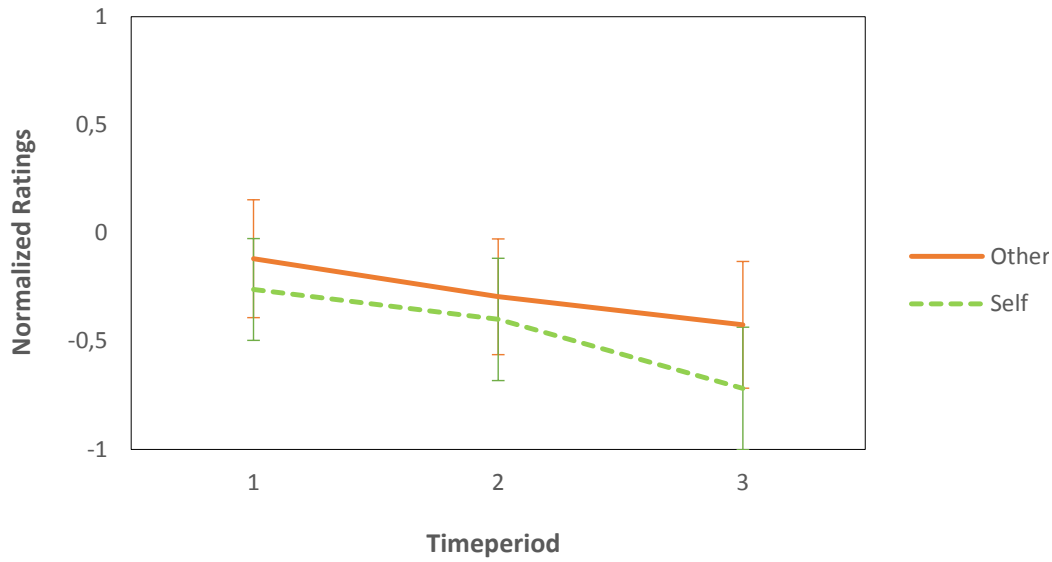

Figure 6: Interaction graph of G1 Sora 6con. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 13.25, p < .01$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .67$).

There was a significant main effect of time, $F(1.33, 26.63) = 73.42, p < .01$. Looking at the interaction graph (Figure 6), it is seen that ratings (other-rating and self-rating), was more

negative with increasing time. Contrasts revealed that each time-period differed significantly to the other ones; period 2 vs. 1, $F(1, 20) = 47.77, r = .84$, period 3 vs. 1, $F(1, 20) = 96.93, r = .91$, and period 2 vs. 3, $F(1, 20) = 61.66, r = .87$, all at $p < .01$.

There was also a significant main effect of rating (other vs. self), $F(1, 20) = 18.88, p < .01, r = .70$.

There was no significant interaction effect between type of rating (other vs. self) and time-period, $F(2, 40) = 2.93, p \geq .05$. To break down this interaction, contrasts were performed comparing all time-periods to each other. Contrasts revealed no significant interaction term, when comparing perceivers' other-ratings to their self-ratings for period 2 compared to period 1, $F(1, 20) = 1.60, p \geq .05, r = .27$, nor for period 3 compared to period 1, $F(1, 20) = 1.07, p \geq .05, r = .22$. However, these contrasts yielded small-to-medium and medium effect sizes. When comparing perceivers' other-ratings to their self-ratings for period 2 compared to period 3, a significant interaction effect was revealed, $F(1, 20) = 6.97, p < .05, r = .51$. Looking at the interaction graph (Figure 6), this effect reflects that other-ratings lowered significantly more from period 2 to period 3 than self-ratings. Or rather, the ratio of the difference between other-ratings and self-ratings at period 2 differed significantly from that at period 3.

Correlations of questionnaires/tasks with $|SOD|$ and $|"EA"|$. (N=21)

The significant correlations with $|SOD|$ are as followed. BVAQ-B identifying and period 3, $r = -.45, p < .05$. Reading the Mind in the Eyes task and period 1, $r = -.53, p < .05$. Correlations with the post video questionnaire (PVQ) revealed the following significant correlations. PVQ-4 (distance) and period 2, $r = -.48, p < .05$. PVQ-4 (distance) and period 3, $r = -.54, p < .05$. PVQ-5 (zone out) and period 3, $r = -.51, p < .05$.

When correlating time-periods of this video with $|"EA"|$, following significant correlations were found. BVAQ-B fantasizing and period 3, $r = .59, p < .01$. BVAQ-B analyzing and period 2, $r = -.48, p < .05$. BVAQ-B affective factor and period 3, $r = .54, p < .05$. BVAQ-B cognitive factor and period 2: $r = -.45, p < .05$. BVAQ-B global score and period 3: $r = .45, p < .05$. Emotional contagion's scale positive and period 1: $r = .53, p < .05$. SPF perspective taking and period 2: $r = .47, p < .05$. Reading the Mind in the Eyes task and period 1: $r = -.61, p < .01$. No significant correlations were found when correlate the post video questionnaire (PVQ) with $|"EA"|$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|"EA"|$ with all videos.

3.3.3 G1 Rudi 3inc – Celebration of a Newborn / Kilias Kindsfeier (incongruent, positive)

As shown in Figure 7 the three periods of interest are seconds 38–46 (1. story-introduction), 58–70 (2. salient event & emotional period), and 76–101 (3. reason & cognitive period).
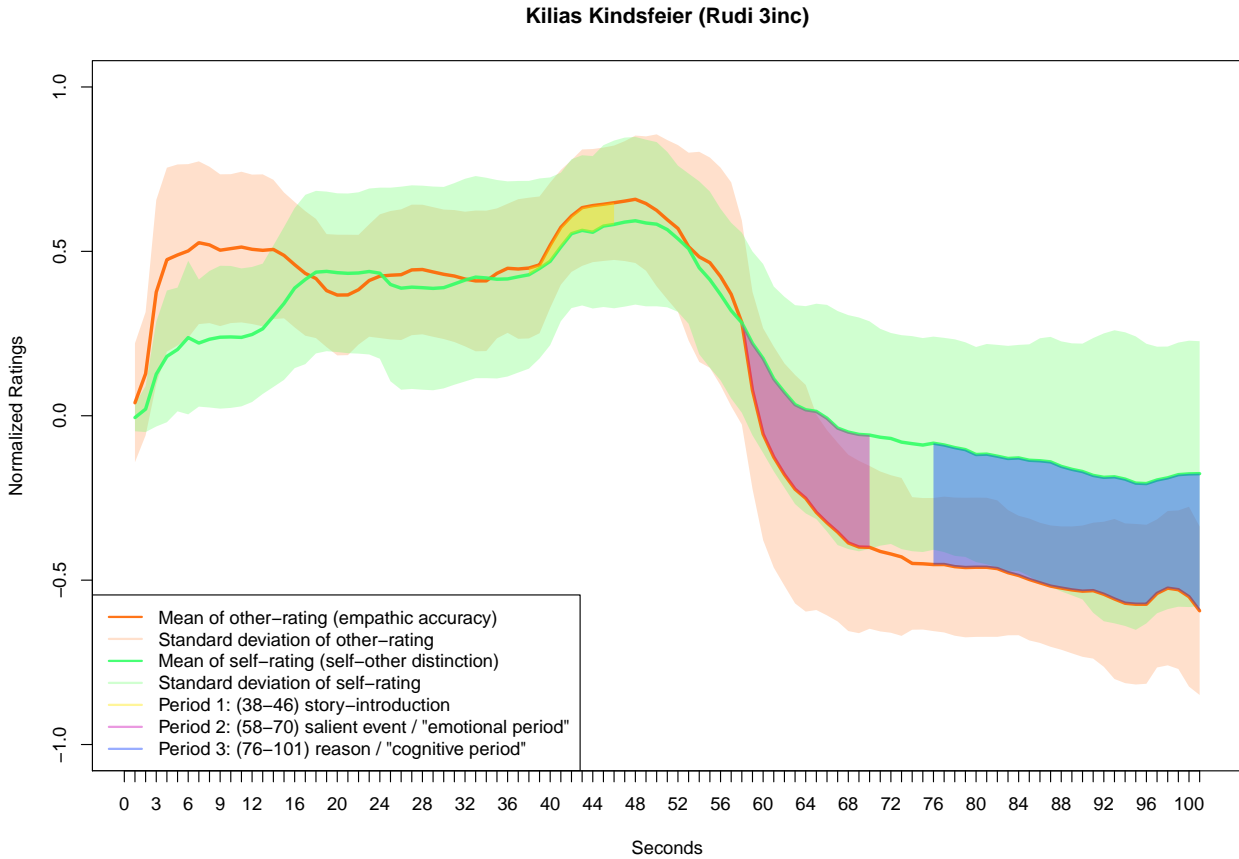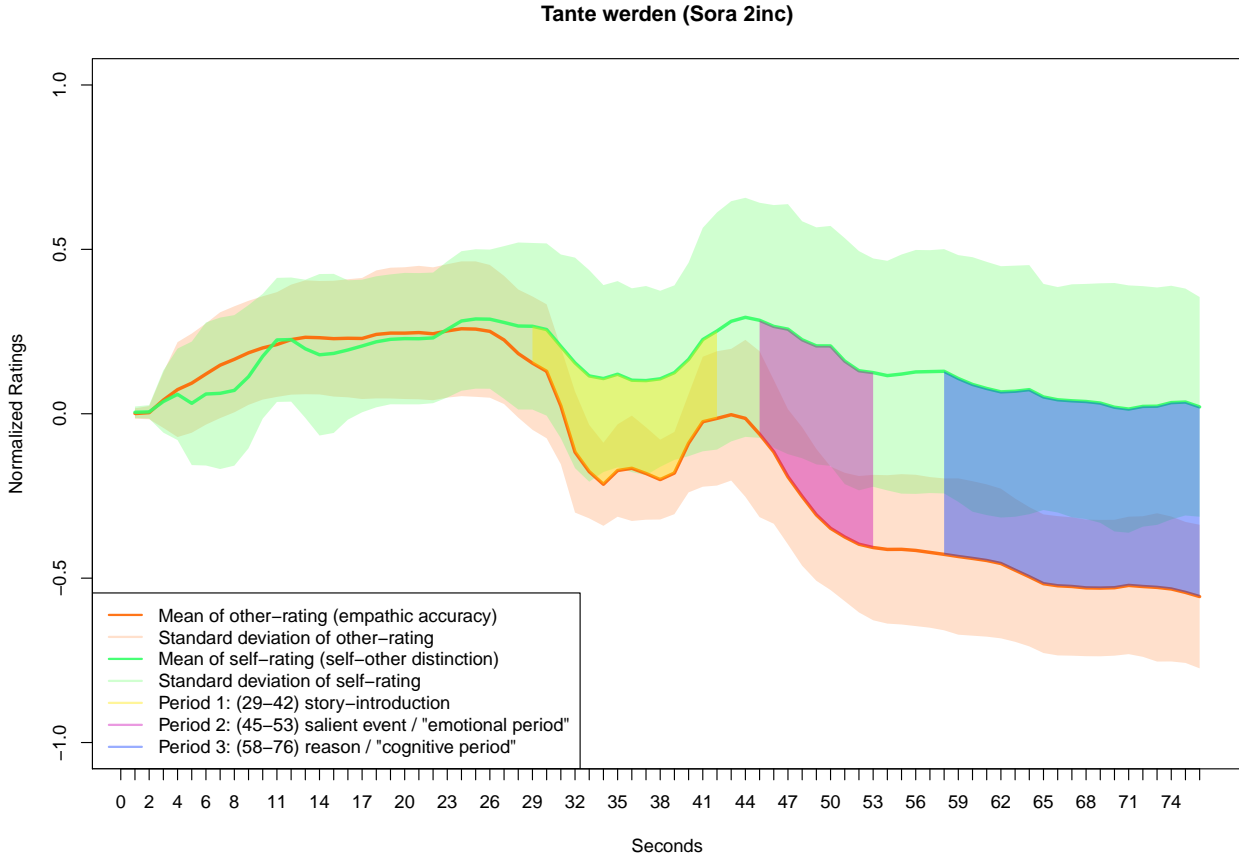

Figure 7: Perceivers' affect-ratings at the video G1 Rudi 3inc (Celebration of a Newborn / Kilias Kindsfeier).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA demonstrated that there was no significant difference between type of rating (other vs. self) in period 1 $F(1, 20) = 2.57, p \geq .05, r = .34$. But other-ratings differed significantly from self-ratings in period 2, $F(1, 20) = 7.08, p < .05, r = .51$, and period 3, $F(1, 20) = 21.52, p < .01, r = 72$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 8.

Table 8: G1 Rudi 3inc: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	0.57 (0.16)	0.52 (0.20)
period 2	-0.20 (0.27)	0.06 (0.29)
period 3	-0.51 (0.20)	-0.15 (0.36)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

Two-way repeated-measures ANOVAs were conducted to reveal interaction effects of Time x Other/Self.
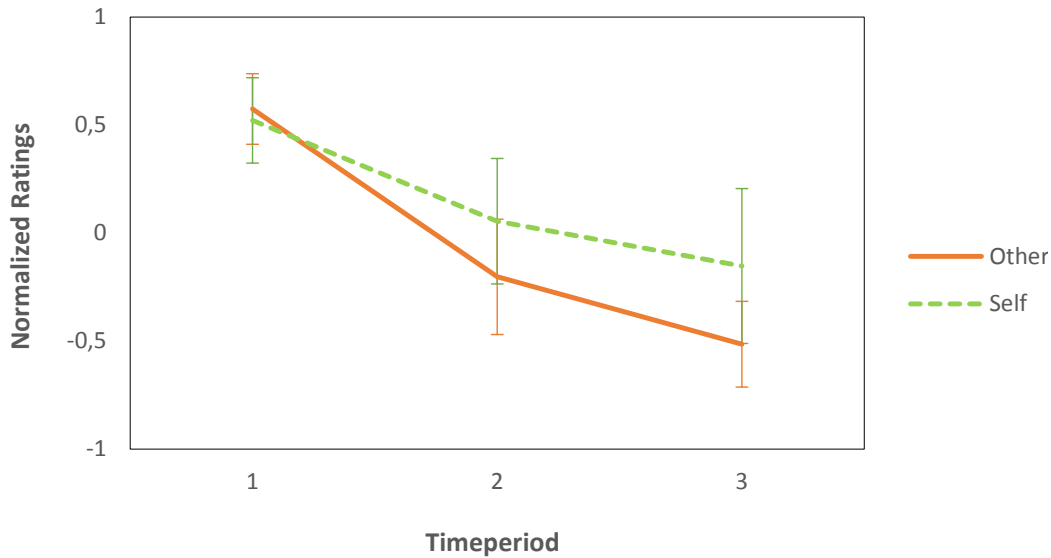
Figure 8: Interaction graph of G1 Rudi 3inc. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent $\pm$ one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 6.25, p < .05$, and for the interaction effect of Time x Other/Self, $\chi^2(2) = 9.56, p < .05$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .78$ for the main effects of time and $\varepsilon = .72$ for the interaction effects of Time x Other/Self.).

There was a significant main effect of time, $F(1.56, 31.24) = 156.84, p < .01$. Looking at the graph (Figure 8), this result reflected that ratings (other-rating and self-rating) were rated more negative with increasing time of the video. Contrasts revealed that each period differed significantly to the other ones, period 2 vs. 1, $F(1, 20) = 164.57, r = .94$, period 3 vs. 1, $F(1, 20) = 197.86, r = .95$, and period 2 vs. 3, $F(1, 20) = 42.87, r = .83$, all at $p < .01$.

There was a significant main effect of type of rating (other vs. self), $F(1, 20) = 10.62, p < .01, r = .59$.

There was also a significant interaction effect of Time x Other/Self, $F(1.43, 28.67) = 14.39, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. This revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = 9.50, p < .01, r = .57$, and for period 3 compared to period 1, $F(1, 20) = 28.12, p < .01, r = .76$. But no significant interaction was observed when comparing other-ratings to self-ratings for period 2 compared to period 3, $F(1, 20) = 3.47, p \geq .05, r = .38$. However, this contrast yielded medium effect size. Looking at the interaction graph (Figure 8), these effects reflect that other-ratings lowered significantly more in period 2 and period 3 than self-ratings. But the ratio between other-ratings and self-ratings was equal at period 2 compared to period 3.

Correlations of questionnaires/tasks with $|SOD|$ and $|``EA"|$. (N=21)

Following correlations with $|SOD|$ were significant. BVAQ-B cognitive factor and period 3, $r = .47, p < .05$. BVAQ-B all sum and period 3, $r = .44, p < .05$. Emotional contagion's positive-scale and period 3, $r = -.47, p < .05$.

Correlations with $|``EA"|$ revealed one significant effect. SPF's fanatsy scale correlated negatively with period 1, $r = -.54, p < .05$. Correlations with the post video questionnaire (PVQ) revealed that there was one significant relationship between PVQ-6 (similar situation) and period 3, $r = .48, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|``EA"|$ with all videos.

3.3.4 G1 Michi 1inc – Mouldy Gouda / verschimmelter Gouda (incongruent, negative)

As shown in Figure 9 the three periods of interest are seconds 30–54 (1. story-introduction), 69–74 (2. salient event & emotional period), and 78–83 (3. reason & cognitive period).
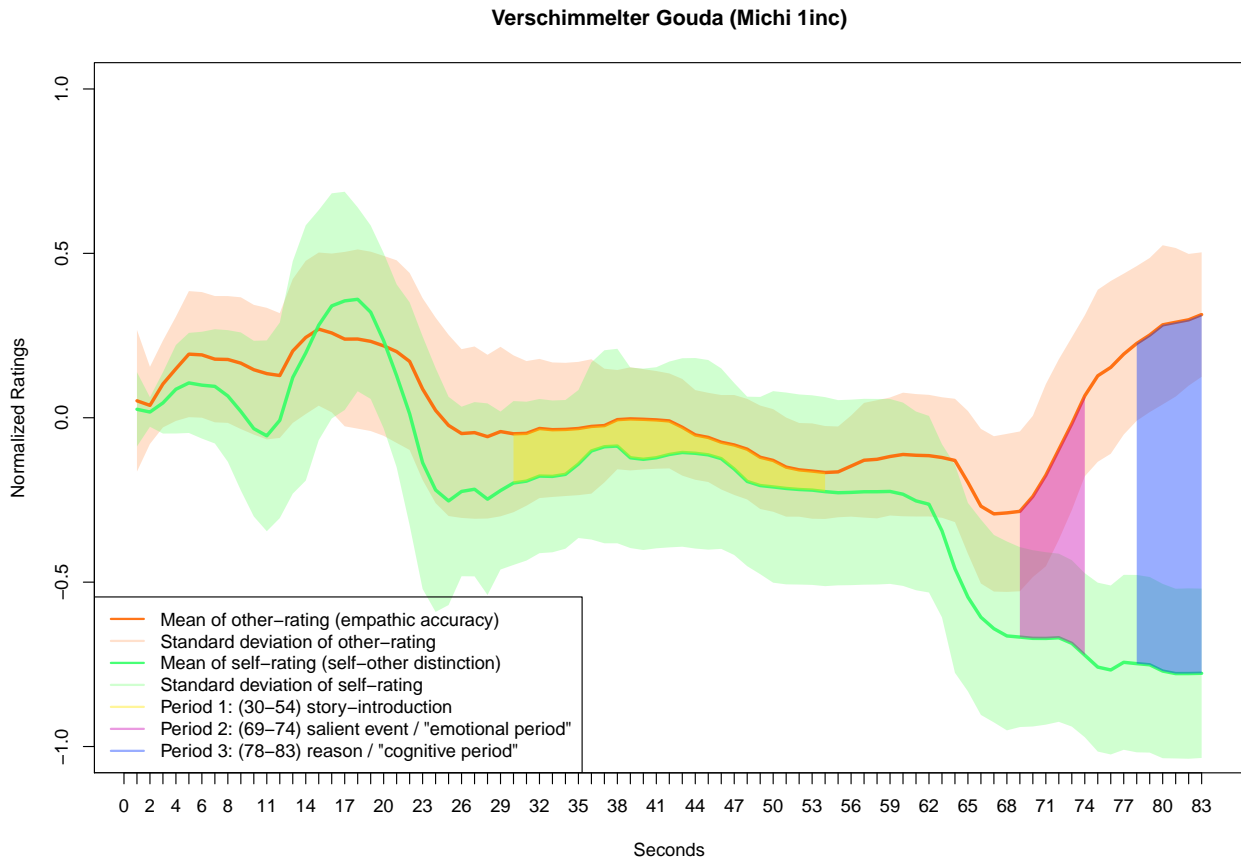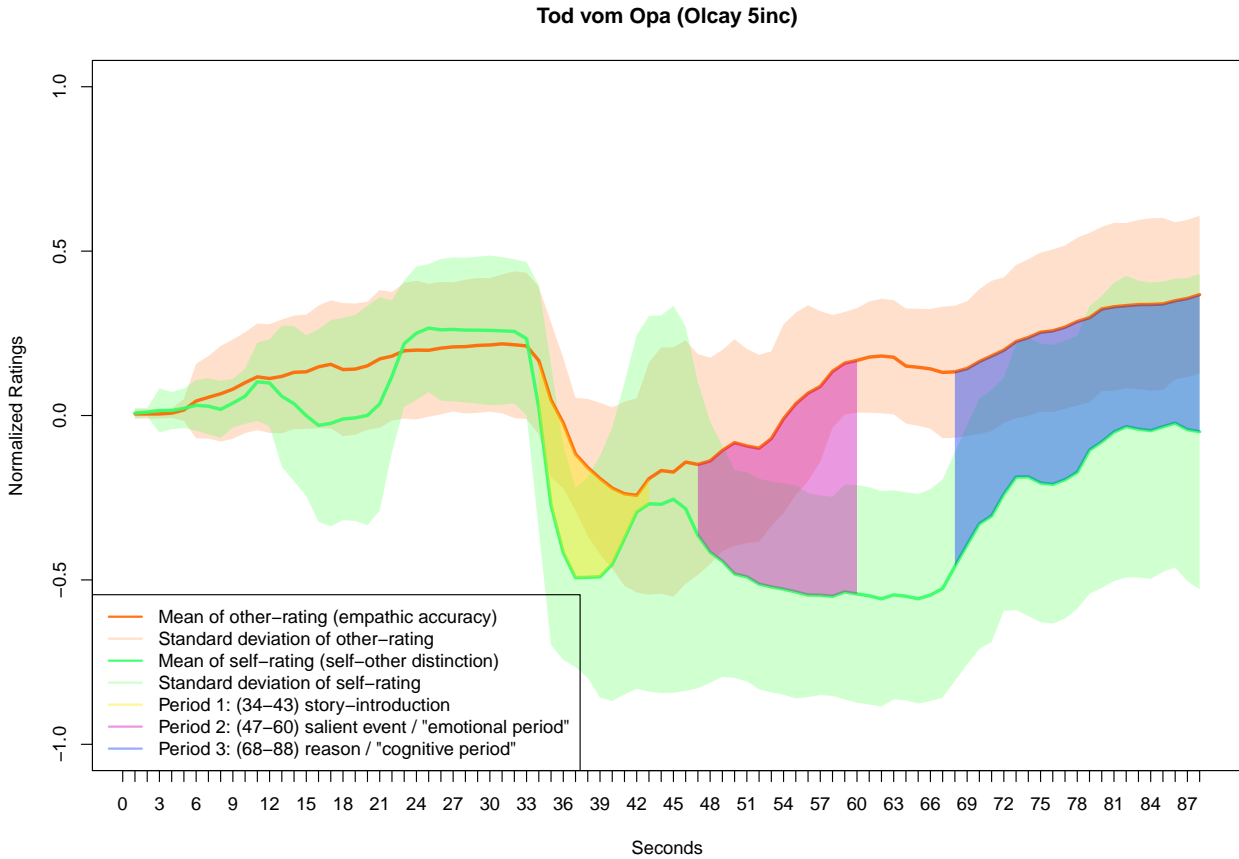

Figure 9: Perceivers' affect-ratings at the video G1 Michi 1inc (Mouldy Gouda / verschimmelter Gouda).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA demonstrated that other-ratings did not differ significantly from self-ratings at period 1, $F(1, 20) = 3.53, p \geq .05, r = .39$. However this effect did yield medium effect size. But other-ratings differed significantly from self-ratings in period 2, $F(1, 20) = 87.45, p < .01, r = .90$, and period 3, $F(1, 20) = 176.72, p < .01, r = .95$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 9.

Table 9: G1 Michi 1inc: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	-0.06 (0.14)	-0.16 (0.24)
period 2	-0.12 (0.22)	-0.68 (0.25)
period 3	0.28 (0.22)	-0.77 (0.26)

Two-way repeated-measures ANOVA: Time x Other/Self. **(N=21)**

To reveal interaction effects of Time x Other/Self two-way repeated-measures ANOVAs were conducted.
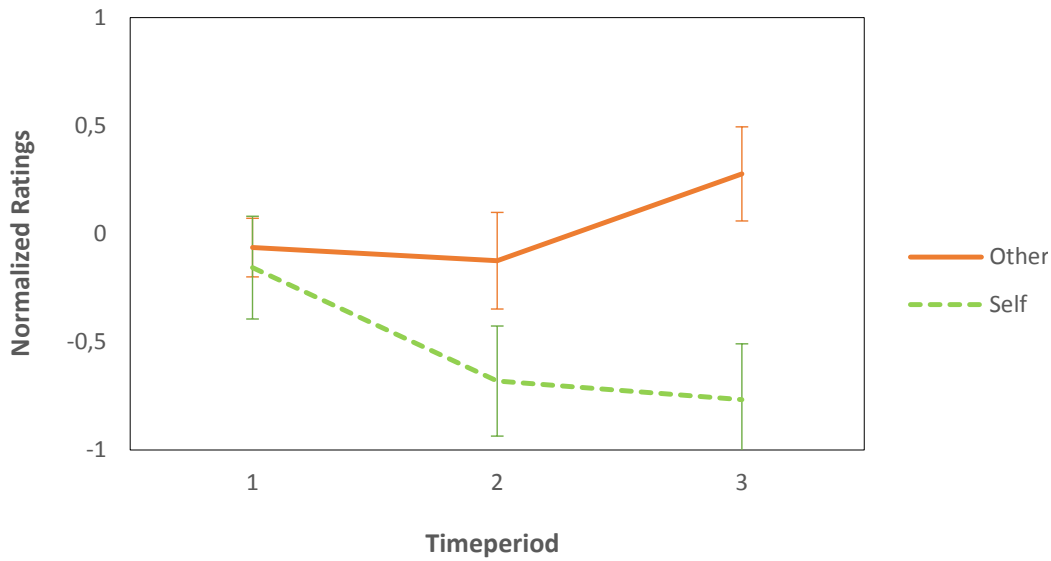
Figure 10: Interaction graph of G1 Michi 1inc. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been met for both, the main effect of time and the interaction effect of Time x Other/Self.

There was a significant main effect of time, $F(2, 40) = 33.02, p < .01$. Contrasts revealed that each time-period differed significantly to the other ones, period 2 vs. period 1, $F(1, 20) =$

$76.49, p < .01, r = .89$, period 3 vs. period 1, $F(1, 20) = 12.12, p < .01, r = .61$, and period 2 vs. period 3, $F(1, 20) = 19.48, p < .01, r = .70$.

There was also a significant main effect of type of rating (other/self), $F(1, 20) = 139.35, p < .01, r = .94$.

Finally there was also a significant interaction effect between time-period and type of rating (other/self), $F(2, 40) = 85.69, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. Contrasts revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = 46.05, p < .01, r = .83$, and for period 3 compared to period 1, $F(1, 20) = 133.79, p < .01, r = .93$, and for period 2 compared to period 3, $F(1, 20) = 53.91, p < .01, r = .85$. Looking at the interaction graph (Figure 10), these effects reflect that self-ratings lowered significantly more in period 2 and period 3 than other-ratings. At each time-period the ratio of other-ratings to self-ratings was significantly different compared to the other time-periods.

Correlations of questionnaires/tasks with $|SOD|$ and $|$"EA"$|$. (N=21)

There were following significant relationships between questionnaires/tasks and $|SOD|$. BVAQ-B emotionalizing and period 1, $r = .50, p < .05$. BVAQ-B verbalizing and period 2, $r = .45, p < .05$. SPF personal distress and period 1, $r = -.59, p < .01$. Shelve task and period 1, $r = -.46, p < .05$. Correlations with the post video questionnaire (PVQ) revealed following significances. PVQ-1 (interesting) and period 1, $r = -.48, p < .05$. PVQ-3 (likeable) and period 3, $r = -.46, p < .05$.

Correlations with $|$"EA"$|$ revealed following significant relationships. Reading the Mind in the Eyes task and period 1, $r = .46, p < .05$. Regarding the post video questionnaire (PVQ) following significant correlations were found. PVQ-3 (likeable) and period 1, $r = -.55, p < .01$, and period 3, $r = -.51, p < .05$. PVQ-4 (distance) and period 3, $r = .65, p < .01$. PVQ-6 (similar situation) and period 3, $r = -.45, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|$"EA"$|$ with all videos.

3.3.5 G2 Rudi 3con – Celebration of a Newborn / Kilias Kindsfeier (congruent, positive)

As shown in Figure 11 the three periods of interest are seconds 42–53 (1. story-introduction), 56–64 (2. salient event & emotional period), and 69–82 (3. reason & cognitive period).

Kilias Kindsfeier (Rudi 3con)

Figure 11: Perceivers' affect-ratings at the video G2 Rudi 3con (Celebration of a Newborn / Kilias Kindsfeier).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA revealed that other-ratings did not differ significantly from self-ratings in all three time-periods, in period 1, $F(1, 20) = .61, p \geq .05, r = .17$, in period 2, $F(1, 20) = .65, p \geq .05, r = .18$, and in period 3, $F(1, 20) = .01, p \geq .05, r = .02$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at

each period are displayed in Table 10.

Table 10: G2 Rudi 3con: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	0.54 (0.27)	0.51 (0.26)
period 2	0.69 (0.26)	0.64 (0.29)
period 3	0.79 (0.25)	0.79 (0.20)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

To reveal interaction effects of Time x Other/Self two-way repeated-measures ANOVAs were conducted.
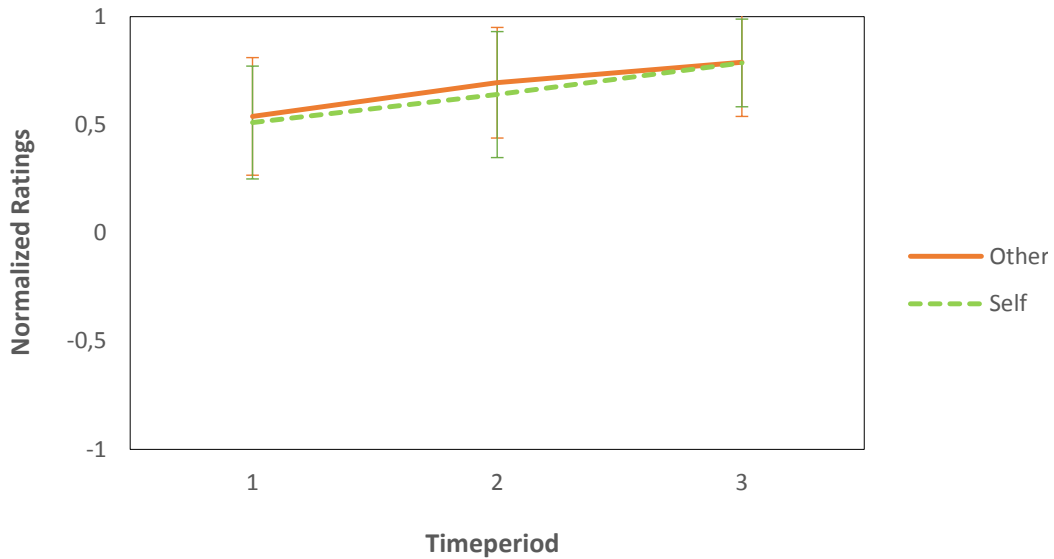
Figure 12: Interaction graph of G2 Rudi 3con. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 7.66, p < .05$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .75$).

There was a significant main effect of time, $F(1.50, 30.04) = 26.95, p < .01$. Contrasts revealed that each time-period differed significantly to the other ones, period 2 vs. period 1,

$F(1, 20) = 12.21, p < .01, r = .62$, period 3 vs. period 1, $F(1, 20) = 41.88, p < .01, r = .82$, period 2 vs. period 3, $F(1, 20) = 26.63, p < .01, r = .76$.

The main effect of type of rating (other/self) was not significant, $F(1, 20) = .51, p \geq .05, r = .16$.

There was no significant interaction effect of type of rating (other/self) and time-period, $F(2, 40) = .51, p \geq .05$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed no significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = .21, p \geq .05, r = .10$, for period 3 compared to period 1, $F(1, 20) = .40, p \geq .05, r = .14$, and for period 2 compared to period 3, $F(1, 20) = .95, p \geq .05, r = .21$. Looking at the interaction graph (Figure 12), these effects reflect that other-ratings and self-ratings increased in all three time-periods with equal ratio.

Correlations of questionnaires/tasks with $|SOD|$ and $|\text{``}EA\text{''}|$. (N=21)

Following significant relationships between the questionnaires/tasks and $|SOD|$ were revealed. BVAQ-B emotionalizing and period 3, $r = .54, p < .05$. BVAQ-B analyzing and period 1, $r = .47, p < .05$. EC negative and period 3, $r = -.49, p < .05$. EC sum and period 3, $r = -.49, p < .05$. SPF fantasy and period 1, $r = -.60, p < .01$. SPF empathy and period 3, $r = -.45, p < .05$. Correlations with the post video questionnaire (PVQ) revealed following significant relationships. PVQ-2 (attractive) and period 2, $r = -.65, p < .01$. PVQ-4 (distance) and period 2, $r = .51, p < .05$. PVQ-5 (zone out) and period 3, $r = .57, p < .01$.

Correlations with $|\text{``}EA\text{''}|$ revealed following significant relationships. Finger Lifting Task and period 2, $r = -.45, p < .05$. Finger Lifting Task and period 3, $r = -.53, p < .05$. PVQ-5 (zone out) and period 1, $r = .60, p < .01$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{``}EA\text{''}|$ with all videos.

3.3.6 G2 Michi 1con – Mouldy Gouda / verschimmelter Gouda (congruent, negative)

As shown in Figure 13 the three periods of interest are seconds 31–39 (1. story-introduction), 41–49 (2. salient event & emotional period), and 57–82 (3. reason & cognitive period).

Figure 13: Perceivers' affect-ratings at the video G2 Michi 1con (Mouldy Gouda / verschimmelter Gouda).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA revealed that other-ratings differed significantly from self-ratings at period 1, $F(1, 20) = 6.61, p < .05, r = .50$. At period 2 other-ratings did not differ significantly from self-ratings, $F(1, 20) = 2.93, p \geq .05, r = .36..$ However this effect yielded medium effect size. At period 3 other-ratings differed another time significantly from

self-ratings, $F(1, 20) = 42.89, p < .01, r = .83$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 11.

Table 11: G2 Michi 1con: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	-0.12 (0.27)	-0.26 (0.24)
period 2	-0.29 (0.27)	-0.40 (0.28)
period 3	-0.42 (0.29)	-0.72 (0.28)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

Two-way repeated-measures ANOVAs were conducted to reveal interaction effects of Time x Other/Self.

Figure 14: Interaction graph of G2 Michi 1con. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent $\pm$ one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 14.64, p < .01$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .65$).

There was a significant main effect of time, $F(1.30, 26.02) = 48.20, p < .01$. Contrasts and Figure 14 revealed that each time-period lowered significantly more compared to the periods before, period 2 vs. period 1, $F(1, 20) = 51.75, p < .01, r = .85$, period 3 vs. period 1, $F(1, 20) = 60.96, p < .01, r = .87$, and period 2 vs. period 3, $F(1, 20) = 29.40, p < .01, r = .77$.

The main effect of type of rating (other/self) was significant too, $F(1, 20) = 14.32, p < .01, r = .65$.

There was a significant interaction effect between time-period and type of rating (other/self), $F(2, 40) = 10.01, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = 1.36, p \geq .05, r = .25$. But there were significant interactions when comparing other-ratings to self-ratings for period 3 compared to period 1, $F(1, 20) = 10.10, p < .01, r = .58$, and for period 2 compared to period 3, $F(1, 20) = 13.24, p < .01, r = .63$. Looking at the interaction graph (Figure 14), these effects reflect that at period 3 self-ratings lowered significantly more than other-ratings. While the ratio of other-ratings to self-ratings was equal at period 2 compared to period 1, it differed significantly both at period 3 compared to period 1 and at period 3 compared to period 2.

Correlations of questionnaires/tasks with $|SOD|$ and $|\text{"}EA\text{"}|$. (N=21)

No significant correlations between the questionnaires/tasks and $|SOD|$ were found. Correlations between questionnaires/tasks and $|\text{"}EA\text{"}|$ revealed one significant relationship, PVQ-6 (similar situation) and period 3, $r = .49, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{"}EA\text{"}|$ with all videos.

3.3.7 G2 Sora 2inc – Become an Aunt / Tante werden (incongruent, positive)

As shown in Figure 15 the three periods of interest are seconds 29–42 (1. story-introduction), 45–53 (2. salient event & emotional period), and 58–76 (3. reason & cognitive period).

Figure 15: Perceivers' affect-ratings at the video G2 Sora 2inc (Become an Aunt / Tante werden). The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA revealed that other-ratings differed significantly from self-ratings at all three time-periods, period 1, $F(1, 20) = 18.52, p < .01, r = .69$, period 2, $F(1, 20) = 24.04, p < .01, r = .74$, and period 3, $F(1, 20) = 31.93, p < .01, r = .78$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 12.

Table 12: G2 Sora 2inc: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	mean (sd)	
	other-ratings	*self-ratings*
period 1	−0.09 (0.09)	0.16 (0.24)
period 2	−0.27 (0.18)	0.21 (0.35)
period 3	−0.50 (0.21)	0.05 (0.34)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

Two-way repeated-measures ANOVAs were conducted to reveal interaction effects of Time x Other/Self.
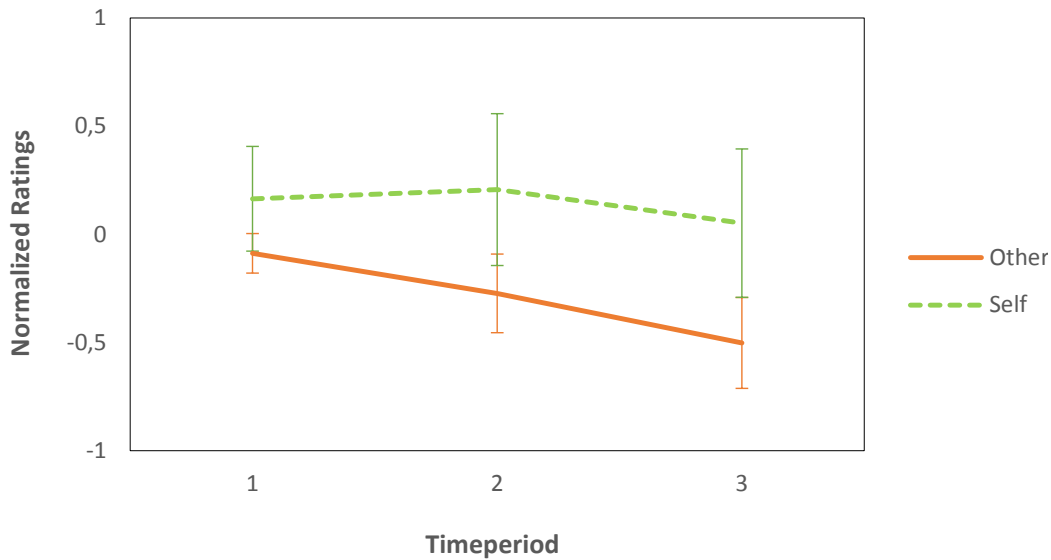
Figure 16: Interaction graph of G2 Sora 2inc. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been met for both the main effect of time and the interaction effect of Time x Other/Self.

There was a significant main effect of time, $F(2, 40) = 26.12, p < .01$. Contrasts revealed that each time-period differed significantly to the other time-periods, period 2 vs. period 1, $F(1, 20) = 5.39, p < .01, r = .46$, period 3 vs. period 1, $F(1, 20) = 36.53, p < .01, r = .80$, and period 2 vs. period 3, $F(1, 20) = 25.99, p < .01, r = .75$.

The main effect of type of rating (other/self) was also significant, $F(1, 20) = 30.80, p <$

.01, $r = .78$. At Figure 16 it is apparent that other-ratings differed from self-ratings at each time-period.

There was a significant interaction effect between time-period and type of rating (other/self), $F(2, 40) = 10.47, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = 12.06, p < .01, r = .61$ and for period 3 compared to period 1, $F(1, 20) = 17.38, p < .01, r = .68$. But no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 3 was revealed, $F(1, 20) = 1.18, p \geq .05, r = .24$. Looking at the interaction graph (Figure 16), this effect reflects that other-ratings lowered significantly more at period 2 and period 3 compared to self-ratings. But more importantly the proportion of other-ratings to self-ratings differed significantly both at period 2 compared to period 1, and at period 3 compared to period 1, but not when comparing period 2 to period 3. At the last two periods (period 2 and period 3) the ratio between the ratings was parallel.

*** Correlations of questionnaires/tasks with $|SOD|$ and $|\text{"}EA\text{"}|$. (N=21)***

There were two significant relationships when correlating the questionnaires/tasks and $|SOD|$, Finger Lifting Task and period 2, $r = .46, p < .05$, and also with period 3, $r = .49, p < .05$.

No significant correlations between the questionnaires/tasks and $|\text{"}EA\text{"}|$ were found.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{"}EA\text{"}|$ with all videos.

3.3.8 G2 Olcay 5inc – Death of Grandfather / Tod vom Opa (incongruent, negative)

As shown in Figure 17 the three periods of interest are seconds 34–43 (1. story-introduction), 47–60 (2. salient event & emotional period), and 68–88 (3. reason & cognitive period).

Figure 17: Perceivers' affect-ratings at the video G2 Olcay 5inc (Death of Grandfather / Tod vom Opa).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=21)

The one-way repeated-measures ANOVA revealed that other-ratings differed significantly from self-ratings in all three time-periods, in period 1, $F(1, 20) = 19.55, p < .01, r = .70$, in period 2, $F(1, 20) = 46.04, p < .01, r = .83$, and in period 3, $F(1, 20) = 23.68, p < .01, r = .74$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at

each period are displayed in Table 13.

Table 13: G2 Olcay 5inc: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	−0.12 (0.17)	−0.35 (0.30)
period 2	−0.01 (0.21)	−0.50 (0.31)
period 3	0.27 (0.22)	−0.16 (0.37)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=21)

To reveal interaction effects of Time x Other/Self two-way repeated-measures ANOVAs were conducted.
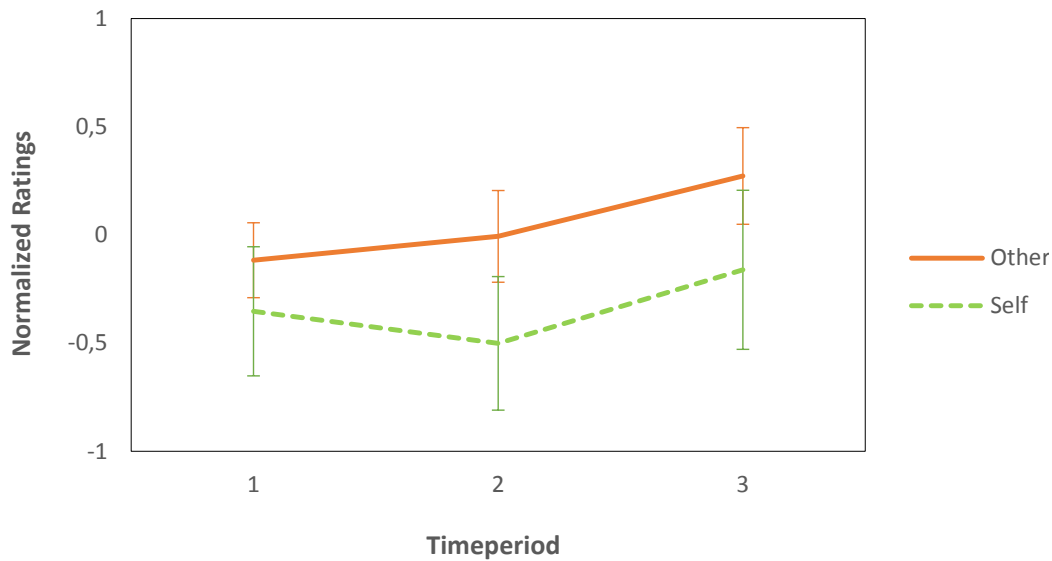
Figure 18: Interaction graph of G2 Olcay 5inc. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 10.45, p < .01$, and the interaction effect of Time x Other/Self, $\chi^2(2) = 8.68, p < .05$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .70$ for the main effect of time and $\varepsilon = .73$ for interaction effect of Time x Other/Self).

There was a significant main effect of time, $F(1.41, 28.11) = 16.16, p < .01$. Contrasts revealed that ratings at period 2 did not differ significantly from them at period 1, $F(1, 20) = .24, p \geq .05, r = .11$. But ratings at period 3 were significantly higher than at period 1, $F(1, 20) = 14.80, p < .01, r = .65$, and at period 2, $F(1, 20) = 24.26, p < .01, r = .74$.

The main effect of type of rating (other/self) was significant too, $F(1, 20) = 41.02, p < .01, r = .82$.

There was also a significant interaction effect between time-period and type of rating (other/self), $F(1.46, 29.27) = 7.17, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 20) = 32.37, p < .01, r = .79$, and for period 3 compared to period 1, $F(1, 20) = 5.50, p < .05, r = .46$, but not for period 3 compared to period 2, $F(1, 20) = .60, p \geq .05, r = .17$. Looking at the interaction graph (Figure 18), these effects reflect while at period 2 (compared to period 1) self-ratings lowered and the difference between other-ratings and self-ratings increased, at period 3 both ratings increased parallel (with maintaining the same ratio to each other).

Correlations of questionnaires/tasks with $|SOD|$ and $|\text{"}EA\text{"}|$. (N=21)

There were following significant relationships when correlating the applied questionnaires/tasks with $|SOD|$. BVAQ-B emotionalizing and period 1, $r = -.46, p < .05$. BVAQ-B affective factor and period 1, $r = -.45, p < .05$. EC negative-scale and period 1, $r = .45, p < .05$. Finger Lifting Task and period 1, $r = .44, p < .05$. Finger Lifting Task and period 2, $r = .56, p < .01$. Correlations with the post video questionnaire (PVQ) revealed the following significant correlations. PVQ-3 (likeable) and period 2, $r = -.52, p < .05$. PVQ-3 (likeable) and period 3, $r = -.51, p < .05$. PVQ-4 (distance) and period 1, $r = .44, p < .05$ and with period 3, $r = .52, p < .05$.

There was one significant relationship when correlating the questionnaires/tasks and $|\text{"}EA\text{"}|$, SPF empathy and period 1, $r = -.47, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{"}EA\text{"}|$ with all videos.

3.3.9 G3 Sora 5con – 1st Boyfriend / 1. Freund (congruent, positive)

As shown in Figure 19 the three periods of interest are seconds 26–47 (1. story-introduction), 50–66 (2. salient event & emotional period), and 69–85 (3. reason & cognitive period).
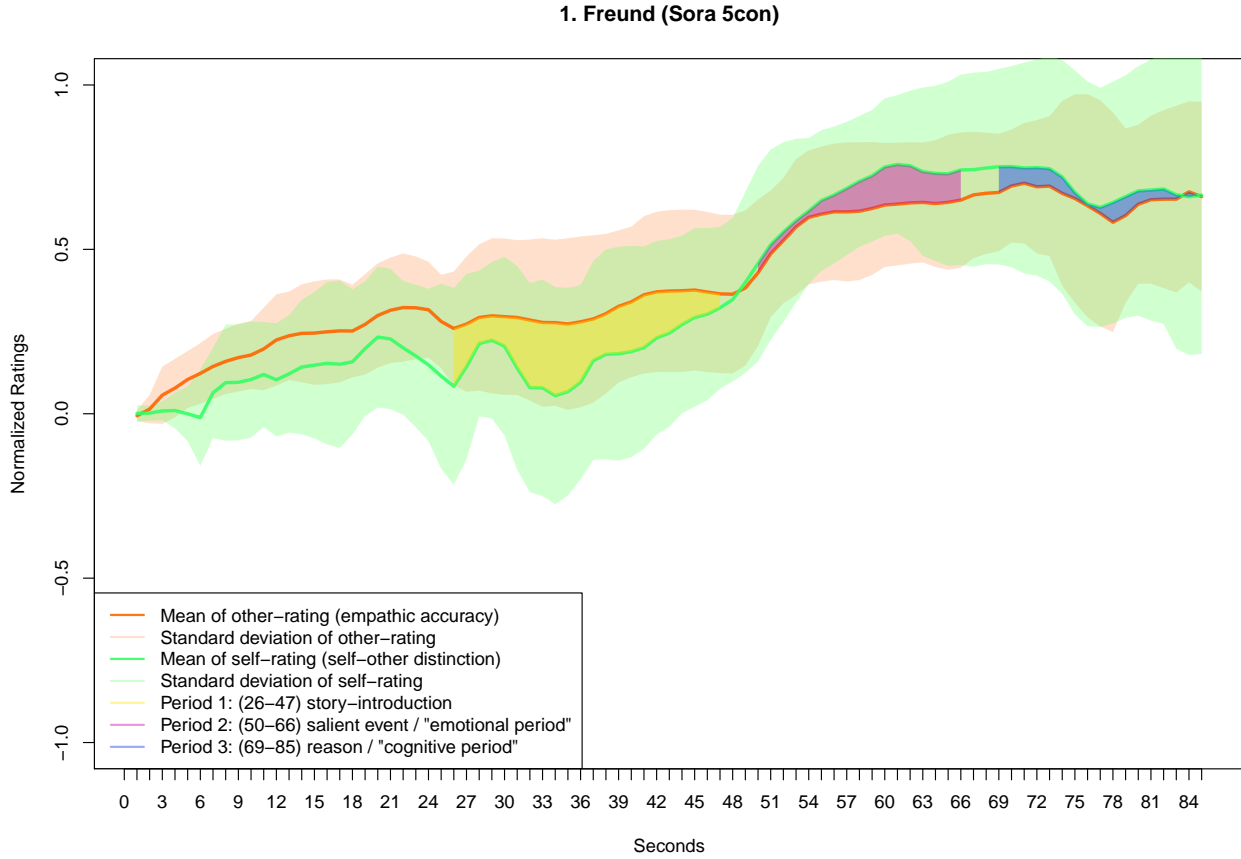
Figure 19: Perceivers' affect-ratings at the video G3 Sora 5con (1st Boyfriend / 1. Freund)
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=19)

The one-way repeated-measures ANOVA revealed that other-ratings differed significantly from self-ratings in all three time-periods, in period 1, $F(1, 18) = 7.15, p < .05, r = .53$, in period 2, $F(1, 18) = 6.10, p < .05, r = .50$, and in period 3, $F(1, 18) = 7.30, p < .05, r = .54$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 14.

Table 14: G3 Sora 5con: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	0.33 (0.20)	0.20 (0.21)
period 2	0.61 (0.19)	0.71 (0.18)
period 3	0.69 (0.21)	0.77 (0.20)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=19)

Two-way repeated-measures ANOVAs were conducted to reveal interaction effects of Time x Other/Self.
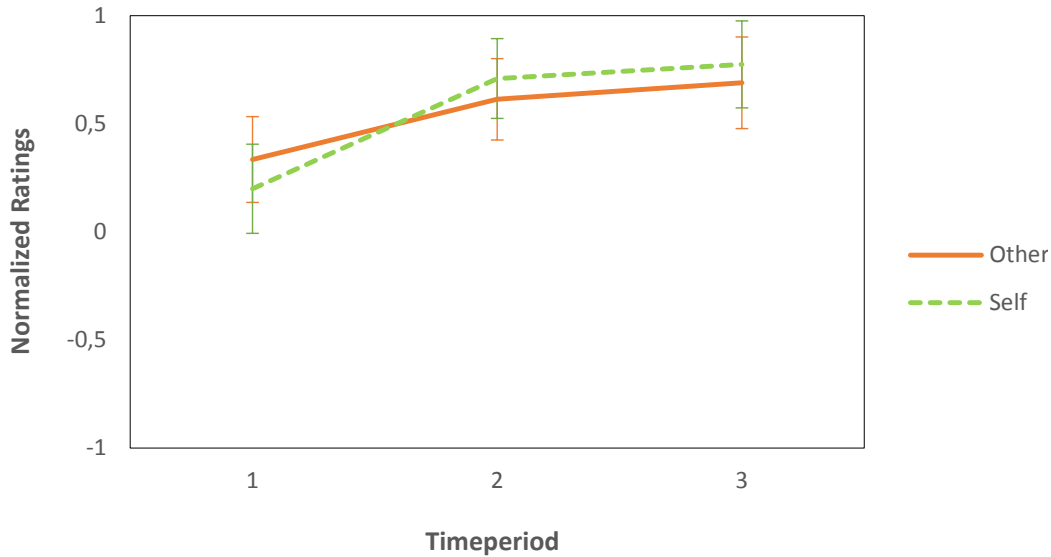
Figure 20: Interaction graph of G3 Sora 5con. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 7.67, p < .05$, and for the interaction effect Time x Other/Self, $\chi^2(2) = 10.28, p < .01$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .73$ for the main effect of time and $\varepsilon = .69$ for the interaction effect of Time x Other/Self).

There was a significant main effect of time, $F(1.47, 26.41) = 74.23, p < .01$. Contrasts revealed that each time-period increased significantly more compared to the time-periods before,

period 2 vs. period 1, $F(1, 18) = 81.68, p < .01, r = .91$, period 3 vs. period 1, $F(1, 18) = 88.28, p < .01, r = .91$, and period 2 vs. period 3, $F(1, 18) = 6.95, p < .05, r = .53$.

There was no significant main effect of type of rating (other/self), $F(1, 18) = .23, p \geq .05, r = .11$.

There was a significant interaction effect between time-period and type of rating (other/self), $F(1.38, 24.76) = 17.22, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 18) = 34.81, p < .01, r = .81$, and for period 3 compared to period 1, $F(1, 18) = 14.76, p < .01, r = .67$. But there was no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 3, $F(1, 18) = .11, p \geq .05, r = .08$. Looking at the interaction graph, Figure 20, these effects reflect that self-ratings increased from period 1 to period 2 significantly more than other-ratings. But other-ratings and self-ratings increased with equal ratio from period 2 to period 3.

Correlations of questionnaires/tasks with $|SOD|$ and $|"EA"|$. (N=19)

Correlations of questionnaires/tasks with $|SOD|$ revealed the following two significant relationships. SPF empathic concern and period 2, $r = -.46, p < .05$. SPF perspective taking and period 2, $r = -.47, p < .05$.

There were the following significant relationships between the applied questionnaires and $|"EA"|$. BVAQ-B fantasizing and period 2, $r = .60, p < .01$. BVAQ-B affective factor and period 2, $r = .48, p < .05$. SPF personal distress and period 2, $r = .50, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|"EA"|$ with all videos.

3.3.10 G3 Michi 3con – Dad Drank Away my Apple-Juice / Papa trinkt Apfelsaft weg (congruent, negative)

As shown in Figure 21 the three periods of interest are seconds 25–47 (1. story-introduction), 56–72 (2. salient event & emotional period), and 75–86 (3. reason & cognitive period).
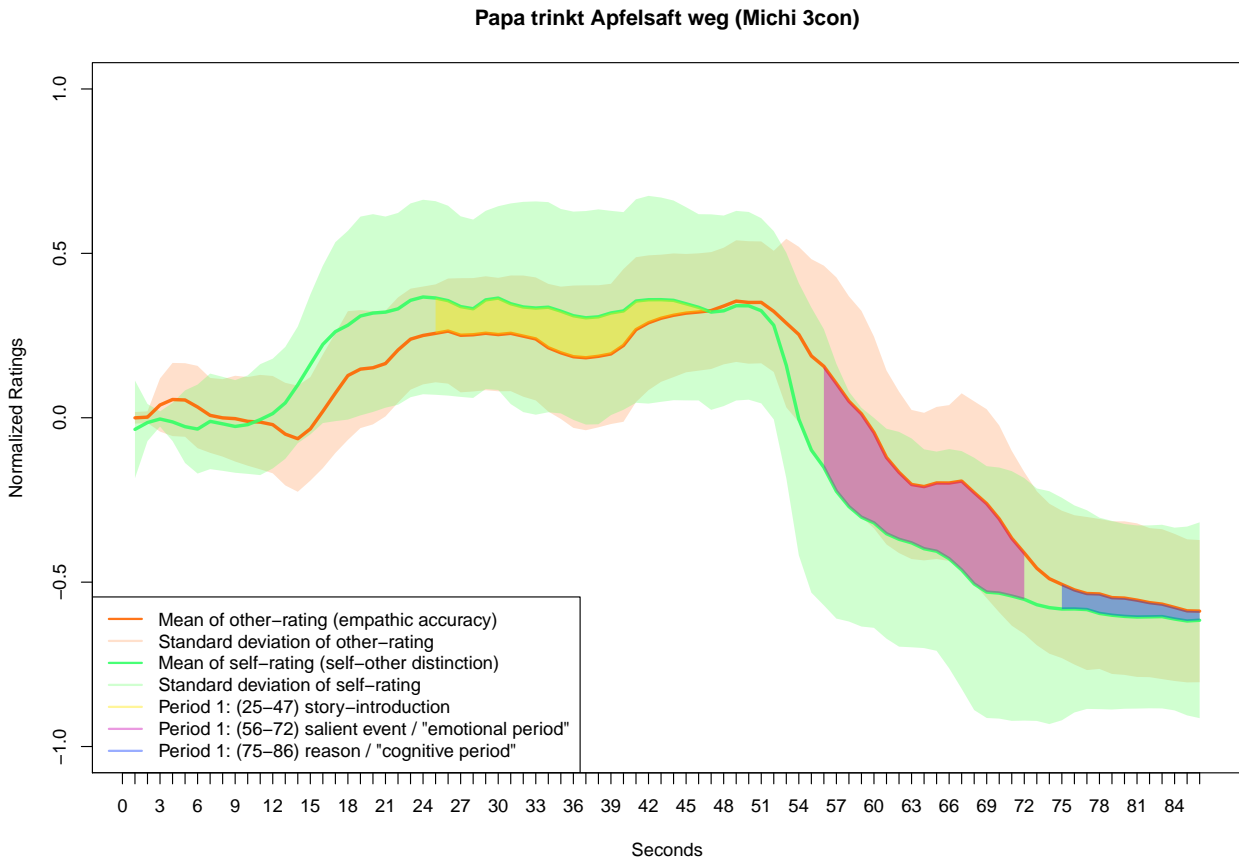
Figure 21: Perceivers' affect-ratings at the video G3 Michi 3con (Dad Drank Away my Apple-Juice / Papa trinkt Apfelsaft weg).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=19)

The one-way repeated-measures ANOVA revealed that other-ratings did not differ significantly from self-ratings at period 1, $F(1, 18) = .79, p \geq .05, r = .20$. But other-ratings and self-ratings differed significantly from each other at period 2, $F(1, 18) = 7.16, p < .05, r = .53$. At period 3 other-ratings did not differ significantly from self-ratings, $F(1, 18) = .35, p \geq .05, r = .14$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 15.

Table 15: G3 Michi 3con: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	0.25 (0.17)	0.31 (0.28)
period 2	−0.16 (0.25)	−0.40 (0.34)
period 3	−0.54 (0.23)	−0.59 (0.30)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=19)

To reveal interaction effects of Time x Other/Self two-way repeated-measures ANOVAs were conducted.
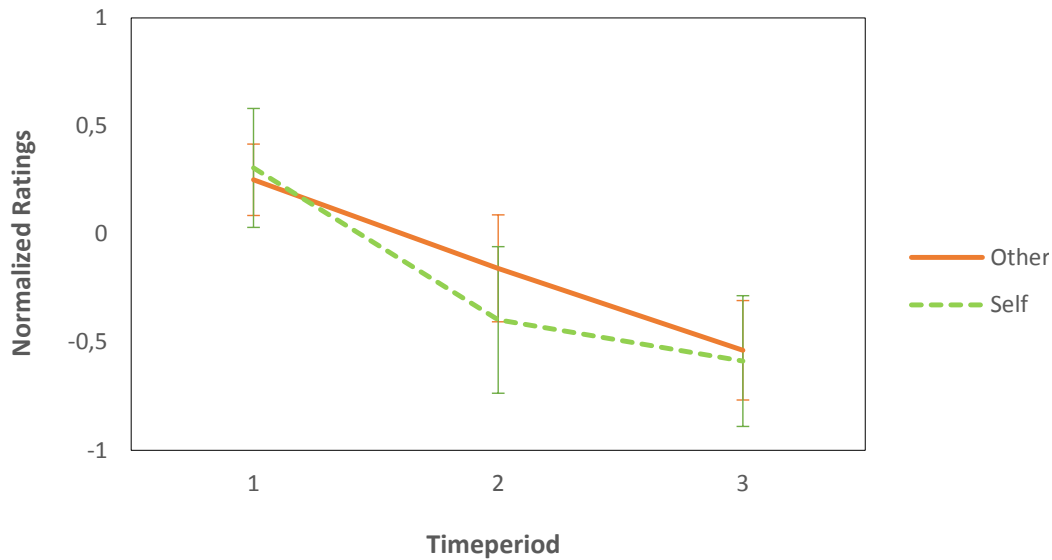
Figure 22: Interaction graph of G3 Michi 3con. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of time, $\chi^2(2) = 12.16, p < .01$, and for the interaction effect of Time x Other/Self, $\chi^2(2) = 6.72, p < .05$. Therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of spericity ($\varepsilon = .66$ for the main effect of time and $\varepsilon = .75$ for the interaction effect of Time x Other/Self).

There was a main effect of time, $F(1.32, 23.83) = 134.00, p < .01$. Contrasts revealed that each time-period lowered significantly compared to the time-period before, period 2 vs. period 1, $F(1, 18) = 82.64, p < .01, r = .91$, period 3 vs. period 1, $F(1, 18) = 192.84, p < .01, r = .96$, period 2 vs. period 3, $F(1, 18) = 103.93, p < .01, r = .92$.

There was no significant main effect of type of rating (other/self), $F(1, 18) = 2.14, p \geq .05, r = .33$. However this effect yielded medium effect size.

There was a significant interaction effect between time-period and type of rating (other/self), $F(1.51, 27.14) = 4.26, p < .05$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 18) = 7.76, p < .05, r = .55$, and for period 2 compared to period 3, $F(1, 18) = 7.22, p < .05, r = .54$. But no significant interaction effect was revealed when comparing other-ratings to self-ratings for period 3 compared to period 1, $F(1, 18) = .72, p \geq .05, r = .20$. Looking at the interaction graph, Figure 22, these effects reflect that although other-ratings and self-ratings lowered significantly from each time-period to the next time-period, in period 2 self-ratings lowered significantly more than other-ratings.

Correlations of questionnaires/tasks with $|SOD|$ and $|\text{``}EA\text{''}|$. (N=19)

There were the following significant relationships between the applied questionnaires/tasks and $|SOD|$. BVAQ-B emotionalizing and period 1, $r = .46, p < .05$. SPF empathic concern and period 2, $r = -.50, p < .05$. SPF perspective taking and period 2, $r = -.53, p < .05$. SPF empathy and period 2, $r = -.54, p < .05$. Two significant relationships with the post video questionnaire (PVQ) were found. PVQ-5 (zone out) and period 2, $r = .56, p < .05$ and & period 3: $r = .62, p < .01$.

Correlations of questionnaires/tasks with $|\text{``}EA\text{''}|$ revealed the following significant relationships. BVAQ-B fantasizing and period 2, $r = .52, p < .05$. BVAQ-B affective factor and period 2, $r = .67, p < .01$. BVAQ-B all sum and period 2, $r = .52, p < .05$. SPF personal distress and period 2, $r = .74, p < .01$. Finger Lifting Task and period 2, $r = .49, p < .05$. PVQ-1 (interesting) and period 3, $r = .47, p < .05$. PVQ-2 (attractive) and period 2, $r = .47, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{``}EA\text{''}|$ with all videos.

3.3.11 G3 Rudi 5inc – Girlfriend Dumped Me / Freundin macht Schluß (incongruent, negative)

As shown in Figure 23 the three periods of interest are seconds 29–41 (1. story-introduction), 45–63 (2. salient event & emotional period), and 67–85 (3. reason & cognitive period).
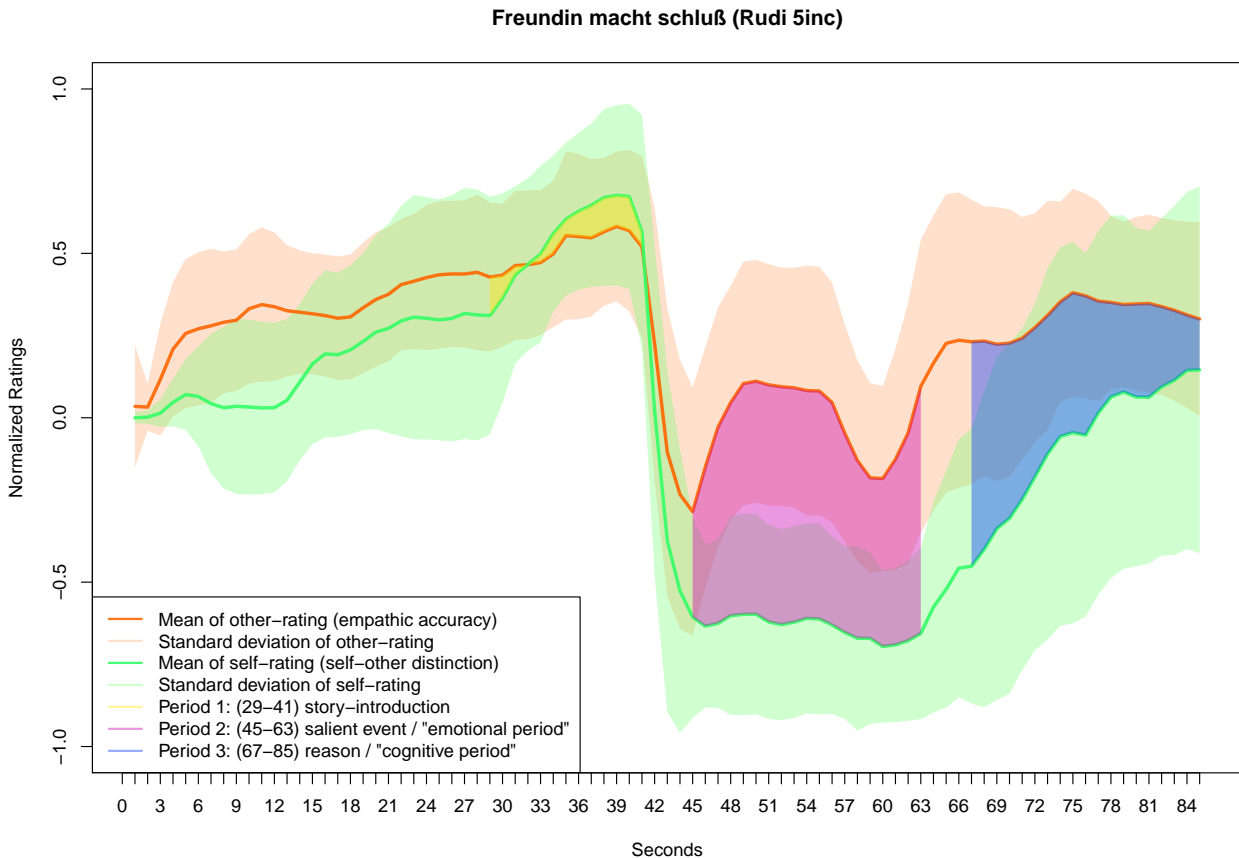
Figure 23: Perceivers' affect-ratings at the video G3 Rudi 5inc (Girlfriend Dumped Me / Freundin macht Schluß).
The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=19)

The one-way repeated-measures ANOVA revealed that other-ratings did not differ significantly from self-ratings at period 1, $F(1, 18) = 1.97, p \geq .05, r = .31$. However this effect did yield a medium effect size. But other-ratings differed significantly from self-ratings at period 2, $F(1, 18) = 57.93, p < .01, r = .87$, and period 3, $F(1, 18) = 10.46, p < .01, r = .61$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 16.

Table 16: G3 Rudi 5inc: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	mean (sd)	
	other-ratings	*self-ratings*
period 1	0.52 (0.21)	0.58 (0.21)
period 2	−0.01 (0.31)	−0.64 (0.25)
period 3	0.34 (0.27)	−0.07 (0.47)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=19)

Two-way repeated-measures ANOVAs were conducted to reveal interaction effects of Time x Other/Self.
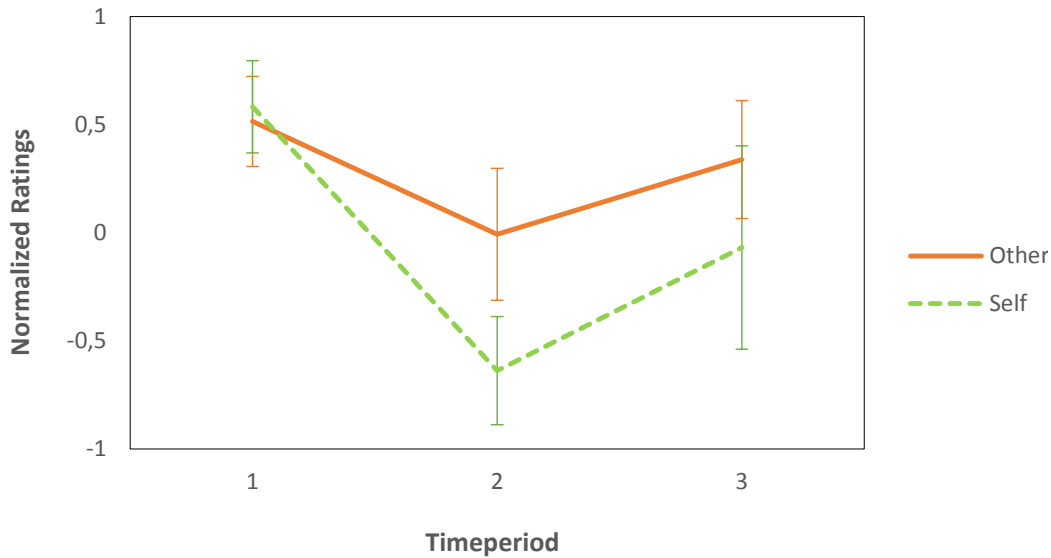
Figure 24: Interaction graph of G3 Rudi 5inc. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been met for the main effect and interaction effect.

There was a significant main effect of time, $F(2, 36) = 88.61, p < .01$. Contrasts revealed that each time-period differed significantly to the other ones, period 2 vs. period 1, $F(1, 18) =$

$170.16, p < .01, r = .95$, period 3 vs. period 1, $F(1,18) = 36.10, p < .01, r = .82$, and period 2 vs. period 3, $F(1,18) = 57.23, p < .01, r = .87$.

The main effect of type of rating (other/self) was also significant, $F(1,18) = 31.80, p < .01, r = .80$.

There was a significant interaction effect between time-period and type of rating (other/self), $F(2,36) = 16.78, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1,18) = 45.08, p < .01, r = .85$, and for period 3 compared to period 1, $F(1,18) = 11.79, p < .01, r = .63$. But no significant interaction effect was revealed, when comparing other-ratings to self-ratings for period 3 compared to period 2, $F(1,18) = 3.22, p \geq .05, r = .39$. However this effect did yield a medium effect size. Looking at the interaction graph, Figure 24, these effects reflect that from period 1 to period 2 self-ratings lowered significantly more compared to other-ratings. And both ratings increased while maintaining the same ratio to each other (i.e., parallel to each other).

Correlations of questionnaires/tasks with $|SOD|$ and $|"EA"|$. (N=19)

Correlations of the applied questionnaires/tasks with $|SOD|$ revealed following significant relationships. BVAQ-B analyzing and period 1, $r = -.48, p < .05$ and & period 3: $r = -.48, p < .05$. BVAQ-B affective factor and period 1, $r = -.53, p < .05$. BVAQ-B all sum and period 1, $r = -.60, p < .01$ and with period 3: $r = -.50, p < .05$. SPF empathic concern and period 1, $r = .47, p < .05$. With the post video questionnaire (PVQ) significant correlations were as followed. PVQ-3 (likeable) and period 2, $r = .50, p < .05$. PVQ-6 (similar situation) and period 2, $r = .56, p < .05$.

There were also significant relationships between $|"EA"|$ and the applied questionnaires/tasks. BVAQ-B Fantasizing and period 2, $r = .72, p < .01$. BVAQ-B affective factor and period 2, $r = .57, p < .05$. SPF personal distress and period 2, $r = .55, p < .05$. PVQ-2 (attractive) and period 2, $r = .51, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|"EA"|$ with all videos.

3.3.12 G3 Stephan 3inc – Banana Cake / Bananenschnitte (incongruent, negative)

As shown in Figure 25 the three periods of interest are seconds 23–45 (1. story-introduction), 61–74 (2. salient event & emotional period), and 79–84 (3. reason & cognitive period).
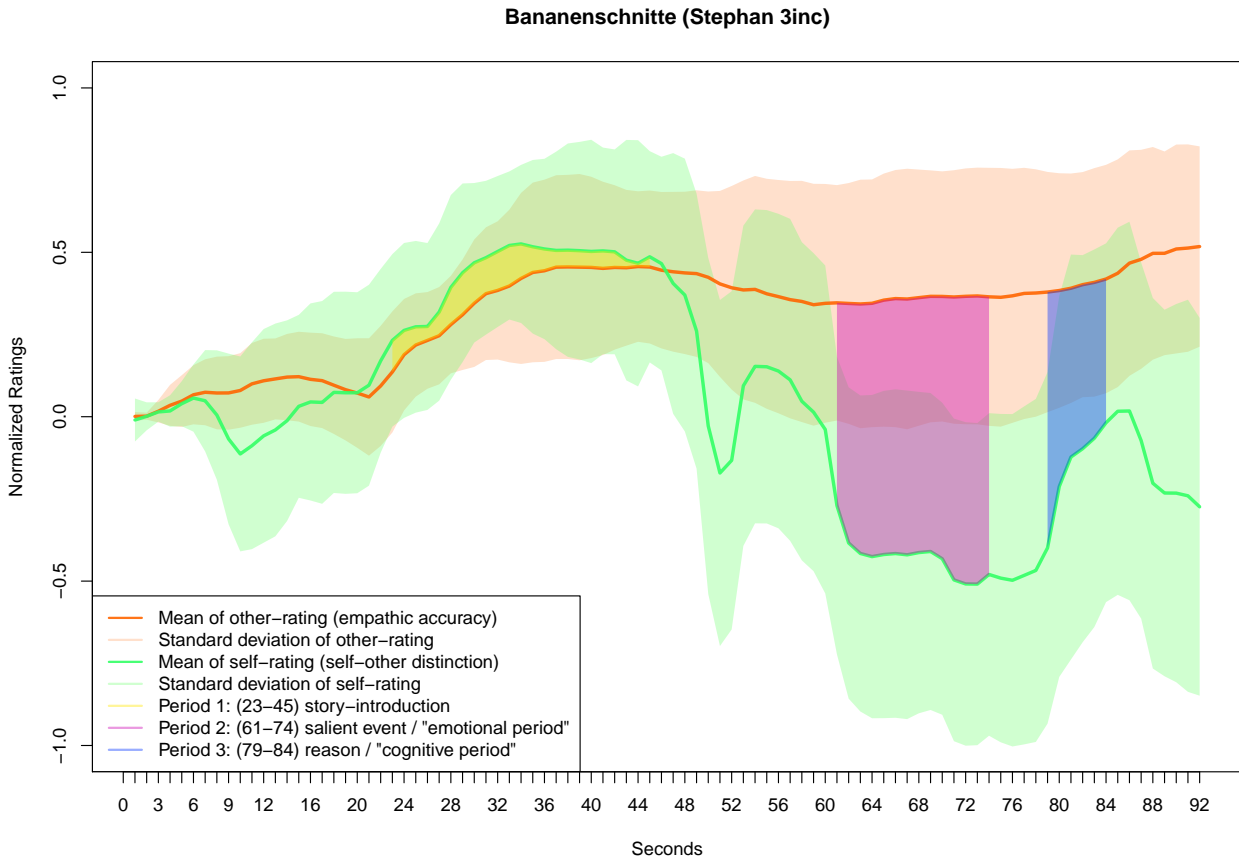
Figure 25: Perceivers' affect-ratings at the video G3 Stephan 3inc (Banana Cake / Bananenschnitte). The three periods of interest are: (1) the story-introduction (yellow), (2) the salient event & emotional period (pink), and (3) the reason & cognitive period (blue).
The mean and standard deviation of perceivers' other-ratings (when rating the emotional valence of the target-person) are in orange colouring. Whereas mean and standard deviation of participants' self-ratings (when rating the own emotional valence) are in green tone.

One-way repeated-measures ANOVA: other/self. (N=19)

The one-way repeated-measures ANOVA revealed that other-ratings did not differ significantly from self-ratings in period 1, $F(1, 18) = 1.22, p \geq .05, r = .25$. But other-ratings differed significantly from self-ratings in period 2, $F(1, 18) = 44.46, p < .01, r = .84$, and in period 3, $F(1, 18) = 19.54, p < .01, r = .72$.

Means and standard deviations of perceivers' other-ratings and perceivers' self-ratings at each period are displayed in Table 17.

Table 17: G3 Stephan 3inc: Means and standard deviations of perceivers' other-ratings and their self-ratings.

time	*mean (sd)*	
	other-ratings	*self-ratings*
period 1	0.40 (0.18)	0.47 (0.24)
period 2	0.46 (0.20)	−0.44 (0.50)
period 3	0.49 (0.21)	−0.15 (0.53)

Two-way repeated-measures ANOVA: Time x Other/Self. (N=19)

Two-way repeated-measures ANOVAs were conducted to reveal interaction effects of Time x Other/Self.
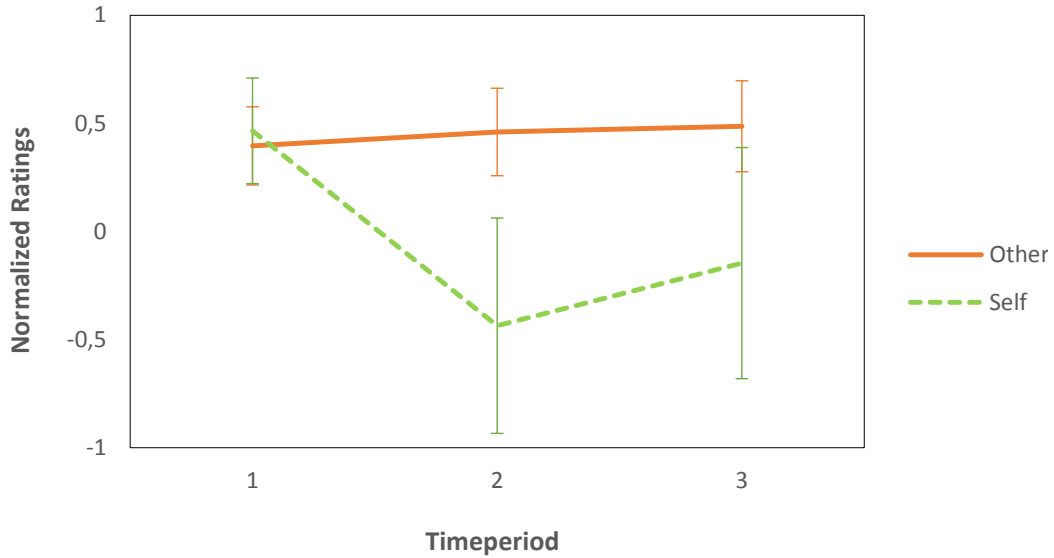
Figure 26: Interaction graph of G3 Stephan 3inc. The x-axis displays the three time-periods of interest while the y-axis presents the mean of type of rating, other-ratings (continuous line) and self-ratings (dashed line). Error bars represent ± one standard deviation.

Mauchly's test indicated that the assumption of sphericity had been met for the main effect and interaction effect.

There was a significant main effect of time, $F(2, 36) = 33.58, p < .01$. Contrasts revealed that each time-period differed significantly from the other ones, period 2 vs. period 1, $F(1, 18) = 59.35, p < .01, r = .88$, period 3 vs. period 1, $F(1, 18) = 22.06, p < .01, r = .74$, and period 2 vs. period 3, $F(1, 18) = 12.78, p < .01, r = .64$.

The main effect of type of rating (other/self) was also significant, $F(1, 18) = 24.17, p < .01, r = .76$.

There was a significant interaction effect between time-period and type of rating (other/self), $F(2, 36) = 37.99, p < .01$. To break down this interaction, contrasts were performed comparing all time-periods to each other. These revealed a significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 1, $F(1, 18) = 64.62, p < .01, r = .88$, for period 3 compared to period 1, $F(1, 18) = 27.61, p < .01, r = .78$, and for period 2 compared to period 3, $F(1, 18) = 9.82, p < .01, r = .59$. Looking at the interaction graph, Figure 26, these effects reflect that at each time-period the difference between other-ratings and self-ratings was different. At period 1 other-ratings and self-ratings didn't differ from each other. At period 2 self-ratings lowered significantly compared to other-ratings. And at period 3 self-ratings increased but still remained with a significant difference to other-ratings.

Correlations of questionnaires/tasks with $|SOD|$ *and* $|\text{"}EA\text{"}|$. (N=19)

Following significant relationships between the applied questionnaires/tasks and $|SOD|$ were found. SPF empathic concern and period 1, $r = -.58, p < .01$. SPF perspective taking and period 1, $r = -.55, p < .05$. SPF empathy and period 1, $r = -.55, p < .05$. Reading the mind in the Eyes and period 3, $r = .54, p < .05$. Shleve task and period 3, $r = .52, p < .05$. Correlations with the post video questionnaire (PVQ) revealed following significant correlations. PVQ-6 (similar situation) and period 2, $r = -.62, p < .01$, and with period 3, $r = -.52, p < .05$.

Correlations of the applied questionnaires/tasks with $|\text{"}EA\text{"}|$ revealed following significant relationships. BVAQ-B identifying and period 3, $r = -.48, p < .05$. BVAQ-B analyzing and period 2, $r = -.50, p < .05$, and with period 3, $r = -.49, p < .05$.

Table 26 (p. 164) and Table 27 (p. 165) give an overview of all significant correlations with $|SOD|$ and $|\text{"}EA\text{"}|$ with all videos.

3.3.13 Post Video Questionnaire (PVQ)

Table 28 at page 166 shows the results of the Post Video Questionnaire. Immediately after watching all four videos, each participant was asked to fill out a Post Video Questionnaire (PVQ) (see p. 152). The questions were with regard to how (1) interesting, (2) attractive, and (3) likeable they found the target-persons. Further if (4) participants wished to distance from the described situation and if (5) they zoned out. And finally participants were asked if (6) they had been in a similar situation once in their life. Results are listed in Table 28. Questions had to be answered on a 9-point Likert scale (except the question concerning if participants had been in a similar situation once, which was to be answered on a 4-point scale). Hence the score may range between 1 and 9, whereat 9 signifies the maximal achievable score.

For the results of the correlations with $|SOD|$ and $|"EA"|$ see the overviews at Table 26 (p. 164) and Table 27 (p. 165).

3.3.14 Validation study: The plausibility of the videos.

At the end of the validation study, participants were asked if they found all videos common or if there was anything strange? Or if they want to comment on anything? Some participants mentioned that they didn't believe one or two videos were true, other participants stated that they think that all videos were true and that they were not very special. Table 18 presents which videos were stated by participants as not-true.

Table 18: Validation study: plausibility of videos.

Video	Participants, who stated following video as not-true
G1 Michi 1inc	319, 332, 338, 344, 349, 361
G2 Sora 2inc	357
G2 Olcay 5inc	327, 340, 357
G3 Rudi 5inc	323, 330, 354
G3 Stephan 3inc	317, 323, 330

4 Discussion

4.1 I. Development of Stimuli

The development of the stimuli was without problems. The realization was well considered in advance. Especially the before deliberated time-structure of the videos was helpful. The emotions and reasons for the incongruent videos were reflected together with the target-person. This preparatory work was important so that the target-persons could identify satisfactorily with this fictitious story.

4.2 II. Pilot Study

The pilot study served to select those videos whose stories were believable and further those at which the target's emotions were comprehensible and it was possible to empathize with the target-person.

Especially the semi-structured interview was helpful to detect those videos which distracted the perceiver's attention or were assumed to be not-true. Because the pilot study video-questionnaire took place immediately after watching the video and before watching a next one, it was considered to have no question in this questionnaire regarding the plausibility of a video. Because a question asking if the just seen video is true or not, may lead to focus on this topic and further to assume too many videos as not true. Only those videos which were obviously not true should be excluded. Interestingly also a few congruent videos were stated as not true by participants of the pilot study. This might be due to the target-persons' excitement of getting filmed while describing a life-event. Or maybe because of the artificial environment of the laboratories for the participants.

One result of the pilot study was that some target-persons and hence all their videos had to be excluded for further research. Twice the t-shirt imprint distracted that much participants attention so that participants focused more on the t-shirt imprints than on the story. This factor can easily be considered by asking the targets to wear simple, maybe one-colored t-shirts. One target-person's mimic and gesture wasn't found to be appropriate. And finally one target-person split participants in two groups with extreme emotions towards her. While one group liked her and her style most of all target-persons, the other group couldn't stand her at all. This target-person had a quite expressive or rather said histrionic way of describing her life-events and her corresponding emotions. To analyze the way of describing life-events and its impact on empathy on an individual level and in reference to the target-person's personality

might be a very interesting topic for future research.

Participants also reported some very interesting statements (see also Table 25 at p 163). Like that the attractiveness of one target-person changed in one video depending on the story. To the video G2 Olcay 5inc one participant stated that at first she wanted to rate her attractiveness with a 8 (which would be the second highest score at the used 9-point Likert-scale). But as Olcay described her feelings she had as her beloved grandfather died, the participant didn't know anymore if she found Olcay attractive anymore. Regarding the possibility to empathize with the target-person, to Rudi it was said that it didn't seem that he was acting, maybe because of his accent, and that he was the most natural of all. To Sora once it was said that her expression was neutral, hence it was difficult to estimate her emotions. And one participant pointed out that he is quite unsure if the stories were true. Subsequent he said that this factor doesn't matter for the task of feeling with the person in the video. A more concrete interpretation would need further qualitative research on this topic.

4.3   III. Validation Study

4.3.1   Data collection.

As described in section 2.5 each participant watched four videos, in random order. Two of them were congruent and the other two incongruent. While videos of the congruent condition consist of a story with feelings and explanations one "normally" expects and which suite the described event, contrary thereto in the incongruent condition target's feelings are not suiting the told life-event. The target-person describes a life-event and his/her emotional reaction thereto is in an unexpected and mostly in an opposite way to the feelings which are commonly expected in that situation. After about 20–30 seconds the target-person gives a reason for the uncommon feelings and clears away a possible incomprehension.

As shown in Figure 1 (p. 17) and described at page 17, the structure of each video consists of target's self-introduction (to enable the perceiver to get familiar with the target's way of describing and facial expression) and following of three periods of interest. In the congruent condition after an introduction into the story (period 1), the emotions will be pointed out (period 2) and finally a reason and explanation will be given (period 3). Incongruent videos have the same structure except the one difference that at incongruent stimuli in period 2 the emotions are described in an unexpected way and do not fit the told life-event.

These three time-periods of interest (period 1, period 2, and period 3) serve as a basic material to collect data for the topic of this diploma thesis—self-other distinction (at incongruent

stimuli) and empathic accuracy (at congruent and incongruent stimuli).

To make their emotions visible and measurable, an affect rating dial was constructed. While watching the video participants could rate the requested emotional valence on-line and a second-by-second attunement was possible.

Each video was watched by the target-person her-/himself. While watching they rated their own emotional valence they had while they presented their story by using the affect rating dial (target's self-rating). Further their ECG and EDA data were recorded (target's physiological data). Physiological data will be analyzed in another master thesis.

Perceivers watched each video twice. In the first run participants were instructed to feel with the person in the video and to rate the emotional valence the target-person displays (other-rating). In a second run they got instructed to put themselves in the described situation and to rate the own emotional valence they would have then (self-rating). Like the targets, perceivers' physiological data (ECG & EDA) were recorded too (perceiver's physiological data). Physiological data will be analyzed in another thesis.

4.3.2 Operationalization & Hypotheses.

Self-other distinction (SOD) is operationalized as the difference between perceiver's other-rating (when rating the emotional valence they ascribed to the target-person) and their self-rating (rating their own emotional valence) (see page 22). In concrete this means that, when at an incongruent video SOD is observable (= perceiver's other-rating differs significantly from perceiver's self-rating) at period 2 and period 3, this video provides a useful stimulus to examine SOD.

Concerning empathic accuracy (EA), first it was thought to operationalize EA as the match of perceiver's other-rating and target's self-rating (see p. 23). But due to the fact that targets' self-rating did not work the expected way (see section 4.3.5), an ad hoc hypothesis to calculate an alternative variation of empathic accuracy was considered. Hence this alternative so-called empathic accuracy = "empathic accuracy" ("EA") and was defined as the match of perceiver's other-rating and the mean of all perceivers' other-ratings (for the definition of "EA" see p. 25).

Further if at a congruent video SOD is not visible and therefore not significant ($\neg SOD$), this video serves as a useful stimulus for information and comparison to examine participant's commitment to engage with these kind of stimuli at all (see p. 26).

4.3.3 Data analysis.

To analyze the data, at first the results were displayed graphically. In a second step one-way repeated-measures ANOVAs were calculated to analyze the difference of perceivers' other-ratings and perceivers' self-ratings (SOD) at each time-period. By conducting two-way repeated-measures ANOVAs the interaction of SOD x Time was calculated. And as a last step, correlations (two-tailed) were calculated to measure the extend of the relation of $|$"EA"$|$ or $|SOD|$ with the applied questionnaires and tasks. For further information see section 2.5.

4.3.4 Exclusion of participants.

Two participants of the validation study had to be excluded. Because the validation of these videos uses the statistical norm, participants (TP) who were outlier had to be excluded. The affect rating data of TP 311 differed too much from the "norm". This behaviour was observed already at the rating dial practice-tasks (2 packages of pasta at the price of one was rated more positive than being in love). But the inconsistency of rating behaviour was also seen in the ratings of the videos in the validation study. And the other participant who had to be excluded was TP 330. This participant's rating behaviour was influenced by the realization of the non-truth of both incongruent videos. This participant saw through the experiment and experimental manipulation didn't succeed. Hence these both participants were excluded for further data analysis.

4.3.5 Targets' self-ratings.

On the basis of the graphs target's self-rating was analyzed. Figures 27–38 (p. 155–160) show target's self-ratings embedded in perceivers' other-ratings and perceivers' self-ratings. These figures revealed that target's self-ratings rarely reflected the emotional valence they displayed and therefore were not that suitable as an anchor for comparing participants' rating-data with them. It will be interesting, if the physiological data are a useful measure for comparisons. One aspect may be that the targets couldn't differ enough between the real emotions they had while discussing and the emotions they prevented they had. Maybe it didn't work because the instruction was not clear enough (a new instruction could be, "Please rate the emotions, you are displaying." Instead of, "Please rate the emotions you had while describing."). Or maybe it didn't work because the targets weren't trained and so they couldn't differentiate enough from the feelings they should display in the story and the own and real feelings. It is suggested to train target-persons' skill of self-monitoring and self-evaluation by self-awareness.

4.3.6 Consequences of section 4.3.5

Further calculations were conducted by using only perceiver's self-rating and perceiver's other-rating to examine SOD, $\neg SOD$, and the alternative so-called "EA". Hence ANOVAs were conducted to calculate SOD at incongruent videos by computing if perceivers' other-ratings differed significantly from perceivers' self-ratings. SOD at incongruent videos was calculated to measure if this video provides a useful stimulus to examine self-other distinction. And at congruent stimuli $\neg SOD$, as the match of perceivers' other-ratings and perceivers' self-ratings, was calculated to measure the utility of this video as a stimulus for information and comparison. Further SOD and "EA" were used to correlate videos' self-other distinction and "empathic accuracy" with questionnaires and tasks which are related to self-other distinction or empathy.

4.3.7 Discussion of the single stimuli.

From section 4.3.8 to 4.3.19 the single videos will be discussed.

4.3.8 G1 Lesly 4con – Entrance Exam / Aufnahmeprüfung FH (congruent, positive)

The content of this video is a positive and congruent life-event. Therefore it is expected that both ratings (other and self) are similarly positive in period 2 (the emotional period) and period 3 (the cognitive period). Signifying that both ratings (other and self) do not differ significantly from each other and both are in the positive continuum of the affect rating dial (displayed as a value above 0 at the y-axis at Figure 3).

As seen in Figure 3 (page 61), in this video the duration of the three periods is not well-balanced. The emotional period (period 2) from the onset-point on (that is the time-point when the message could be captured by the participant, 2–3 seconds after the target starts describing it) lasts only one second which is very short (one clear sentence). And the part of the story-introduction (period 1) takes over half a minute which is quite long.

One-way repeated-measures ANOVAs. Figure 3 (page 61) and the results of the ANOVAs showed that other-ratings and self-ratings differed significantly at period 1, but that there was a strong overlap at period 2 and period 3.

Although the period of the story-introduction (period 1) doesn't deal of any event with unexpected or disruptive emotions, the two ratings differed significantly from each other. The split starts at the onset-point when Lesly tells that she is, different to her siblings, more active in the technical field. The division lasts as long as the focus lies on construction engineering (&

technical field). It stops when the focus shifts to the feelings when Lesly tells that she passed the acceptation exam (at period 2, the emotional period). Ratings also matched when the focus lies on the explanation of this eminently feelings.

It seems that working in a technical field isn't very attractive to psychology students. Being instructed to focus on their own personality and rate their own emotional valence they would have when they were busy in a technical context, participants of this study rated their own emotional valence significantly more negative compared to the emotional valence of the targets.

Two-way repeated-measures ANOVAs.   Figure 4 revealed that the emotional valence (the own one and those which are ascribed to the target) got rated more positive with increasing time. The significant main effect of time underlined this observation. Contrasts revealed that each time-period's ratings differed significantly from the other ones. While at period 1 both ratings were in the negative area of the rating dial (mean of other-ratings $= -0.17$ and of self-ratings $= -0.43$), they ascended and at period 2 both were in the positive field around 0, and increased further as in period 3 they were both above 0.5 (see also Table 6). These results are in line with the positive valence of the video's story.

Nevertheless there was no significant main effect of type of rating (other vs. self), but there was a significant interaction effect of Time x Other/Self. Contrasts revealed that there were significant interaction effects when comparing other-ratings to self-ratings for period 2 compared to period 1 and for period 3 compared to period 1. But no significant interaction effect of other-ratings comparing to self-ratings for period 2 compared to period 3 was observed. This signifies that there was no significant difference between other-ratings and self-ratings in the emotional period compared to the cognitive period. In both periods the ratings matched in a similar way.

Summing up.   In this video other-ratings differed significantly from self-ratings in the period of the story-introduction (period 1) (period 1 is discussed apart at section 4.3.22, p. 123). But the two ratings matched in period 2 (salient event & emotional period) and in period 3 (reason & cognitive period). There was no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 3, signifying that in period 2 and period 3 both ratings overlapped similarly. Hence in these two periods self-other distinction is not observable, as anticipated. Due to the fact that other-ratings and self-ratings matched at period 2 and period 3, this video provides a useful stimulus for information and comparison to examine participant's commitment to engage with these kind of stimuli.

4.3.9   G1 Sora 6con – Boss Criticizes Me in front of a Colleague / Chefin kritisiert mich vor Kollegin (congruent, negative)

This video is about a negative but congruent story. Hence if other-ratings and self-ratings match at all periods of interest, this video would present a useful stimulus for information and comparison to examine participant's commitment and engagement into this kind of stimulus. Because the valence of the story is negative, it is expected that other-ratings and self-ratings are negative (between 0 and −1 at the rating-dial scale).

One-way repeated-measures ANOVAs. Results revealed that perceivers' other-ratings differed significantly from their self-ratings at every of the three time-periods of interest. When looking at Figure 5, it is seen that when perceivers feel with the target-person and rated Sora's emotional valence, these ratings were significantly more positive than when rating the own emotional valence at these periods (or the own emotional valence was rated significantly more negative than the target-person ones).

This story is a about a very emotional-loaded and negative life-event. When participants get instructed to identify with this situation, they rated their own emotional valence significantly more negative than the emotional valence they referred to the target-person.

Two-way repeated-measures ANOVAs. There was a significant main effect of time. Contrasts revealed that each time was rated significantly different to the other ones. As seen in Figure 6 and the means, which were all in the negative range of the rating dial (see Table 7), these results reflect that the video's content was a life-event of negative valence and successfully told as a negative one.

The significant main effect of rating (other/self) reflects the above mentioned significant differences of ratings at each time-period of interest.

There was no significant interaction effect between type of rating (other vs. self) and time-period. But when looking at the contrasts there was a significant interaction when comparing type of rating for period 2 compared to period 3. When looking at the graphic of the video (Figure 5), it is seen that the two ratings approximated in the cognitive period (period 3) compared to the emotional period (period 2). So in the emotional period, perceivers' rated their own emotional valence significantly more negative (or target's emotional valence significantly more positive) compared to the cognitive period.

Summing up. Perceivers rated their own emotional valence significantly more negative than the emotional valence they ascribed to the target-person as soon as the story about the negative life-event begun and the discrepancy continued further in the emotional and cognitive period. Anyhow at period 2, when the emotions are focused, participants rated the own emotional valence significantly more negative compared to the emotional valence of the target person when comparing to the cognitive period.

Two possibilities to interpret this results are suggested. One option concerns the perception of feelings. When participants put themselves in someone's position and concentrate on their own emotions, they feel the "own" emotions more intensely than when feeling with another person. In this video the content of the story is an emotional-loaded negative life-event. Participants rated the emotional valence when the situation would concern the own life significantly more negative than the emotional valence they ascribed to the target-person. Further they rated the own emotional valence significantly more negative when the focus lies on the emotions (emotional period/period 2) compared to the cognitive period (period 3).

The second possibility concerns the target-person and their emotional expressivity. Maybe Sora felt very negative while describing—Sora rated her own emotional valence even more negative than participants' did for themselves (see Figure 28 at p. 155)—but her emotional expressivity is just not that strong. But as documented at section 3.1 at Table **??** (p. **??**), Sora's emotional expressivity measured by the BEQ is 64% for the facette *Negative Expressivity* and 64% for *General Expressivity*. When looking at the FFEE the results are as followed: Expressivity is 73%; Darstellungsfähigkeit is 52%; Vorstellungsfähigkeit is 80%; Gefühlsklarheit is 100%; and erlebte Unfähigkeit, Gefühle zu vermitteln is 5%. Due to Darstellungsfähigkeit is just around 50%, this explanation can be considered too. As written in section 3.2 (or Table 25 at p. 163) it was mentioned by a participant of the pilot study that Sora describes emotional situations rather reserved compared to other targets[8]; e.g., Rudi, whose way of story-telling is very natural[9]. Further one more video of Sora (G3 Sora 5con) of positive valence shows the same pattern, in terms of that participants' self-ratings was significantly rated more positive than perceivers' other-ratings and Sora rated herself even more positive compared to the perceivers. But on the other hand, Sora's expressivity (with 64%) is quite high when compared to the other targets. Further it has to be considered that these questionnaires are self-rating questionnaires, meaning that not the ability itself is measured but the beliefs someone has about his/her abilities. E.g., Rudi's Expressivität (BEQ) is 54%, his Darstellungsfähigkeit (FFEE) is 67%,

---

[8]One participant said that it was difficult to asses Sora in general because her expression was so neutral.

[9]At the pilot study one participant meant that it doesn't appear that Rudi acted because of his dialect. Another participant stated that he is the most cool person because he is absolutely natural.

and erlebte Unfähigkeit, Gefühle zu vermitteln (FFEE) is 38%. And Rudi's videos don't show this effect of discrepancy.

Hence it is considered most likely that both explanations contribute to this result.

Due to the fact that other-ratings and self-ratings differed significantly from each other at all three periods this video does not represent a proper stimulus for information and comparison to examine perceiver's commitment to engage with these kind of stimuli.

4.3.10   G1 Rudi 3inc – Celebration of a Newborn / Kilias Kindsfeier (incongruent, positive)

This video is pursued to be an incongruent stimulus with a positive valence for perceivers. This implies that as soon as the salient event starts (period 2, the salient event & emotional period) and further at period 3 (reason & cognitive period), self-ratings should differ from other-ratings. Not only that these two ratings should discriminate in these periods, but also participants' self-ratings should be more positive than their other-ratings.

One-way repeated-measures ANOVAs.   Figure 7 (p. 67) and one-way repeated-measures ANOVAs revealed that at period 1 (the story-introduction) other-ratings did not differ significantly from self-ratings. Whereas at period 2 and period 3 type of ratings (other vs. self) differed significantly from each other, and self-ratings was more positive than other-ratings, as expected.

Compared to the two videos before (G1 Lesly 4con and G1 Sora 6con), in this video other-ratings and self-ratings overlapped at period 1. In this video period 1 (the story-introduction) focuses on a positive life-event concerning the life of the best friend ("the best friend became a father"). Results displayed that this event is generally positive, displayed by similar ratings when participants imagined themselves in this situation and rated their own emotional valence (self-ratings) as well as the emotional valence they referred to the target-person (other-ratings) (mean of both ratings was > 0.5). A consideration for the overlap is that the target-person discusses a situation in which the best friend became a father. So the event itself actually doesn't affect the target's own life nor would it actually concern the participant's own life, when their best friend becomes a father. Regardless of own life-style, wishes, preferences, and liking of children, this event belongs to another person's life—regardless of rating the own emotional valence when putting oneself in this situation (self-rating) or when feeling with the target-person and rate their emotional valence (other-rating), actually ratings always concern the best friend's life, not the own. It seems to result in a more homogeneous rating when the

(focused) story concerns another person's life.

As intended, at period 2 (the salient event & emotional period) and period 3 (reason & cognitive period) other-ratings differed significantly from self-ratings, moreover self-ratings were more positive than other-ratings. Figure 7 (p. 67) revealed that from the onset-point of the salient event on ("not being happy when the best friend becomes a father") other-ratings and self-ratings went apart. Different to period 1, in period 2 and 3 the focus of the discussed emotions and reason lie on the own person ("I'm not happy because..."); or rather the target's life when feeling with the target at other-rating, and the perceiver's own life when put themselves in this situation at self-rating. Therefore at these two periods actually the own life is concerned and the valence of the "own emotions" are rated.

Two-way repeated-measures ANOVAs.   The significant main effect and the contrasts of time revealed that the valence of the content of this video is changing significantly at every time-period compared to the other ones. Looking at the interaction graph (Figure 8 at p. 68) and Table 8 (p. 68), it is seen that the video's content got rated more negatively with increasing time.

The significant main effect of type of rating (other vs. self) goes in line with the intention of having a separation in rating-type (to make self-other distinction observable due to the fact that this is an incongruent stimulus).

There was also a significant interaction effect of Time x Other/Self. Contrasts revealed significant interactions when comparing type of rating (other vs. self) for period 2 (salient event & emotional period) compared to period 1 (story-introduction), as well as for period 3 (reason & cognitive period) compared to period 1. No significant interaction effect was observed when comparing type of rating for period 2 compared to period 3. Participants rated their own emotional valence significantly different to the emotional valence they ascribed to the target at these two periods. But within these two periods the difference remained similar. Regardless of knowing a reason (period 3) or not (period 2), participants' self-ratings were significantly different to their other-ratings—self-other distinction is observable in both periods.

Summing up.   While period 1 reflects a story-introduction of positive valence where other-ratings and self-ratings matched, at period 2 and period 3 self-ratings differed significantly from other-ratings. These results were intended due to the fact that this video is an incongruent stimuli, at which a differentiation between other and self is aimed. Further when comparing period 2 to period 3, it is seen that in both periods participants rated their own emotional

valence (self-ratings) significantly different to the target's emotional valence (other-ratings), but quite parallel. It seems that it is possible to feel with another person and rate the valence of their "special emotions" (other-rating) regardless of knowing the concrete reason of the target for them (period 3) or not (period 2). Further participants also achieved to remain by themselves and to rate their own emotional valence (self-rating) which are significantly different to the other's ones.

In this video self-other distinction is observable in the salient event & emotional period (period 2) and the reason & cognitive period (period 3), but not in the story-introduction (period 1). Hence this video presents an appropriate stimulus to examine self-other distinction at period 2 and period 3.

4.3.11 G1 Michi 1inc – Mouldy Gouda / Verschimmelter Gouda (incongruent, negative)

This video is about an incongruent and, for perceivers', negative story. Therefore it is expected that ratings of perceivers when rating the own emotional valence (self-ratings) compared to other-ratings (rating the emotional valence of the target) are significantly more negative at the salient event & emotional period (period 2) and the reason & cognitive period (period 3).

One-way repeated-measures ANOVAs. The one-way repeated-measures ANOVA revealed that other- and self-ratings didn't differ significantly at the first period. This indicates that the period of the story-introduction (period 1) was rated by participants nearly equally when rating the own emotional valence (self-rating) compared to when rating the target's emotional valence (other-rating). But type of rating (other vs. self) differed significantly at period 2, when the target-person reports the salient event, and the differentiation continued at period 3, when the target explains the behaviour and feelings. According to the hypotheses participants rated their own emotional valence significantly more negative compared to the emotional valence of the target-person in both periods, the emotional (period 2) and the cognitive period (period 3).

Two-way repeated-measures ANOVAs. The two-way repeated-measures ANOVA showed a significant main effect of time. Contrasts revealed that each time-period differed significantly from the other ones.

Also the main effect for type of rating (other/self) was significant.

The interaction effect of Time x Other/Self was significant too. Contrasts revealed signifi-

cant interactions comparing other-ratings to self-ratings for each time-period compared to the other ones (for period 2 compared to period 1, for period 3 compared to period 1, and for period 3 compared to period 2). When looking at the interaction graph (Figure 10 at p. 71) and Table 9 (p. 71) it is seen that while the other-ratings lowered at period 2 (mean $= -0.12$) compared to period 1 (mean $= -0.06$), it increased at period 3 (mean $= 0.28$). In contrast thereto self-ratings decreased more at each period compared to the periods before. The difference between the other-ratings and self-ratings increased at period 2 compared to period 1, and increased even more at period 3 (compared to period 2 and 1). These results revealed that perceivers rated their own emotional valence when imagining to be in the situation of eating a cheese full of mould significantly more negative compared to the emotional valence they ascribe to the target-person. At both periods, the emotional period and the cognitive period. The significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 3 and when looking at Figure 10, signifies that there is an increasing difference of the ratings in the cognitive period (period 3) compared to the emotional period (period 2). Figure 10 revealed that the target's emotional valence got rated more positive at period 3 compared to period 2, while perceivers' self-ratings remained negative. It seems that the target showed an increasing emotional valence at period 3 compared to period 2, hence self-other distinction is observable at both periods (period 2 and period 3), but self-other distinction increased at period 3 compared to period 2. Due to the fact that only target's emotional valence changed, I would not suggest to interpret that as knowing a reason leads to an increased self-other distinction. Rather participants differentiated themselves from the target's feelings in both situations and according to the instruction of feeling with the target and rate his emotional valence, other-ratings appear to reflect appropriate the emotional valence the target displays.

Summing up. In this video self-ratings were significantly more negative compared to other-ratings at period 2 (salient event & emotional period) and period 3 (reason & cognitive period), but ratings matched at period 1 (story-introduction). The significant interaction effect when comparing other-ratings to self-ratings for period 2 compared to period 3 is understood as the reflection of target's emotional valence which increased at period 3 compared to period 2 (rather than the fact of knowing a reason in period 3 compared to period 2).

In this video self-other distinction is observable at the last two periods, hence this video seems to be a useful stimulus to examine self-other distinction at period 2 (salient event & emotional period) and period 3 (reason & cognitive period).

4.3.12 G2 Rudi 3con – Celebration of a Newborn / Kilias Kindsfeier (congruent, positive)

This video displays a congruent stimulus with positive valence with the aim to examine empathic accuracy. According to the hypotheses it is expected that other-ratings and self-ratings match, i.e., there is no significant difference of ratings (other/self) at any time-period of interest.

One-way repeated-measures ANOVAs. Corresponding to the hypotheses, at none of the periods of interest (period 1, 2, and 3) self-ratings differed significantly from other-ratings. At each time-period these two ratings matched.

G1 Rudi 3inc and G2 Rudi 3con are two different videos who are about the same life-event. While G1 Rudi 3inc is an incongruent stimuli, and therefore the story has an unexpected and fictitious twist, in G2 Rudi 3con the story deals with the real life-event (congruent stimuli). Therefore in both videos, the story-introduction is about the best friend who became a father. Both ratings, other-rating and self-rating, concern the feelings the target-person or oneself would have when the best friend becomes a father. Hence ratings at period 1 are independent of the own wish of becoming a parent or the own life-style, the ratings are about being happy with and for someone else.

In the emotional (period 2) and cognitive period (period 3) the target-person describes the positive emotions and how happy he and everyone was. At these two periods the "own" emotions are addressed. Self-ratings and other-ratings matched here too.

Two-way repeated-measures ANOVAs. The significant main effect of time and contrasts revealed that each time-period differed significantly from the other ones. Looking at the interaction graph (Figure 12 at p. 74) and Table 10 (p. 74) it is seen that at each time-period both ratings were more positive compared to the period(s) before. This reflects the video's content, which is a story of positive valence.

There was no significant main effect of rating (other/self).

Furthermore there was no significant interaction effect of Time x Other/Self. Contrasts revealed no significant interaction term when comparing other-ratings to self-ratings for all periods compared to each other, for period 2 compared to period 1, for period 3 compared to period 1, and for period 2 compared to period 3. These results signify that the two ratings acted to each other parallel in every period compared to the other periods.

Summing up. This video is about a positive and congruent story. As intended, self-ratings did not differ significantly from other-ratings at any of the three time-periods of interest (= $\neg SOD$ at period 1 (the story-introduction), period 2 (the salient event & emotional period), and period 3 (the reason & cognitive period)). The positive valence of the story is well reflected by the significant main effect and contrasts of time. There was no significant main effect of type of rating (other/self), nor a significant interaction effect of Time x Other/Self. Contrasts of interaction effect revealed that participants rated the own emotional valence (self-ratings) similarly and not significantly different to the emotional valence they ascribed to the target-person (other-ratings) at all time-periods of interest.

Hence this video is a convenient stimulus for information and comparison to examine participant's commitment and engagement for this kind of stimulus at all three time-periods of interest.

4.3.13 G2 Michi 1con – Mouldy Gouda / verschimmelter Gouda (congruent, negative)

This video deals with the same story as G1 Michi 1inc, but this time the story is congruent, as it actually happened. In this video Michi discusses a negative life-event in a congruent way. Hence it is expected that at all time-periods of interest, participants' ratings when imagining to be in this situation and to rate their own emotional valence (self-ratings) matches with the ratings when feeling with the target and rate the target's emotional valence (other-ratings). Because the valence of the story is negative, it is anticipated that both ratings are rated negatively (between 0 and −1 at the scale of the rating dial).

One-way repeated-measures ANOVAs. The one-way repeated-measures ANOVA revealed that other-ratings differed significantly from self-ratings in period 1 (story-introduction) and period 3 (reason & cognitive period). At period 2 (salient event & emotional period) self-ratings and other-ratings do not differ significantly from each other.

When looking at Figure 13 (p. 76) it is seen that the split of the two ratings started already before the story begins. Although self-ratings and other-ratings approximated in the period of the story-introduction (period 1) it was still a significant difference between them.

At period 2 there was no significant difference of other-ratings compared to self-ratings, but a medium effect size was observable. Participants rated the situation of eating cheese which tastes tart equally when rating the own emotional valence (self-ratings) or when rating the emotional valence they refer Michi (other-ratings).

But self-ratings differed significantly from other-ratings at period 3, that is when Michi explains that when he took a look at the cheese the first time, he saw that on one side it was fully covered by mould. As seen in Figure 13 other-ratings and self-ratings dropped both at the beginning of period 3 and they both continued quite parallel. Anyway, eating mouldy cheese is rated significantly more negative when participants put themselves into the situation and rated their own emotional valence (self-ratings) compared to when they felt with Michi and rated his emotional valence (other-ratings) (see also Figure 13).

Two-way repeated-measures ANOVAs.   The significant main effect of time and contrasts revealed that the video was rated significantly more negative with increasing time. This reflects the negative valence of the story.

There was a significant main effect of type of rating (other/self). Meaning that other-ratings differed significantly from self-ratings. Figure 14 (p. 77) revealed that self-ratings were always significantly more negative than other-ratings.

There was also a significant interaction effect of Time x Other/Self. Contrasts revealed that there was no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 1. But there were significant interactions when comparing other-ratings to self-ratings for period 3 compared to period 1 and for period 2 compared to period 3.

Summing up.   This video's story is congruent and therefore it is expected that other-ratings do not differ significantly from self-ratings. But rating-type didn't match at period 1 and period 3. Maybe disgust is a too strong emotion and experience which distresses. In any case, due to the fact that other-ratings and self-ratings differed significantly at period 1 and period 3 (self-other distinction was observable), this video does not constitute a proper stimulus for information and comparison to examine participant's commitment and ability to engage with this kind of stimulus (a stimulus for information and comparison aims $\neg SOD$, meaning that SOD is not visible).

4.3.14   G2 Sora 2 inc – Become an Aunt / Tante werden (incongruent, positive)

This video is aimed to be an incongruent stimulus. Therefore a positive life-event was told in an incongruent way (Sora tells that it was negative for her), so that self-other distinction gets observable. The distinction between other-ratings and self-ratings is expected at period 2 and period 3.

One-way repeated-measures ANOVAs.   The one-way repeated-measures ANOVA revealed that other-ratings differed significantly from self-ratings at all time-periods of interest (period 1, 2, and 3).

Already in the period of the story-introduction (period 1) the two ratings (other vs. self) differed significantly from each other (see also Figure 15 at p. 79). When participants watched Sora reporting the situation of phoning with a very close person who comes out becoming a mother, participants' self-ratings (when they put themselves in this situation and rated their own emotional valence) was significantly more positive than participants' other-ratings (when they felt with Sora and rated her emotional valence). In period 2, the salient event & emotional period, Sora tells that she couldn't be happy for her friend. And in period 3, the reason & cognitive period, Sora explains her reason for these feelings. As expected, in both periods other-ratings differed significantly from self-ratings.

Two-way repeated-measures ANOVAs.   There was a main effect of time. Contrasts revealed that each time-period of interest differed significantly from the other ones. Ratings were more negative at each time-period compared to the other periods before.

The main effect of type of rating (other vs. self) and Figure 16 (p. 80) revealed that self-ratings were significantly more positive compared to other-ratings. These results go in line with the valence of the story, which is positive for perceivers.

There was a significant interaction effect of Time x Other/Self. Contrasts revealed significant interactions when comparing other-ratings to self-ratings for period 2 compared to period 1, and for period 3 compared to period 1. Looking at Figure 16, it is seen that, compared to the story-introduction (period 1), at the emotional period (period 2) and the cognitive period (period 3) the distance between other-ratings and self-ratings was more. But there was no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 3. In both periods, the emotional period (period 2) and the cognitive period (period 3), participants' ratings for the own emotional valence (self-ratings) differed in a parallel way from participants' ratings of the target-person's emotional valence (other-ratings). Regardless of knowing Sora's reason why she felt that way (period 3) or not knowing it (period 2), at both periods participants rated their own emotional valence they wold have in this situation, similar more positive compared to the emotional valence they refer the target-person.

Summing up.   This video deals with an incongruent story and for perceivers' it is of positive valence. Self-ratings were significantly more positive compared to other-ratings at period 1, 2,

and 3, signifying that self-other distinction is observable at all three time-periods of interest. The significant difference of other/self at period 1 (the story-introduction) was unexpected. It seems that the period of the story-introduction is not that "neutral" as planned. The non-significant interaction effect when comparing other-ratings to self-ratings for period 2 (salient event & emotional period) compared to period 3 (reason & cognitive period) implies that participants' self-other distinction is observable in a similar degree in both periods. It seems that participants' didn't need to know Sora's specific reason for her uncommon feelings to be able to feel with her and rate the valence of her emotions significantly more negative compared to the valence of the emotions they would have themselves.

Hence this video is a suitable stimulus to examine self-other distinction at period 2 (salient event & emotional period) and period 3 (reason & cognitive period).

4.3.15   G2 Olcay 5inc – Death of Grandfather / Tod vom Opa (incongruent, negative)

This videos' story is incongruent and about the death of a beloved person. The incongruity is due to the valence of the story, which is negative for perceivers but Olcay tells the story as positive for her. Because it is told incongruently, it is expected that there is a significant difference for participants' self-ratings (when participants put themselves into this situation and rated their own emotional valence) compared to their other-ratings (when participants felt with Olcay (the target-person) and rated her emotional valence) from the salient event (period 2) on. Hence self-other distinction is observable at period 2 (the salient event & emotional period) and at period 3 (the reason & cognitive period).

One-way repeated-measures ANOVAs. Results showed that in period 1 (the story-introduction) other-ratings differed significantly from self-ratings. Figure 17 (p. 82) revealed that the two ratings ran apart as soon as the story begun. Participants' self-ratings were already at this time-period significantly more negative than their other-ratings (or perceivers' other-ratings were significantly more positive compared to their self-ratings). Maybe the death of a beloved person can't be reported as "neutral"; "neutral" in terms of other-ratings and self-ratings match. Especially when participants had seen the video once and knew Olcay's reaction and emotions which are uncommon and different to the own emotions one would have (in terms of a bias for feeling (and rating) the own emotions in the second cycle).

As expected, at period 2 (salient event & emotional period) and period 3 (reason & cognitive period) other-ratings differed significantly from self-ratings. In both periods participants'

ratings of the own emotional valence (self-ratings) was significantly more negative than their ratings of the target's emotional valence (other-ratings).

Two-way repeated-measures ANOVAs. There was a significant main effect of time. Contrasts revealed that there was no significant difference comparing period 2 to 1, but there was a significant effect for period 3 compared to period 1 and for period 2 compared to period 3. The main effect of rating-type (other vs. self) was also significant. Figure 18 (p. 83) revealed that self-ratings was lower than other-ratings.

There was also a significant interaction effect. Contrasts revealed a significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 1, and for period 3 compared to period 1, but not for period 3 compared to period 2.

The interaction graph (Figure 18, p. 83) revealed that the difference between other-ratings and self-ratings was the smallest at period 1, compared to period 2 and period 3. Although type of rating differed significantly at period 1. At period 2 and period 3 the difference between other-ratings and self-ratings were similar. It seems that participants felt with Olcay and distanced from her in a similar way, irrespective of knowing her reason (period 3) or not (period 2).

Summing up. This video treats of Olcay's story when her beloved grandfather died. It is an incongruent stimulus of negative valence for perceivers with the aim to make self-other distinction observable. Results revealed that at all time-periods (period 1, 2, and 3) perceivers' self-ratings were significantly more negative compared to their other-ratings—SOD is observable at all three time-periods. It seems that in this story, participants were able to feel with Olcay and to rate her emotional valence (other-ratings), which were significantly more positive than their own one (self-ratings). When Olcay explains her reason for the uncommon feelings, participants' self-ratings rose (from -0.50 at period 2 to -0.16 at period 3, see Table 13 at p. 83), and as seen at Figure 17 it showed a tendency towards positive, parallel to perceivers' other-ratings. Anyway, the maximum of the difference between perceivers' other-ratings and their self-ratings were at the end of period 2 (as seen in Figure 17).

Self-ratings were significantly more negative compared to other-ratings at the story-introduction (period 1), the salient event & emotional period (period 2), and at the reason & cognitive period (period 3). Hence this video represents a useful stimulus to examine self-other distinction at period 1, period 2, and period 3.

4.3.16 G3 Sora 5con − 1st Boyfriend / 1. Freund (congruent, positive)

This video's story is about a positive and congruent life-event. Hence if self-ratings do not differ significantly from other-ratings at the time-periods of interest (period 1, 2, and 3), this video would provide a useful stimulus for information and comparison (to test if a participant is able to engage with this kind of stimulus and story). Further because the story's valence is positive, it is expected that both ratings (other-rating and self-rating) are in the positive continuum of the scale of the rating dial (between 0 and +1).

One-way repeated-measures ANOVAs. Results revealed that other-ratings differed significantly from self-ratings at all three time-periods, the story-introduction (period 1), the salient event & emotional period (period 2), and the reason & cognitive period (period 3). Although other-ratings and self-ratings were very close (see Figure 19 at p. 19), they differed significantly from each other. These results are unwanted since this video was thought as a congruent stimulus where self-ratings and other-ratings should match (SOD should not be observable ($\neg SOD$)).

Two-way repeated-measures ANOVAs. The significant main effect of time and contrasts revealed that the video got rated differently by every time-period compared to the other time-periods. Figure 20 (p. 86) revealed that both ratings, other and self, got more positive with increasing time. This reflects the positive valence of the story.

The main effect of type of rating (other/self) was not significant. Looking at the interaction graph (Figure 20) it is seen that both ratings were quite close in general.

There was a significant interaction effect of Time x Other/Self. Contrasts revealed a significant interaction when comparing self-ratings to other-ratings for period 2 (salient event & emotional period) compared to period 1 (story-introduction) and for period 3 (reason & cognitive period) compared to period 1, but not for period 2 compared to period 3. Participants' self-ratings (when they put themselves in this situation and rated their own emotional valence) and participants' other-ratings (when they felt with Sora and rated her emotional valence) differed significantly at period 2 and period 3, but within these two periods type of rating behaved to each other in a parallel way.

Summing up. Although the story of this video is rated clearly positive (means of other-ratings and self-ratings were between +0.20 and +0.77, see Table 14 at p. 86), self-ratings differed significantly from other-ratings at all three time-periods of interest (period 1, 2, and

3). Maybe having the first boyfriend means different for different participants. Or another explanation for these results could be that the way of reporting the story and/or target's expressions were not clear enough. As already mentioned in section 4.3.9 (G1 Sora 6con, p. 104), Sora's expressivity might be not that clear to participants. Although she rated herself at self-rating questionnaires (BEQ and FFEE) quite high in expressivity (64% at BEQ's *General Expressivity* and 73% at FFEE's *Expressivity*, see therefore Table **??** (p. **??**) at section 3.1), a participant of the pilot study declared that she found Sora's way of expressing not clear (see Table 25, p. 163 at section 3.2).

However these results showed that self-other distinction was visible (SOD was significant) at all three time-periods of interest (period 1, 2, and 3). This is not intended for a stimulus for information and comparison, where the aim is that self-other distinction is not observable ($\neg SOD$). Hence this video is not a proper stimulus for information and comparison to examine participants' commitment and ability to engage with this kind of stimulus for further examinations in this field (e.g., with incongruent stimuli).

4.3.17 G3 Michi 3con – Dad Drank Away my Apple-Juice / Papa trinkt Apfelsaft weg (congruent, negative)

The story of this video is about a negative life-event of Michi, told in a congruent way (the way the event actually happened). Hence it is expected that participants' other-ratings (when participants feel with Michi and rate his emotional valence) and their self-ratings (when participants put themselves into this situation and rate their own emotional valence) do not differ significantly from each other, i.e., type of rating matches. Meaning SOD is not observable ($\neg SOD$). Due to the story's valence is negative, it is further expected that at period 2 (salient event & emotional period) and period 3 (reason & cognitive period) both ratings were in the negative spectrum of the rating dial's scale (i.e., between 0 and -1).

One-way repeated-measures ANOVAs. The ANOVA showed that other-ratings did not differ significantly from self-ratings at the period of the story-introduction (period 1) and period 3 (reason & cognitive period), but at period 2 (salient event & emotional period). This video was intended as a congruent stimulus, and therefore the ratings should match (SOD should not be observable). Hence the significant difference of other-ratings and self-ratings at period 2 is not conform to the hypotheses for a useful stimulus for information and comparison.

Two-way repeated-measures ANOVAs. The main effect of time and contrasts revealed that each time-period's ratings (at period 1, 2, and 3) differed significantly from the ratings of the other time-periods. According to the negative valence of the story, Figure 22 indicates that ratings were more negative at each time-period compared to the periods before. Besides in period 2 and period 3, both ratings were in the negative area of the scale of the rating dial (means were between -0.16 and -0.59, see Table 15 at p. 89).

Although the main effect of type of rating (other/self) was not significant, it revealed a medium effect size.

There was a significant interaction effect of Time x Other/Self. Contrasts revealed a significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 1 and for period 2 compared to period 3. No significant interaction was observed when comparing other-ratings to self-ratings for period 3 compared to period 1. Figure 22 it is seen that other-ratings and self-ratings were quite close to each other in period 1 and 3 (which is reflected by the non-significant effects of the one-way ANOVA as written above in these periods). But ratings differed significantly at period 2, where participants' self-ratings were significantly more negative compared to their other-ratings.

Summing up. This videos was assumed as a congruent stimulus with negative valence. The negative valence is well reflected by the significant effects of time and Figure 22. But results revealed that other-ratings differed significantly from self-ratings at period 2. At congruent stimuli it is aimed that self-other distinction is not observable ($\neg SOD$). Hence this leads to the consequence to reject this video as a useful stimulus for information and comparison.

4.3.18 G3 Rudi 5inc – Girlfriend Dumped Me / Freundin macht Schluß (incongruent, negative)

This video is about a negative life-event but Rudi tells it in an incongruent way, i.e., the content is actually (and therefore for perceivers) of negative valence but Rudi reports that the event was positive for him. The incongruity starts at period 2 (salient event & emotional period). Hence it is expected that at period 2 and period 3 (reason & cognitive period) participants' self-ratings (when participants put themselves into this situation and rate their own emotional valence) are significantly more negative compared to participants' other-ratings (when participants feel with Rudi and rate his emotional valence).

One-way repeated-measures ANOVAs. Looking at Figure 23 (p. 91), other-ratings and self-ratings matched at period 1 but differed a lot at period 2 and 3. These is statistically reflected by the non-significant effect when comparing participants' other-ratings to their self-ratings at period 1, but the significant effects at period 2 and period 3. At period 2 (salient event & emotional period) and period 3 (reason & cognitive period) participants' self-ratings were significantly more negative compared to participants' other-ratings.

Two-way repeated-measures ANOVAs. The significant main effect of time and contrasts revealed that ratings at each time-period differed significantly to them at the other time-periods. Looking at Figure 24 it is observable that ratings at period 1 were positive (means of both, other-ratings and self-ratings, were above $+0.5$, see Table 16, p. 92), than they lowered at period 2 (mean of other-ratings $= -0.01$, mean of self-ratings $= -0.64$) and at period 3 they rose (mean of other-ratings $= +0.34$, mean of self-ratings $= -0.07$). That ratings lowered at period 2 compared to period 1 goes in line with the negative valence of the story.

There was a significant main effect of type of rating (other/self). Figure 24 revealed that while other-ratings and self-ratings matched at period 1, they differed at period 2 and 3. Participants' self-ratings were more negative compared to their other-ratings.

Further the interaction effect of Time x Other/Self was significant. Contrasts revealed a significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 1, and for period 3 compared to period 1. But there was no significant interaction when comparing other-ratings to self-ratings for period 2 compared to period 3. In both periods, period 2 (salient event & emotional period) and period 3 (reason & cognitive period), participants rated the target's emotional valence (other-ratings) in a similar way differently to their own emotional valence (self-ratings).

Summing up. The story of this video is a negative life-event, told in a way one commonly wouldn't expect (incongruent). Hence it was expected that at period 2 (where the incongruity starts) and period 3 self-other distinction is observable, represented by a significant difference between perceivers' self-ratings (when participants put themselves in this situation and rated their own emotional valence) and their other-ratings (when participants felt with Rudi and rated his emotional valence). And due to the negative valence of the story, it was expected that perceivers' self-ratings are more negative compared to their other-ratings.

Results revealed that perceivers' self-ratings and their other-ratings matched at period 1 (story-introduction). But at period 2 (salient event & emotional period) and period 3 (reason

& cognitive period), perceivers' self-ratings were significantly more negative compared to their other-ratings.

Statistically there is no significant interaction effect when comparing other-ratings to self-ratings for period 2 compared to period 3, but this contrast yielded a medium to large effect size. As seen in Figure 23 (p. 91) at the reason & cognitive period (period 3), participants' self-ratings approximated to their other-ratings. When Rudi explains his reason for this specific uncommon emotions (i.e., feeling good when the love of the life dumps someone), perceivers rated their own emotional valence increasingly more positive. It seems that a "good" explanation for feelings which are different to the own ones, can lead to an approximation of the own feelings towards the feelings of the other person. Although perceivers' self-ratings approximated to their other-ratings, the difference remained significant.

In any case self-other distinction was visible (SOD was significant) at period 2 and period 3. Hence this video represents a useful stimulus to examine self-other distinction at period 2 (salient event & emotional period) and period 3 (reason & cognitive period).

4.3.19   G3 Stephan 3inc – Banana Cake / Bananenschnitte (incongruent, negative)

This video is aimed to be an incongruent stimuli of negative valence for perceivers. Therefore a difference between other- and self-ratings is expected especially in period 2 and period 3. At period 2, which gets opened by the salient event, the target's emotions are focused. At period 3, a reason for the uncommon emotions is given and hence this period depict a cognitive period with explanations.

Figure 25 (p. 94) reveals that after the initial seconds of the salient event (the first onset-point is at second 50), period 2 is interrupted by a "congruent-part" (chocolate) in the middle of the incongruent narration (mayonnaise and pickled gherkin). This is well reflected by the self-ratings of second 51–55 where they approximated to the other-ratings. Therefore period 2 is counted from second 61 on.

At Figure 25 it is also observable that participants' self-ratings started to decrease and to differ from their other-ratings at second 46. This is very interesting due to the fact that the starting-point of the salient event (that is the point of time when Stephan *starts* to narrate the salient event) lies between second 47 and 48. Furthermore the onset-point of this salient event is determined at second 50 (the onset-point is always defined 2–3 seconds after the starting-point, reflecting the time perceivers need to catch the message). It has to be considered that the self-rating was always the second cycle. So participants already watched the video once and became acquainted with the story. Because they knew what will happen, they were prepared.

Because of this reason it seems that the self-ratings were quicker at changes compared to the other-ratings, which was done in the first run.

One-way repeated-measures ANOVAs.   Looking at Figure 25 at page 94, other-ratings and self-ratings matched at period 1 but differed a lot at period 2 and 3. These is statistically reflected by a non-significant effect when comparing other-ratings to self-ratings at period 1, but by significant effects at period 2 and period 3.

Two-way repeated-measures ANOVAs.   The significant main effect of time and contrasts revealed that each time-period's ratings differed significantly from the other ones. Looking at Figure 26 (page 95) it is observable that both data of the rating dial (other and self) at period 1 were around 0.5, which is about the middle of the positive valence of the affect rating dial. Other-ratings remained at this level for period 2 and 3. Contrary thereto self-ratings lowered from $+0.47$ to -0.44 at period 2 and rose at period 3 to -0.15. Due to this video's content which is a story of negative valence (for perceivers), the findings are in line with the intention.

The significant main effect of other/self was observed in period 2 and period 3. As written above, at period 1 other-ratings and self-ratings matched, while at period 2 and period 3 the ratings differed significantly from each other. At the period of the story-introduction (period 1) participants rated their own emotional valence when they put themselves in the place of Stephan's birthday and getting a cake, similar positive as the emotional valence they found the target-person displayed. At period 2 and 3 the own emotional valence (self-ratings) differed significantly from the emotional valence which were ascribed to the target-person (other-ratings).

The significant interaction effect of Time x Other/Self and the contrasts revealed significant interactions when comparing other-ratings to self-ratings for all periods compared to the other periods (for period 2 compared to period 1, for period 3 compared to period 1, and for period 3 compared to period 2). This signifies that other-ratings and self-ratings related to each other significantly different in every time-period. While at period 1 the ratings didn't differ significantly from each other, at period 2 and period 3 they differed significantly. Figure 26 reveals that at period 2 the difference was the maximum and at period 3 self-ratings rose and approximated to other-ratings, although they still remained significantly different. These findings indicate that participants rated their own emotional valence (self-ratings) significantly different from the target ones (other-ratings) in the period of the salient event (period 2) and the cognitive period (period 3). Although the difference between the ratings were significantly

higher in period 2, with a distance of 0.90 points between the two ratings (at the y-axis) (reflecting 45% of the maximal achievable distance by the rating dial's range), compared to period 3, where the distance was 0.64 points of the rating scale (reflecting 32% of the maximal distance).

Summing up.   These findings indicate that participants rated their own emotional valence significantly more negative compared to when participants rated the target's emotional valence for both, the emotional period (period 2) and the cognitive period (period 3), but the ratings matched for the period of the story-introduction (period 1). The significant effects of time reflects that the video is emotionally loaded. The content is of negative valence, which is well reflected by the self-ratings in period 2, the period where the emotions are focused. In period 3 when the target explains and gives a reason the self-ratings rose significantly, compared to period 2. It seems that when participants get a good explanation why they could like this cake like Stephan does, they approximate to the target. Self-other distinction is significantly reduced in the cognitive period (period 3) compared to the emotional period (period 2).

Hence this video is an appropriate stimulus to examine self-other distinction at period 2 and period 3.

4.3.20   Bottom-up vs. top-down.

Regarding the question of the role of bottom-up and top-down processes for empathy (see p. 18), the following observations were made. Assuming the visible self-other distinction at incongruent videos is connected with perceivers' ability to empathize, perceivers were able to empathize with the target-person at period 2 and period 3. This means that participants empathized with the target-person although not knowing the target's reason (which was given in period 3). It is assumed that participants have automatically an explanation for the target's feelings, based on their own life-experiences. This explanation must not match the target's explanation and can be adjusted and corrected as soon as knowing the target's reason. The proof of this assumption would need further qualitative research.

4.3.21   The balance of period 1, period 2, and period 3.

Especially at incongruent videos it was considered to divide the story into three parts. These three periods were not considered enough at congruent videos. Hence it happened that e.g., at G1 Lesly 4con the three periods were not well-balanced and period 2 lasts only one second. A more careful consideration concerning the duration of these periods while developing the

videos, would help to have more uniform periods to compare them afterwards to each other and between the conditions (congruent and incongruent).

4.3.22 Period 1 (the story-introduction).

Another point which didn't work as expected concerns period 1, the period of the story-introduction. It was thought that this period could provide a control-period with a "neutral introduction into the story of the life-event" where perceivers' other-ratings and perceivers' self-ratings match. But only in 6 of 12 videos in this period perceivers' other-ratings matched perceivers' self-ratings.

In the following videos perceivers' other-ratings matched perceivers' self-ratings at period 1 (plus period 1 is shortly characterized):

- G1 Rudi 3inc: His best friend became a father.
- G1 Michi 1inc: Michi gripped a package of cheese and bit into it.
- G2 Rudi 3con: His best friend became a father.
- G3 Michi 3con: The father drank away the apple juice which belonged to Michi.
- G3 Rudi 5inc: The girlfriend, the love of his life, dumped him from one day to the other.
- G3 Stephan 3inc: Having birthday and getting a banana cake.

In this videos perceivers' other-ratings differed significantly from perceivers' self-ratings (plus period 1 is shortly characterized):

- G1 Lesly 4con: Being interested in the technical field / construction engineering.
- G1 Sora 6con: The boss criticized her in front of a colleague.
- G2 Michi 1con: Coming home late, eat something without looking.
- G2 Sora 2inc: A close person becomes a mother.
- G2 Olcay 5inc: Death of a beloved person.
- G3 Sora 5con: 1st boyfriend at the age of 18, which Sora finds quite late.

It seems that there doesn't exist something as a "neutral situation" in which everybody has the same emotions and is therefore emotional valence is rated equally. Every person prefers different life-styles, countries, surroundings, has different interests (e.g., technical field), etc. The instruction to concentrate on their own feelings they have when they put themselves in this situation, explicitly asks for participant's interests and emotions. Therefore when a participant prefers something different in comparison to the target person, the ratings (other vs. self) do not match although nothing "special" happens in the story of this period.

Hence period 1 (or any other period with any content that can be liked differently) is not a useful period for information and comparison. Or in more general terms it has to be considered

that a "neutral situation" is not really achievable. If a period for information and comparison is desired, a possibility might be the loophole that the target-person describes a story of a life-event of a third person. This would provide rating of emotional valence of a situation in which participant's own life, life-plan, wishes, and preferences are not asked directly. So irrespectively if the instruction is to feel with the target-person and rate his/her emotional valence (perceivers' other-ratings) or when perceivers are asked to put themselves in this situation and rate their own emotional valence (perceivers' self-ratings), the life-event wouldn't actually affect their personal life (as in G1 Rudi 3inc and G2 Rudi 3con, where Rudi's best friend became a father).

4.3.23 Congruent condition.

The reason described in section 4.3.22 might be the same reason why only two of six congruent videos reached the aim that perceivers' other-ratings did not differ significantly from perceivers' self-ratings. It was rather an exception than the rule. It seems almost impossible to find situations which are felt similarly by different persons, so that there is a match when rating the emotional valence of another person (perceivers' other-ratings) and when they put themselves in this situation and rate their own emotional valence (perceivers' self-ratings).

To find situations which evokes that similar emotions that ratings don't differ significantly from each other seems very challenging due to the individuality of participants.

Hence congruent videos as stimuli for information and comparison are hardly working due to individuality and individual preferences.

4.3.24 Correlation of the questionnaires/tasks with $|SOD|$ and $|"EA"|$.

All significant correlations of questionnaires and tasks with $|SOD|$ are summarized in Table 26 (p. 164). All significant correlations of $|"EA"|$ and the applied questionnaires and tasks are presented in Table 27 (p. 165).

As seen in both, Table 26 (p. 164) and Table 27 (p. 165), there was no uniform answer-pattern and often the direction of the correlation was opposed to the expected one. Two possible reasons will be given. The first reason concerns the questionnaires itself. When assumed that participants filled out the questionnaires in all conscience, it can be considered that the questions of these self-rating questionnaires might be not precisely enough and therefore for answering them exactly there is too much different interpretation possible. Further self-rating questionnaires do not measure the actual ability rather the belief of a person of this ability. But the non-stringent pattern of the correlations with the Finger Lifting Task (Imitation Inhibition Task) and Shelve Task (Perspective Taking Task) cannot be explained with this argument. The

second reason concerns only empathic accuracy. As written before and at page 25, in this study the definition of empathic accuracy is based on an ad hoc hypothesis and therefore named a so-called "empathic accuracy" ("EA"). The ad hoc hypothesis uses a plausibility-argument and is not tested. Hence interpreting this "empathic accuracy" as a kind of empathy is vague and results have to be interpreted carefully.

4.3.25 Post Video Questionnaire (PVQ).

After watching and rating the four videos, participants were asked to fill out a post video questionnaire (PVQ) (the PVQ is presented at p. 152 and the results of the PVQ in Table 28 at p. 166). The questions of the PVQ concerned how *(1.) interesting*, *(2.) attractive*, and *(3.) likebale* the perceivers found the target-person. The physical attractiveness was asked because Ickes et al. (1990) demonstrated in a study that partner's physical attractiveness correlated positively with content accuracy. It was expected that all three the target-person concerning questions correlated positively[10] with the redefined "empathic accuracy" (see therefore p. 25). The questions *(4.) the wish to distance* and *(5.) if participants zoned out* were expected to correlate negatively and *(6.) if the perceiver was once in a similar situation* was expected to correlate positively with "empathic accuracy". As seen in Table 27 (p. 165), only at two videos, G1 Michi 1inc and G2 Rudi 4con, results showed significant correlations with expected directions. G1 Rudi 3inc and G3 Rudi 5inc showed correlations with direction contra to the expectation. The three other significant correlations were contra the expected direction and belong to videos which are assumed as no appropriate stimuli.

The results of the correlation of the PVQ with self-other distinction (for the operationalization of self-other distinction see p. 22) are presented at Table 26 (p. 164). Concerning self-other distinction the two questions *(4.) the wish to distance* and *(5.) if participants zoned out* were hypothesized to correlate positively with self-other distinction. But as seen in Table 26 the answer pattern not only at these two questions, but at all questions of the PVQ regarding self-other distinction was quite inconsistent.

In general the correlations of the questions of the PVQ with self-other distinction or "empathic accuracy" showed an inconsistent answer-pattern. It is open for future research what exactly lead to this inconsistency. It might be the ad hoc operationalization of "empathic accuracy". But this argument doesn't give an answer to the even more contradicting results of the correlation of the PVQ with self-other distinction.

---

[10]A relation was displayed by a negative correlation (see also p. 25).

4.3.26   Plausibility of the videos as stated by participants of the validation study.

At the end of the validation study, participants were asked if any video was a bit unusual for them or all four videos rather common. As seen in Table 18 (p. 97), five videos were mentioned as "not true" at least once by a participant of the validation study. The videos with the most no-true-stating were G1 Michi 1inc, G2 Olcay 5inc, G3 Rudi 5inc, and G3 Stephan 3inc. These four videos had also the highest scores at the item *(4.) want to distance* at the Validation-Study's Post Video Questionnaire (PVQ) (see Table 28, p. 166). While G1 Michi 1inc and G3 Stephan 3inc deal with a story in which the target-persons pretend that something is tasty what most would find disgusting, the other two videos (G2 Olcay 5inc and G3 Rudi 5inc) were concerned with a story in which a beloved person died or dumped someone. It seems that these two topics, disgust and the loss of a beloved person, mostly evokes perceiver's attention regarding the plausibility of a story. Maybe because especially these stories (to be happy when a beloved person dies or dumps someone) constitute a moral problem.

Interesting is that although some participants mentioned that they didn't believe these videos, mostly this opinion did not interfere the ability to empathize or to sense the own emotions, reflected by the affect rating data. From a total of 63 participants, only one participant (TP 330) had to be excluded because both incongruent videos were detected as not true and ratings of this participant varied noticeable from the other participants.

As seen, the lack of knowledge of a story's truthfulness does not necessarily interfere with the ability to empathize. This argument is also underlined by a statement of a participant (TP 17) of the pilot study, who pointed out that he is quite unsure if the stories were true. And further he said that this fact doesn't affect the ability to feel with the person in the video (see Table 25, p. 163). Hence it seems that the plausibility of another person's story is no prerequisite to be empathic, but the willingness to empathize with another person is a necessary requirement.

4.3.27   Emotions displayed by perceivers at the validation study.

When looking at the graphs of perceivers' other-ratings and perceivers' self-ratings, presented in the Results section of the validation study (section 3.3, p. 59ff.), it is quite noticeable that the affect ratings show some kind of emotional flatness and further some way of emotional contagion. To explain this kind of emotional contagion, an example will be given. In the video G1 Rudi 3inc (see Figure 7, p. 67), Rudi describes an actual positive story with a negative valence. At the two pivotal periods, period 2 and period 3, perceivers' self-ratings were more positive relative to perceivers' other-ratings, but in total with a mean of $M = 0.06$ at period 2

and $M = -0.15$ at period 3, perceivers' self-ratings were not rated clearly in the positive continuum of the rating dial's scale (which would be 0 to +1). Although the described event is a positive life-event, participants rated their emotional valence they would have when being in this situation not positive. Especially this video is good for explaining this point due to its counterpart—the video G2 Rudi 3con, which has the same story but told in a congruent way (see Figure 11, p. 73). There perceivers' self-ratings as well as perceivers' other-ratings were rated clearly positive (means at period 2, $M = 0.64/M = 0.69$, and at period 3 both, $M = 0.79$). It seems that perceivers let themselves influence by the emotions the target describes/displays. Hence it seems that the own emotions in a situation are affected by the emotions which are described by another person in this situation.

In regard to emotional flatness it is seen that although this event (described in the video G2 Rudi 3con) treats of a very positive event (for Rudi is was one of the most positive life-events he remembered when listing his six most positive life-events), perceivers' other-ratings as well as perceivers' self-ratings did not exceed 0.80 of the rating dial's scale. Other speaking the most positive ratings in this video were only 80% of the total possible positive affect, so only a part of the rating scale's range was used. Further research would be needed to reveal the reason therefor, but the following factors may influence those two points. It might be due to the instruction of empathizing with an unknown person (May I trust him/her?, Who is she/he?), or that the stimuli are videos (trying to be empathic without having a dialog), or because of the artificial context (laboratory) and that some videos are strange (incongruent condition).

4.3.28 Affect rating dial.

The affect rating dial seems to be a useful instrument to enable a visualization of emotional valence. It seems that the possible continuously and infinitely adjustment was used proper by the participants. Anyway, I observed that the four rating dial practice tasks, described at p. 45, were very useful and even necessary. With these practice tasks not only the handling was exercised, but also that the continuous use of the rating dial was not forgotten. Hence when using the rating dial in further experiments, it is recommended to undertake those or similar practice tasks.

4.3.29 Résumé of Discussion

In general some points were noticed and will be summed up now.

The balance of period 1, 2, and 3. Especially in congruent videos a more careful consideration of the duration of each period would provide a better basis for a comparison of the periods to each other and between the conditions (incongruent and congruent).

Period 1 (the story-introduction). Only at half of the videos at period 1 (the story-introduction) perceivers' self-ratings did not differ significantly from perceivers' other-ratings. So at half of the videos self-other distinction was visible in this period, although it was aimed to have a "neutral" introduction where perceivers' self-ratings and their other-ratings match. It seems that there doesn't exist such a "neutral" situation which evokes similar emotions in everyone (or rather evokes emotions that are that similar so that the affect ratings of different people do not differ significantly from each other).

Congruent condition. The reason which lead to the significant differences in perceivers' other-ratings and perceivers' self-ratings at period 1 might be the same reason why only two of six congruent videos revealed a match of perceivers' other-ratings and their self-ratings (at Table 20, p. 130 an overview is given at which videos a match was found).

Target's self-rating. Another topic which has to be considered are targets' self-ratings. Figures 27 to Figure 38 (at page 155–160) revealed that target's self-rating rarely reflected the emotional valence the target was describing. Hence targets' self-ratings were no suitable basis to compare perceivers' other-ratings with it and to examine empathic accuracy (see also the discussion at section 4.3.5). On the one hand a new instruction should be considered, e.g., "Please rate the emotions you are displaying." On the other hand targets were not trained in sensing their own emotions. Hence a training in self-awareness might help to achieve affect ratings displaying more exactly the emotional valence the target-person had while describing the life-event.

Correlation of questionnaires/tasks with SOD and EA. Concerning the correlations of self-other distinction and "empathic accuracy" with the applied questionnaires and tasks, results revealed an answer-pattern which was whether expected nor stringent/consistent (see also the discussion at section 4.3.24). Two considerations are listed. First "empathic accuracy"

is defined without taking into account target's self-rating (see also p. 25). This definition of "empathic accuracy" bases on a plausibility-argument and therefore it is not allowable to use it as a synonym to empathic accuracy. Second, the questionnaires are self-rating questionnaires and hence not the actual ability is measured, but the believes a person has about his/her abilities.

The plausibility of a video. Although some participants mentioned that they didn't believe some videos, mostly this belief didn't interfere whether their ability to empathize with the person in the video nor to put themselves in this situation and to sense and rate their own emotional valence. Only ratings of one participant were influenced by their disbelief of the videos' stories. In general it seems that for being empathic the plausibility of a story is no necessary requirement to be empathic, but the willingness to empathize with another person is a prerequisite for empathy.

Affect rating dial. The affect rating dial was a very useful instrument to enable the visualization of emotional valence. As has been seen in this study, it is recommended to undertake some rating dial practice tasks so that participants can become acquainted with the handling of the rating dial and further to get used to its continuous application.

4.3.30 Overview of appropriate stimulus-videos.

Finally two tables are presented which sum up if an incongruent video provides an appropriate stimulus to examine self-other distinction (Table 19) or a congruent video a useful stimulus for information and comparison to examine participant's commitment to engage with these kind of stimuli (Table 20).

Due to the fact that period 1 is whether a suitable period for information and comparison nor a part of the actual stimulus to provoke self-other distinction, this period is not listed in the following tables. Only period 2 and period 3 are relevant to measure SOD or $\neg SOD$.

Table 19: Validation Study: Incongruent videos and its appropriateness as a stimulus to examine self-other distinction.

Group	Video	Valence*	SOD significant at period 2	at period 3	appropriate stimulus**
Group 1	Rudi 3inc	positive	yes	yes	yes
	Michi 1inc	negative	yes	yes	yes
Group 2	Sora 2 inc	positive	yes	yes	yes
	Olcay 5inc	negative	yes	yes	yes
Group 3	Rudi 5inc	negative	yes	yes	yes
	Stephan 3inc	negative	yes	yes	yes

* The *valence* is always out of the perspective of the perceiver.
** The last column sums up if this video provides an appropriate stimulus to examine self-other distinction.

Table 20: Validation Study: Congruent videos and its appropriateness as a stimulus for information and comparison to examine participant's commitment to engage with these kind of stimuli.

Group	Video	Valence*	SOD significant at period 2	at period 3	appropriate stimulus**
Group 1	Lesly 4con	positive	no	no	yes
	Sora 6con	negative	yes	yes	no
Group 2	Rudi 3con	positive	no	no	yes
	Michi 1con	negative	no	yes	no
Group 3	Sora 5con	positive	yes	yes	no
	Michi 3con	negative	yes	no	no

* The *valence* is always out of the perspective of the perceiver.
** The last column sums up if this video provides an appropriate stimulus for information and comparison.

5 Conclusion

The aim of this study was to develop and validate videos as stimuli to examine self-other distinction. But self-other distinction, a prerequisite for empathy, is not visible under normal circumstances. Hence a way was considered to make self-other distinction visible. The solution was found in incongruity i.e., the emotional reaction of a person is in an unexpected and uncommon (mostly opposite) way. For example instead of being happy when the best friend became a father, the target-person describes that he was not happy. This describes a uncommon emotional reaction to this event. This unexpected turn into incongruity was incorporated into the stories of some videos to make self-other distinction visible (incongruent condition).

A further prerequisite to make self-other distinction visible is to make perceivers' emotions measurable. Therefore an affect rating dial was constructed, at which the index was precisely adjustable from the pole *negative* via *neutral* to the pole *positive*. Due to this affect rating dial a second-by-second attunement of the emotional valence was possible and therefore a continuous rating could be demonstrated. In any case rating dial practice tasks, as undertaken in this study, are highly recommended to help participants to become acquainted with the handling of the rating dial and to keep in mind its continuous usage.

For the development of the videos 11 target-persons described about 4 highly negative or positive emotionally loaded autobiographical life-events in front of a camera. They described it in a congruent way i.e., the way they actually experienced this life-event. And also in an incongruent way, where the emotional reaction of the target-person is unexpected. While describing these life-events the target-persons got filmed.

After the development of 101 congruent and incongruent videos, 55 were preselected and with them a pilot study was undertaken. The pilot study served to exclude those videos which were not believable and confusing at an emotional level. Out of this selection, 12 videos (6 incongruent and 6 congruent) were chosen to become validated. In the validation study 63 participants watched a counterbalanced selection of four of these 12 videos in a randomized order. Each video was presented twice. While in the first run participants rated the emotional valence of the target-person (perceivers' other-ratings), in the second run they were asked to put themselves in this situation and to rate their own emotional valence (perceivers' self-ratings).

It was hypothesized that by comparing perceivers' other-ratings and perceivers' self-ratings, perceivers' self-other distinction can be made visible at incongruent stimuli which provides the basis for the validation of these videos. Data analyses revealed that the experiment worked for all six incongruent videos. Hence all six incongruent videos provide appropriate stimuli to

examine self-other distinction.

Open for further research is first the assessment of congruent videos because four out of six showed significant self-other distinction contrary to the initial hypothesis. Second the correlation of empathy and perspective taking related questionnaires and tasks with self-other distinction and empathic accuracy revealed no consistent answer-pattern. Third targets' self-ratings rarely reflected their displayed emotional valence and insofar they weren't used for further analysis.

Summing up two congruent videos can be used for information and comparison to examine the participant's commitment to engage with these kind of stimuli. It shows that being empathic does not mean to abandon the own feelings. The incongruent videos allow to make visible self-other distinction, the ability to feel with the other person and to remain with one's own emotions at the same time. Hence these incongruent videos provide useful and innovative stimuli to examine self-other distinction in the context of empathy.

References

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37–46.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251.

Bischof-Köhler, D. (1994). Selbsobjektivierung und fremdbezogene Emotionen. Identifikation des eigenen Spiegelbildes, Empathie und prosoziales Verhalten im 2. Lebensjahr. *Zeitschrift für Psychologie*, *202*, 349–377.

Blatt, S. J. (1974). Levels of object representation in anaclitic and introjective depression. *Psychoanalytic Study of the Child*, *29*, 107-157.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Committee report: Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*, 1017–1034.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), *Cognitive neuroscience of emotion* (pp. 242–276). New York: Oxford University Press.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, *44*, 124–143. doi: 10.1006/brcg.2000.1225

Brass, M., Derrfuss, J., & Von Cramon, D. (2005). The inhibition of imitative and over-learned responses: a functional double dissociation. *Neuropsychologia*, *43*, 89–98. doi: 10.1016/j.neuropsychologia.2004.06.018

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philos Trans R Soc Lond B Biol Sci.*, *364*, 2359–2367. doi: 10.1098/rstb.2009.0066

Carlson, S. M., & Moses, L. J. (2001, July/August). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*, 85–104.

Deborde, A.-S., Berthoz, S., Wallier, J., Fermanian, J., Falissard, B., Jeammet, P., & Corcos, M. (2008). The bermond-vorst alexithymia questionnaire cutoff scores: A study in eating-disordered and control subjects. *Psychopathology*, *41*, 43–49.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*, 71–100.

Decety, J., & Lamm, C. (2006). Human empathy through the lens of social neuroscience. *The Scientific World Journal*, *6*. doi: 10.1100/tsw.2006.221

Decety, J., & Lamm, C. (2011). Empathy versus personal distress: Recent evidence from social neuroscience. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 199–213). Cambridge, MA: MIT Press.

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527–533.

Deci, E., & Ryan, R. (1990). A motivational approach to self: Integration in personality. *Nebraska Symposium on Motivation*, *38*, 237-288.

de Vignemont, F. (2006). When do we empathize? *Novartis Foundation symposium*, *278*, pp.180–195. doi: ijn 00169590

de Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, *10*, 435–441.

Doherty, R. (1997). The emotional contagion scale: A measure of individual differences. *Jounal of Nonverbal Behavior*, *21*(2), 131–154.

E-Prime. (2012). E-prime 2.0 [Computer software manual]. Retrieved from `http://www.pstnet.com`

Eslinger, P. J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology*, *39*, 193–199.

Excel. (2003). [Computer software manual]. Redmond, Washington: Microsoft.

Field, A. (2009). *Disconvering statistics using spss* (3rd edition ed.). London, England: SAGE Publications Ltd.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the level 1-level 2 distinction. *Developmental Psychology*, *, Vol. 17, No. 1,*(1), 99-103.

Gramann, K., & Schandry, R. (2009). *Psychophysiologie. Körperliche Indikatoren psychischen Geschehens* (4., vollständig überarbeitete Auflage ed.). Einheim, Basel: Beltz Verlag.

Gross, J. J., & John, O. P. (1997). Revealing feelings: Facets of emotional expressivity in self-reports, peer ratings, and behavior. *Journal of Personality and Social Psychology*, *72*, 435–448.

Hatfield, E., Cacioppo, J., & Rapson, R. L. (1994). *Emotional contagion.* New York: Cambridge University Press.

Hatfield, E., Rapson, R. L., & Le, Y.-C. L. (2011). Emotional contagion and empathy. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 19–30). Cambridge, MA: MIT Press.

Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, *61*(4), 587–610.

Ickes, W. (2011). Empathic accuracy: Its links to clinical, cognitive, developmental, social,

and physiological psychology. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 57–70). Cambridge, MA: MIT Press.

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, *59*(4), 730–742.

Jennings, J. R., Berg, W. K., Hutcheson, J. S., Obrist, P., Porges, S., & Turpin, G. (1981). Committee report: Publication guidelines for heart rate studies in man. *Psychophysiology*, *18*(3), 226–231.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32–38.

Lamm, C., Meltzoff, A. N., & Decety, J. (2009). How do we empathize with someone who is not like us? a functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *22*(2), 362–376.

Lamm, C., Nusbaum, H. C., Meltzof, A. N., & Decety, J. (2007). What are you feeling? using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *Plos One PLoS ONE,*, *2*(12, e1292). doi: 10.1371/journal.pone.0001292

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiotogy*, *30*, 261–273.

Lanzetta, J., & Englis, B. G. (1989, April). Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *Journal of Personality and Social Psychology*, *56*(4), 543–554.

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, *63*(2), 234–246.

Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clincally relevant setting. *Journal of Personality and Social Psychology*, *68*(5), 854–869.

MATLAB. (2015). Version 8.5 (R2015a) [Computer software manual]. Natick, Massachusetts. Retrieved from `http://de.mathworks.com/products/matlab/`

Müller, J., Bühner, M., & Ellgring, H. (2004). The assessment of alexithymia: psychometric properties and validity of the bermond-vorst alexithymia questionnaire. *Personality and Individual Differences*, *37*, 373–391. doi: 10.1016/j.paid.2003.09.010

Mohiyeddini, C., John, O., & Gross, J. J. (2008). Der „Berkeley Expressivity Questionnaire". Deutsche Adaption und erste Validierungsbefunde. *Diagnostica*, *54*(3), 117–128.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Paulus, C. (2009). *Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity*

Index. Retrieved from `http://psydok.sulb.uni-saarland.de/volltexte/2009/2363` (Online; Visited: 2013, August 22)

Photoshop. (2013). Adobe Photoshop CC (Version 14.0.) [Computer software manual].

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(04), 515–526.

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(01), 1–20.

R Development Core Team. (2015). R: A language and environment for statistical computing (version 0.99.441) [Computer software manual]. Vienna, Austria. Retrieved from `http://www.R-project.org/`

Rauchbauer, B., Majdandžić, J., Hummer, A., Windischberger, C., & Lamm, C. (2015, September). Distinct neural processes are engaged in the modulation of mimicry by social group-membership and emotional expressions. *Cortex*, *70*, 49–67. doi: 10.1016/j.cortex.2015.03.007

Rogers, C. R. (1959). A theory of therapy, personality and interpersonal relationships, as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science. vol. III. Formulations of the person and the social context.* (pp. 184–256). New York: McGraw Hill.

Rogers, C. R., & Roethlisberger, F. J. (1991, November–December). Barriers and gateways to communication. *Harvard Business Review*. (This article originally appeared in HBR July–August 1952.)

Ruef, A., & Levenson, R. (2007). Continuous measurement of emotion. In J. Coan & J. Allen (Eds.), *The handbook of emotion elicitation and assessment* (pp. 286–297). New York: Oxford University Press.

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. *Current Biology*, *22*(23), 1–4.

Schandry, R. (1996). *Lehrbuch Psychophysiologie. Körperliche Indikatoren psychischen Geschehens* (3., korrigierte Auflage ed.). Weinheim: Psychologie Verlags Union.

Shamay-Tsoory, S. G. (2011). Empathic processing: Its cognitive and affective dimensions and neuroanatomical basis. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 215–232). Cambridge, MA: MIT Press.

Simpson, J. A., Ickes, W., & Blackstone, T. (1995). When the head protects the heart: Empathic accuracy in dating relationships. *Journal of Personality and Social Psychology*, *69*(4), 629-641.

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the N.Y. Academy of Science*, *1156*, 81–96.

Sodian, B. (2002). Entwicklung begrifflichen Wissens. In R. Oerter & L. Montada (Eds.),

Entwicklungspsychologie (pp. 443–468). Weinheim, Basel, Berlin: Beltz PVU.

Soto, J. A., & Levenson, R. W. (2009). Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion*, *9*(6), 874–884.

SPSS. (2010). IBM SPSS Statistics for Windows (Version 19.0.) [Computer software manual]. Armonk, NY: IBM Corp..

Starke-Perschke, S. (Ed.). (2001). *Der Brockhaus: Psychologie. Fühlen, Denken und Verhalten verstehen. [the brockhaus: Psychology. understanding feeling, thinking, and behaviour]*. Leipzig, Mannheim: F.A. Brockhaus GmbH.

Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology*, *62*(5), 787-797.

Super. (2013). Simplified Universal Player Encoder & Recorder. (v2013.build.57) [Computer software manual]. Retrieved from `http://www.erightsoft.com/SUPER.html` (Online; Retrieved October 16, 2013)

Tausch, A. (2006). *Facetten emotionaler Expressivität auf Fragebogenebene.* (Inaugural-Dissertation. Johannes Gutenberg-Universität Mainz)

Wilcox, R. (2011). *Modern statistics for the social and behavioral sciences: A practical introduction.* CRC Press.

Zaki, J., & Ochsner, K. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature Neuroscience*, *15*(5), 675–680.

Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *PNAS*, *106*(27), 11382–11387.

List of Figures

List of Tables

Appendix

- Soziodemographisches Datenblatt

- Probandeninformation und Einverständniserklärung

- Validation Study: Post-Video Questionnaire (PVQ)

- Target's Expressivity measured by the BEQ (Score).

- Scores of all Subscales of the FFEE as a Measure of Target's Expressivity

- Graphs of Target's Self-rating.
  Figure 27 to Figure 38 represent the graphics of target's self-rating and simultaneously perceivers' other-ratings and perceivers' self-ratings. The x-axis displays the time while the y-axis presents the valence of the affect rating dial. Target's self-rating is coloured black. The mean of perceivers' other-ratings are in orange tone (the standard deviation in orange shade) and the mean of perceivers' self-ratings are green (the standard deviation in green shade).

- Pilot Study: Pool of possible congruent videos for the validation study.

- Pilot Study: Pool of possible incongruent videos for the validation study.

- Pilot Study: Cites of the participants (TP).

- Validation study: Correlation of questionnaires and $|SOD|$.

- Validation study: Correlation of questionnaires and $|\text{``}EA\text{''}|$.

- Validation study: Results of the Post Video Questionnaire (PVQ).

- Figures at pages 167–229 present the other-ratings and self-ratings of each single participant, also the graphics of the two excluded participants (TP311 and TP330). Participant's individual other-rating is colored pink, single participant's self-rating blue. The mean of all perceivers' other-ratings is displayed as an orange line (and standard deviation as orange shade); the mean of perceivers' self-ratings is green (standard deviation as green shade).

- Curriculum Vitae

Soziodemographisches Datenblatt

Datum _____

Geschlecht: Alter: _____

☐ weiblich   ☐ männlich

Höchste abgeschlossene Ausbildung :
☐ noch Schüler
☐ Pflichtschule
☐ Lehre/Berufsschule
☐ Berufsbildende mittlere Schule
☐ Berufsbildende höhere Schule
☐ Bachelor-Studium
☐ Master/Diplomstudium
☐ Doktorat

Aktuelle Berufstätigkeit:

☐ SchülerIn

☐ StudentIn   Studienrichtung: _____

☐ Selbstständig   Berufsfeld: _____

☐ Angestellter   Berufsfeld: _____

☐ Arbeiter

☐ arbeitslos/Arbeit suchend

Probandeninformation und Einverständniserklärung

Wir laden Sie ein, an der Studie

„Emotionsbewertung"

als freiwillige(r) Proband(in) teilzunehmen.

Im Rahmen dieser Studie sollen der *Hautleitwert* und die *Herzfrequenzvariabilität* in Zusammenhang mit der gestellten Aufgabe untersucht werden. Ziel der Studie ist, den zeitlichen Ablauf bestimmter Verarbeitungsschritte mit Verhaltensdaten zu korrelieren und so Einblick in die *Emotionsbewertung*, zu gewinnen.

Messung des Hautleitwerts (EDA)

Der Hautleitwert ist ein Maß für vorübergehende Veränderungen in der Leitfähigkeit der Haut. Diese Veränderungen sind mit bestimmten Ereignissen verknüpft (z.B. einer erfreulichen Überraschung). Grund für derartige Veränderungen ist die Aktivierung der Schweißdrüsen in emotional bedeutsamen Situationen. Ein den Hautleitwert verändernder Reiz kann sowohl von innen (Gedanken, Erinnerungen, Emotionen) als auch von außen (Bilder, Töne, Ereignisse) kommen.

Erfasst wird der Hautleitwert über zwei Elektroden, die an den mittleren Fingergliedern des Zeige- und Ringfingers der linken Hand angebracht werden. Die Applikation erfolgt mittels elastischer Bänder, die per Klettverschluss entsprechend justiert werden können.

Messung der Herzfrequenzvariabilität (EKG)

Als Herzfrequenzvariabilität wird die Fähigkeit eines Organismus (Mensch, Säugetier) bezeichnet, die Frequenz des Herzrhythmus zu verändern. Über autonome physiologische Regulationswege passt ein gesunder Organismus die Herzschlagrate beständig momentanen Erfordernissen an. Diese Veränderungen treten sowohl bei körperlicher Beanspruchung als auch in emotional bedeutsamen Situationen auf.

Erfasst wird die Herzfrequenz in dieser Studie mittels zwei zusätzlichen Elektroden. Eine Elektrode wird auf der linken Körperhälfte unterhalb des Rippenbogens angebracht, die andere unter dem rechten Schlüsselbein.

Risiken und Unannehmlichkeiten der Hautleitwerts- und EKG-Messung

Sowohl die Hautleitwertmessung als auch die Herzfrequenzvariabilitätsmessung sind schmerzfrei und ohne Gesundheitsrisiko. Sämtliche mit der Haut in Kontakt kommende Materialien sind desinfiziert.
Es werden lediglich Veränderungen des Hautleitwerts und der Herzfrequenz gemessen. Es wird keinerlei Strom zugeführt. Durch speziell geerdete Geräte besteht keine Gefahr elektromagnetischer Induktion. Sehr selten kommt es vor, daß eine vorläufige kleine Hautirritation unter der Elektrode entsteht und es zu einem leichten Juckreiz kommen kann.

Teilnahmebeschränkungen für die Hautleitwert- und EKG-Messung

Sie dürfen nicht an der Untersuchung teilnehmen, wenn Sie:

1.) vorherige negative Erfahrungen mit einer Hautleitwert- und/oder EKG-Messung hatten,
2.) eine Erkrankung des Herzens (Herz-Arrhythmie, Herzschrittmacher) haben,
3.) regelmäßig Psychopharmaka einnehmen,
4.) eine Suchterkrankung (z.B. Alkohol, Drogen, Psychopharmaka) haben,
5.) zuwenig Schlaf und/oder eine große Menge von psychotropen Substanzen in den letzten 24 Stunden zu sich genommen haben.

Bitte legen Sie alle metallische Objekte, die Sie am Körper und/oder in der Kleidung mit sich tragen, ab (z.B. Ketten, Armreif, Handy, …), damit die Signalübertragung nicht gestört wird.

Ihre Rechte

Selbstverständlich können Sie vor und jederzeit während der Untersuchung weitere Informationen über Zweck, Ablauf, etc. der Studie von den durchführenden Personen erfragen. Sie können die Studie jederzeit, auch ohne Angabe von Gründen, von sich aus abbrechen.

Datenschutz

Sämtliche Ihre Person betreffenden Daten werden getrennt von den bei Ihnen erhobenen EDA- und EKG-Daten (und getrennt von eventuellen Fragebogendaten) aufbewahrt, so dass Ihre Anonymität stets gewahrt bleibt. Die bei Ihnen erhobenen Daten fließen in eine Gruppenanalyse ein. Das heißt es werden im Rahmen der Interpretation der Daten und der Verschriftlichung der Studie nur Aussagen über die gesamte Stichprobe getroffen, nicht über einzelne Personen.

Einverständniserklärung

Durch Ihre Unterschrift bestätigen Sie, daß Sie die Probandeninformation gelesen und verstanden haben. Sie erklären sich mit der Teilnahme an dieser Studie sowie der Analyse und Verwendung Ihrer Daten für diese und weitere Studienzwecke durch befugte Personen einverstanden.

Sie wurden darauf hingewiesen, dass Sie den Anweisungen der studiendurchführenden Mitarbeiter im Interesse Ihrer eigenen Sicherheit nachkommen sollen und daß ein Verschweigen von bestehenden Krankheitszuständen oder der Studie unmittelbar vorangegangenen Medikamenteneinnahmen Ihre eigene Sicherheit gefährden kann.

Name (Blockbuchstaben): _____

Geboren am: _____

Datum: _____   Unterschrift: _____

Ich bestätige, daß ich oben genannte(n) ProbandIn über Zweck und Art der Studie informiert habe:

Name des Untersuchungsleiters: Angelika Längle

Datum: _____   Unterschrift: _____

Datum:

Vp-Code:

Bitte beantworten Sie folgende Fragen zu jedem Video.

Bitte die Zahl bzw. den Text ankreuzen

1. Video	gar nicht				mittel				sehr
Finden Sie die dargestellte Person interessant?	1	2	3	4	5	6	7	8	9
Wie attraktiv finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Wie sympathisch finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Hatten Sie das Bedürfnis sich von der erzählten Situation distanzieren zu müssen oder zu wollen?	1	2	3	4	5	6	7	8	9
Hatten Sie das Gefühl nicht „bei der Sache" gewesen zu sein / abgeschweift zu sein?	1	2	3	4	5	6	7	8	9
Waren Sie in Ihrem Leben schon einmal in einer ähnlichen Situation?	Nein, gar nicht		annäh-ernd ähnlich		ähnlich			Ja, genau gleich	

2. Video	gar nicht				mittel				sehr
Finden Sie die dargestellte Person interessant?	1	2	3	4	5	6	7	8	9
Wie attraktiv finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Wie sympathisch finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Hatten Sie das Bedürfnis sich von der erzählten Situation distanzieren zu müssen oder zu wollen?	1	2	3	4	5	6	7	8	9
Hatten Sie das Gefühl nicht „bei der Sache" gewesen zu sein / abgeschweift zu sein?	1	2	3	4	5	6	7	8	9
Waren Sie in Ihrem Leben schon einmal in einer ähnlichen Situation?	Nein, gar nicht		annäh-ernd ähnlich		ähnlich			Ja, genau gleich	

3. Video	gar nicht				mittel				sehr
Finden Sie die dargestellte Person interessant?	1	2	3	4	5	6	7	8	9
Wie attraktiv finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Wie sympathisch finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Hatten Sie das Bedürfnis sich von der erzählten Situation distanzieren zu müssen oder zu wollen?	1	2	3	4	5	6	7	8	9
Hatten Sie das Gefühl nicht „bei der Sache" gewesen zu sein / abgeschweift zu sein?	1	2	3	4	5	6	7	8	9
Waren Sie in Ihrem Leben schon einmal in einer ähnlichen Situation?	Nein, gar nicht		annäh-ernd ähnlich		ähnlich			Ja, genau gleich	

4. Video	gar nicht				mittel				sehr
Finden Sie die dargestellte Person interessant?	1	2	3	4	5	6	7	8	9
Wie attraktiv finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Wie sympathisch finden Sie die dargestellte Person?	1	2	3	4	5	6	7	8	9
Hatten Sie das Bedürfnis sich von der erzählten Situation distanzieren zu müssen oder zu wollen?	1	2	3	4	5	6	7	8	9
Hatten Sie das Gefühl nicht „bei der Sache" gewesen zu sein / abgeschweift zu sein?	1	2	3	4	5	6	7	8	9
Waren Sie in Ihrem Leben schon einmal in einer ähnlichen Situation?	Nein, gar nicht		annäh-ernd ähnlich		ähnlich			Ja, genau gleich	

Table 21: Target's Expressivity measured by the BEQ (Score).

Target	BEQ			
	Negative Expressivity	*Positive Expressivity*	*Impulse Strength*	*General Expressivity*
Daniel	25	25	26	76
Michi	19	21	24	64
Sora	29	25	23	77
Stephan	27	24	26	77
Anna	23	24	33	80
Olcay	24	15	32	71
Lesly	19	20	25	64
Rudolf	28	22	26	76
Konstanze	18	25	25	68
Matthias	26	22	29	77
Rudi	17	21	30	68

Note. Presented are scores of the questionnaire BEQ.

Table 22: Scores of all Subscales of the FFEE as a Measure of Target's Expressivity.

FFEE & FFEE-I Target	I. Kern-Expressivität							II. Soziale Expressivität				III. Kognitive Expressivität		
	Positive Impulsstärke	Fremdbezogene negative Impulsstärke	Selbstbezogene negative Impulsstärke	Positive Expressivität	Nonverbale negative Expressivität	Verbale negative Expressivität	Ärgerexpressivität	Darstellungsfähigkeit	Darstellungstendenz	Verstellungsfähigkeit	Verstellungstendenz	Gefühlsklarheit	Erlebte Unfähigkeit, Gefühle zu vermitteln	Unerwünschter Gefühlsausdruck
Daniel	14	7	11	18	12	18	9	18	17	14	12	24	18	9
Michi	16	10	12	16	13	16	11	24	12	15	13	28	12	6
Sora	14	7	10	18	18	13	12	18	12	17	9	24	8	3
Stephan	14	6	14	16	15	14	13	23	18	12	7	23	11	5
Anna	14	8	11	18	18	18	15	22	13	18	8	24	10	6
Olcay	9	12	12	11	19	18	12	14	7	13	11	16	13	8
Lesly	14	8	8	15	12	16	8	19	9	11	11	25	16	5
Rudolf	13	8	9	13	15	16	13	18	10	14	9	24	13	7
Konstanze	16	9	9	17	13	16	8	21	10	12	14	20	14	7
Matthias	15	5	11	16	14	17	11	15	13	15	11	27	12	6
Rudi	15	3	8	14	16	15	10	21	16	15	12	20	15	4

Aufnahmeprüfung FH (Lesly 4con)
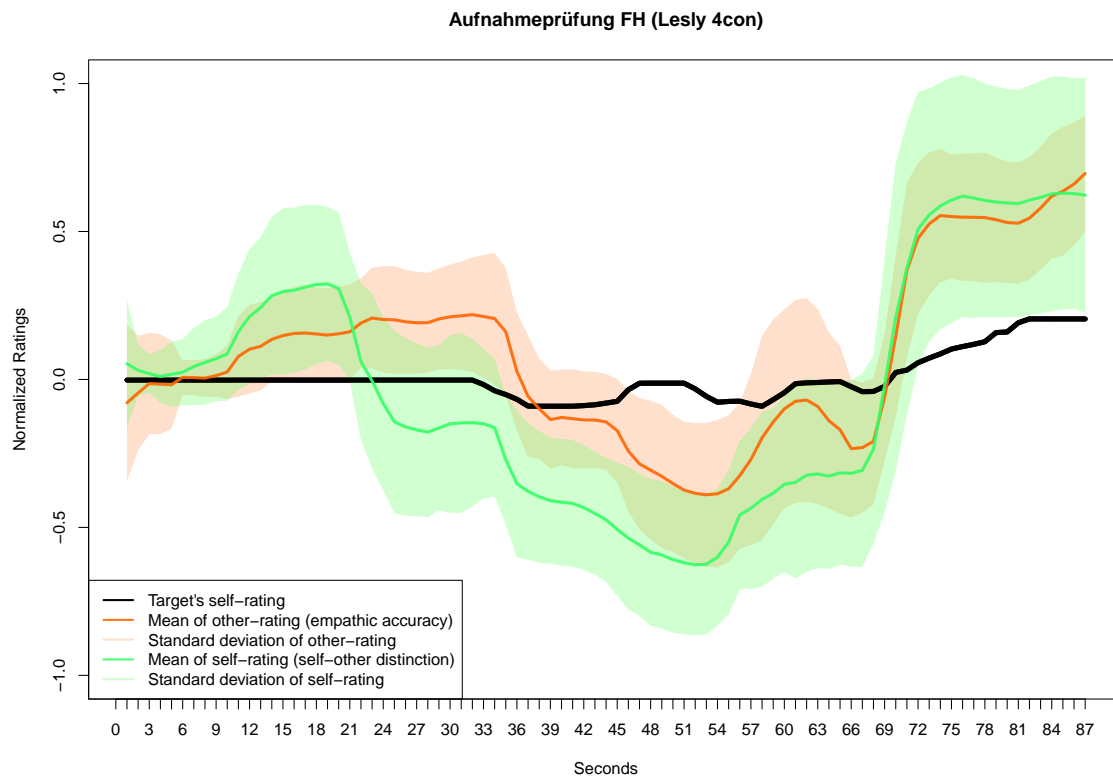
Figure 27: G1 Lesly 4con: Lesly's self-rating.

Chefin kritisiert mich vor Kollegin (Sora 6con)
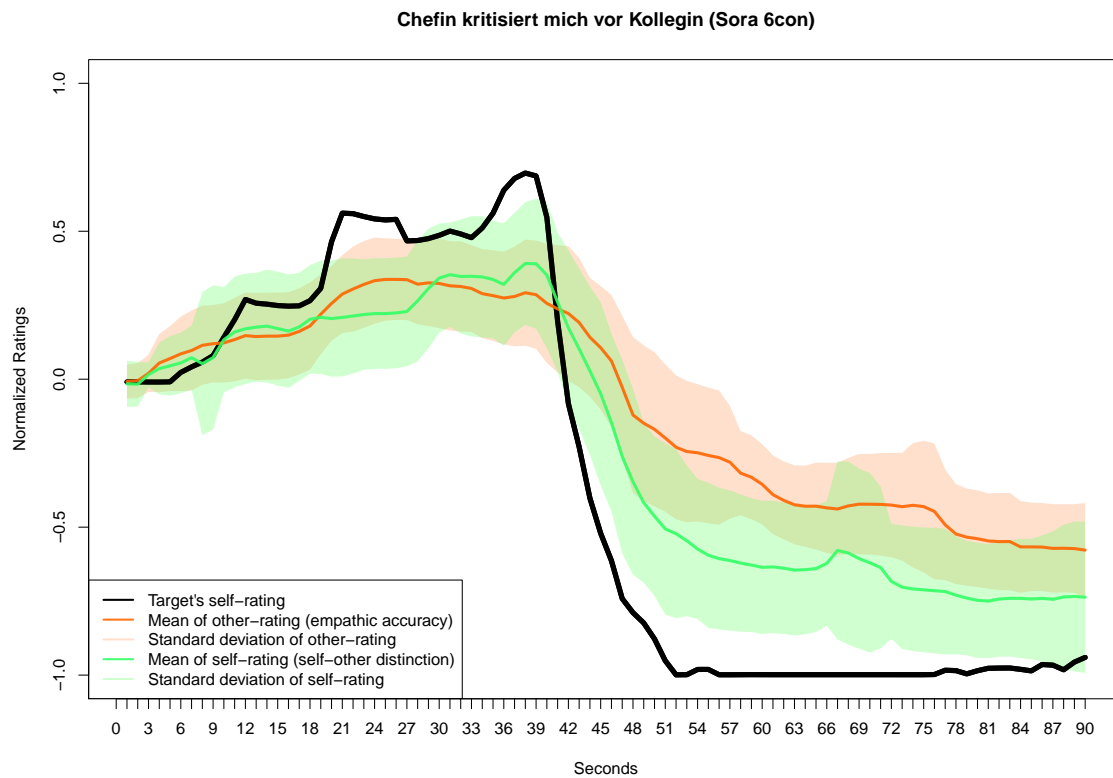
Figure 28: G1 Sora 6con: Sora's self-rating.

Kilias Kindsfeier (Rudi 3inc)
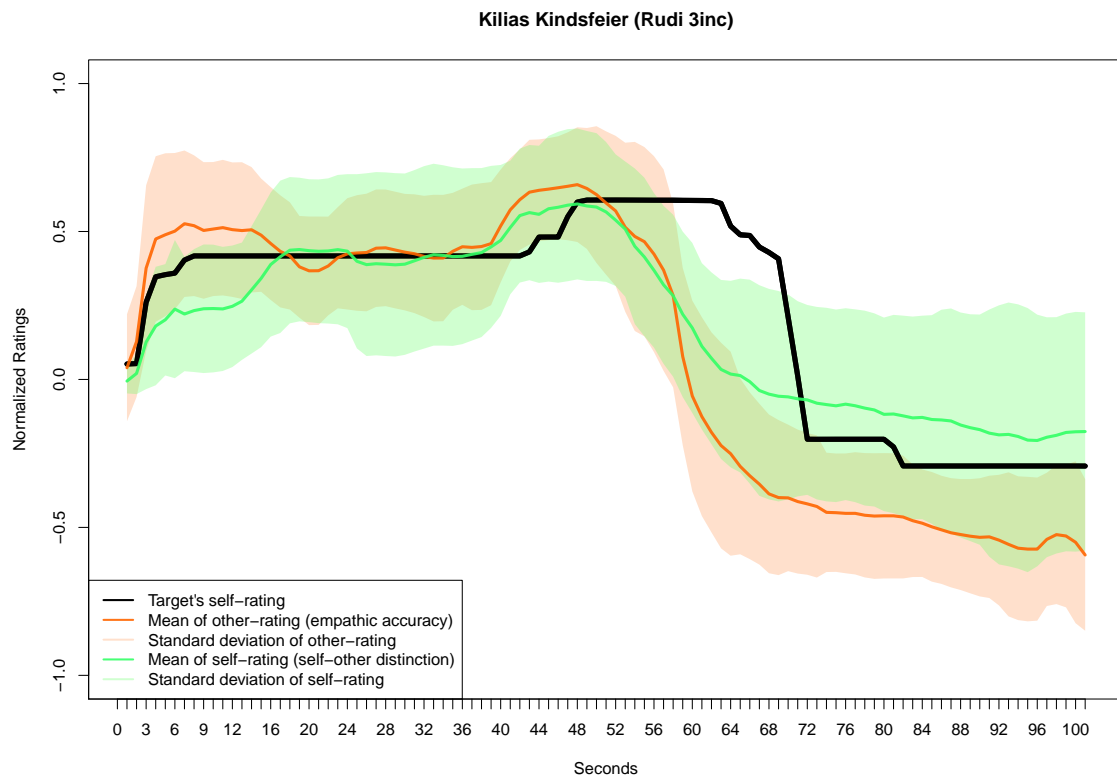
Figure 29: G1 Rudi 3inc: Rudi's self-rating.

Verschimmelter Gouda (Michi 1inc)
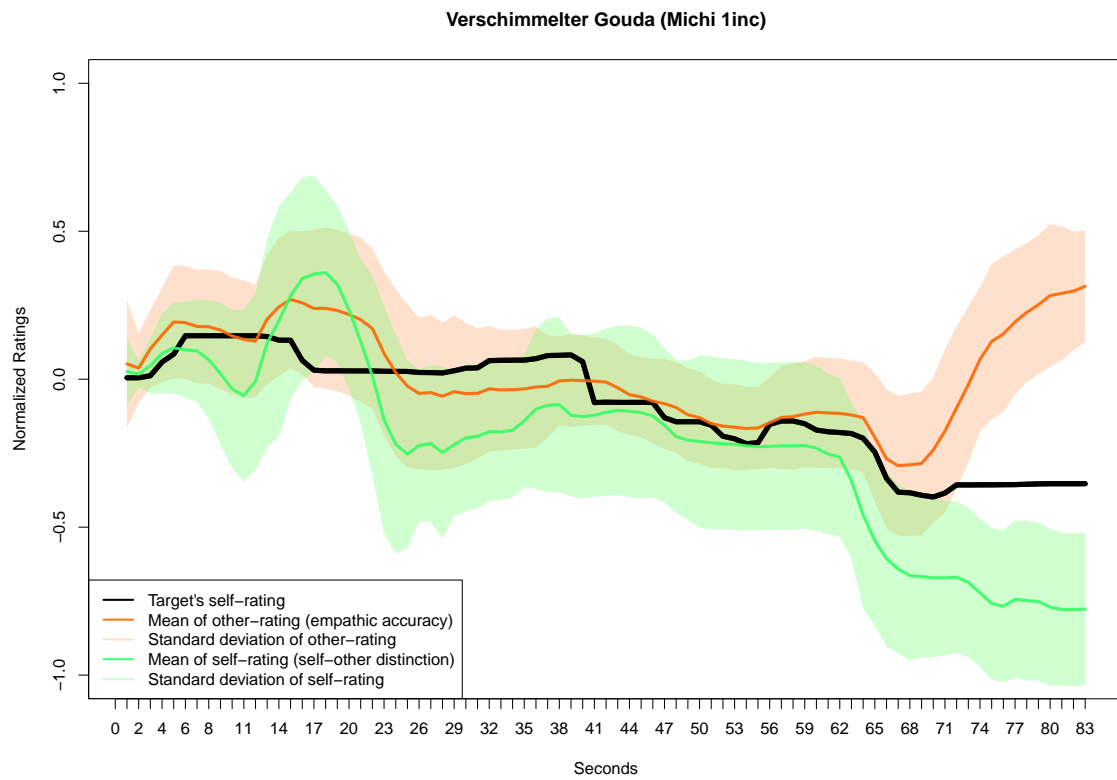
Figure 30: G1 Michi 1inc: Michi's self-rating.

Kilias Kindsfeier (Rudi 3con)
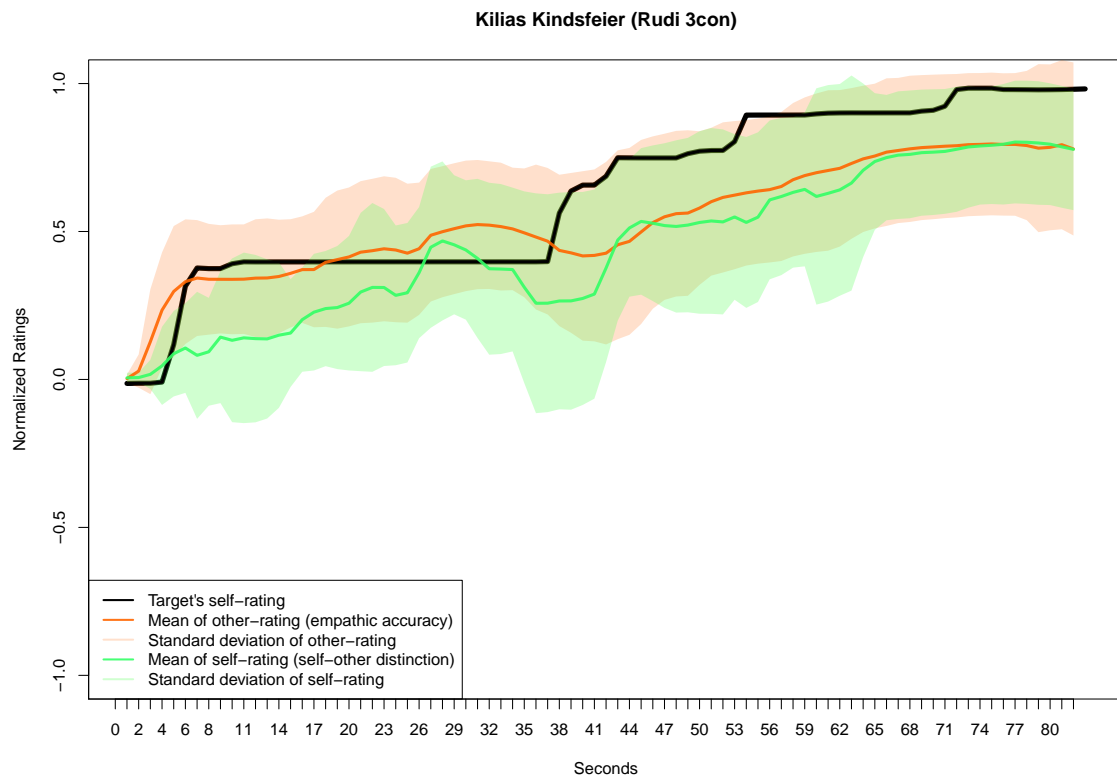
Figure 31: G2 Rudi 3con: Rudi's self-rating.

Verschimmelter Gouda (Michi 1con)
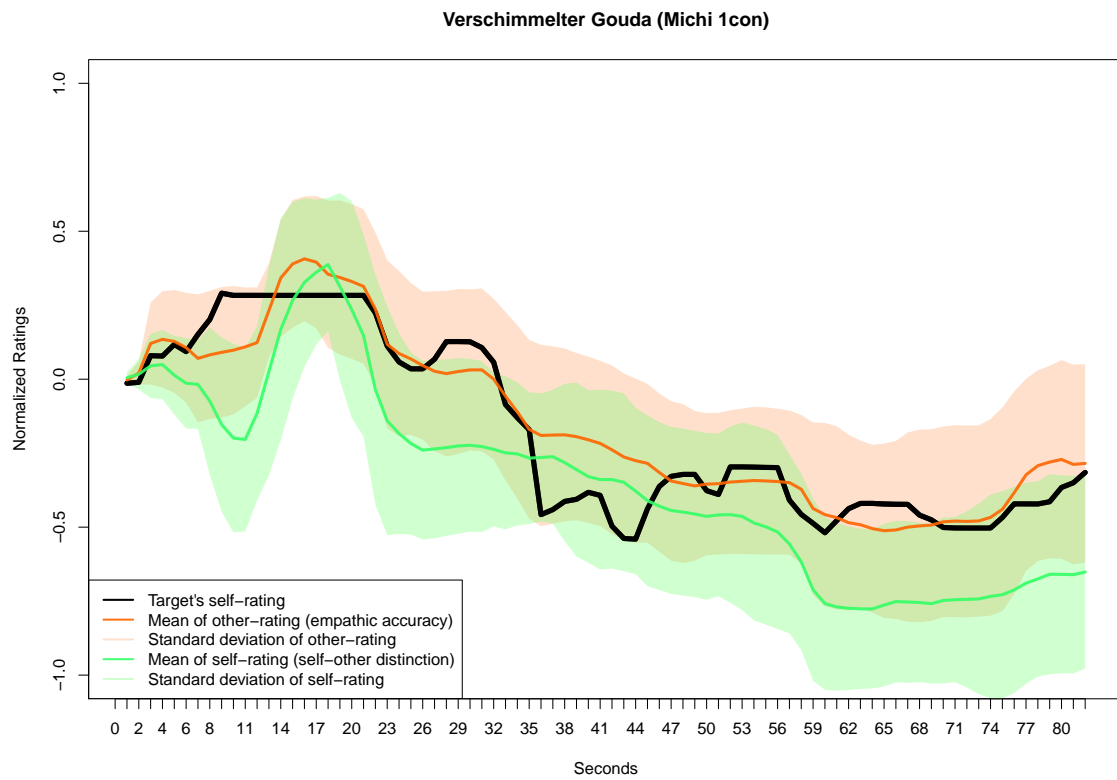
Figure 32: G2 Michi 1con: Michi's self-rating.
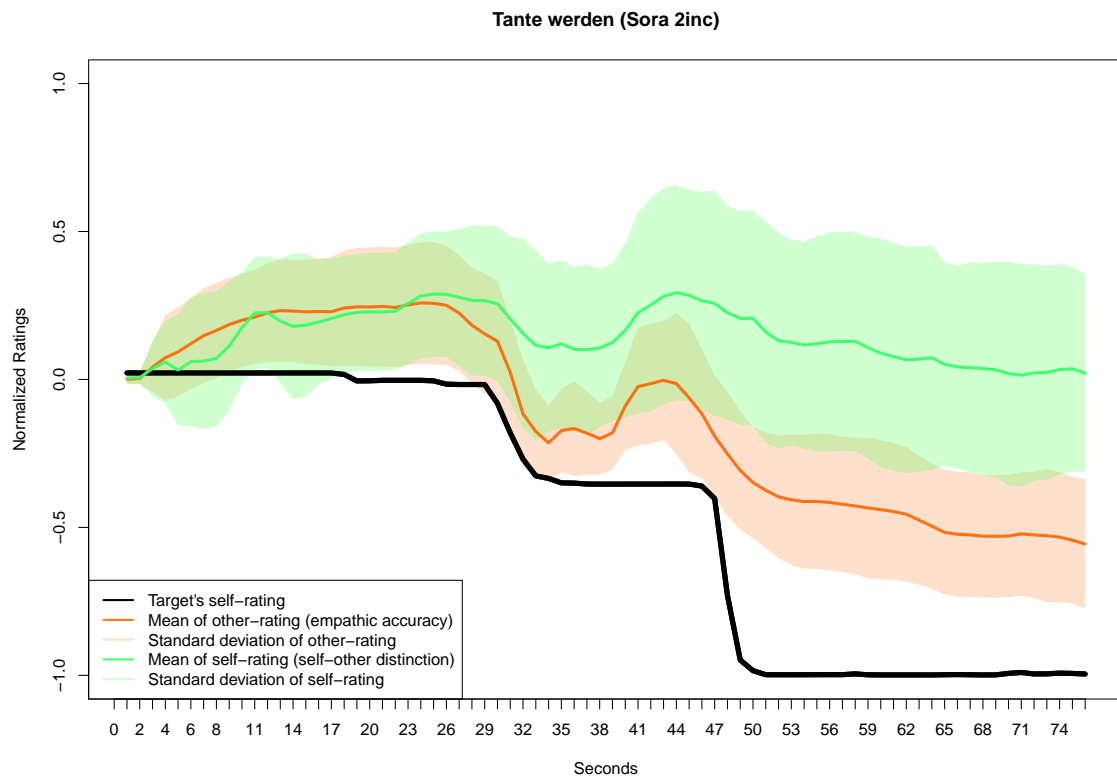
Figure 33: G2 Sora 2inc: Sora's self-rating.
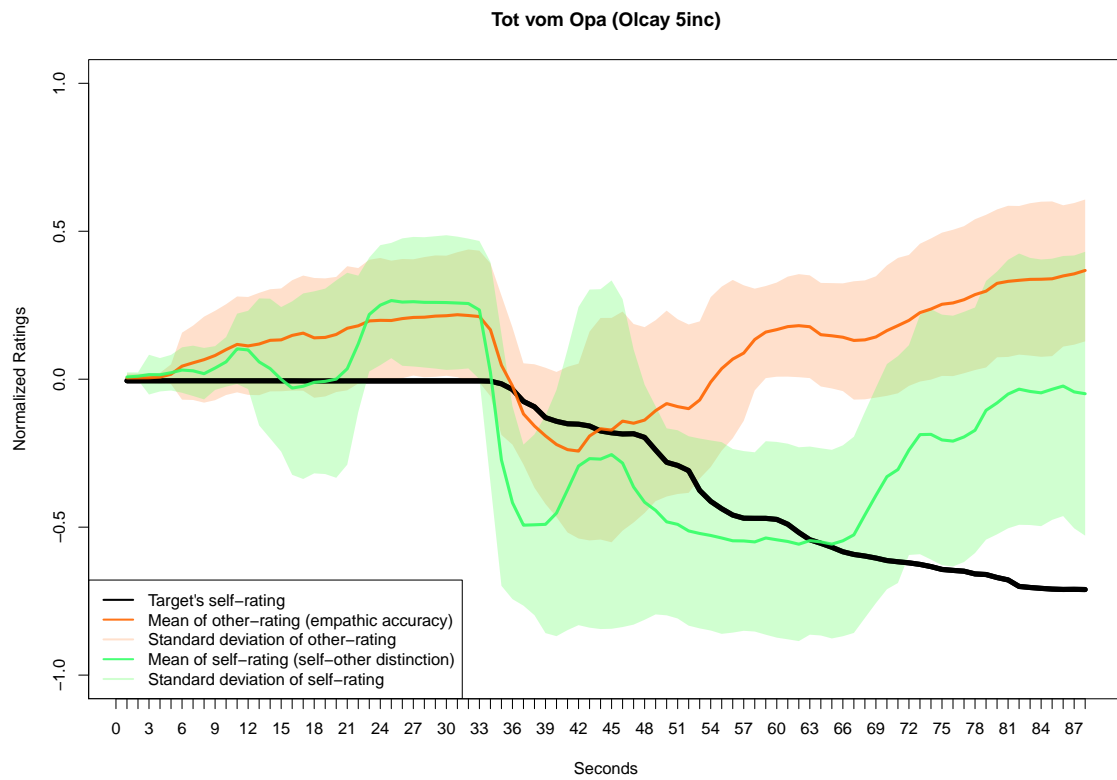
Figure 34: G2 Olcay 5inc: Olcay's self-rating.

1. Freund (Sora 5con)
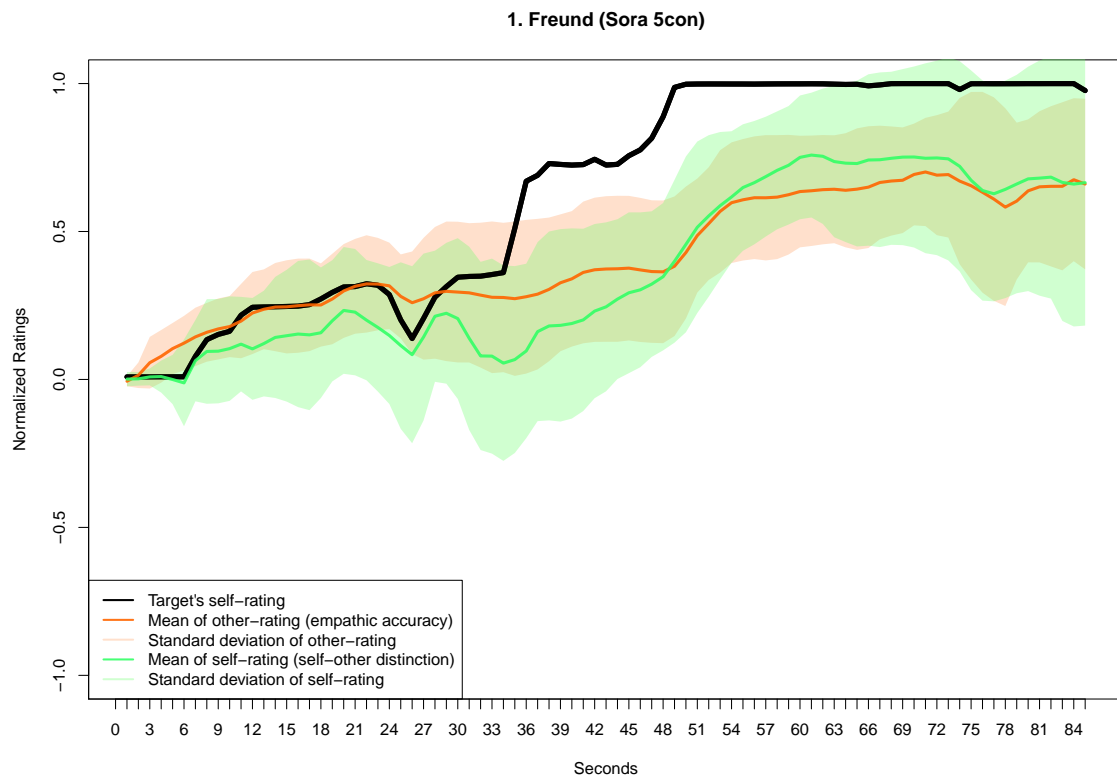
Figure 35: G3 Sora 5con: Sora's self-rating.

Papa trinkt Apfelsaft weg (Michi 3con)
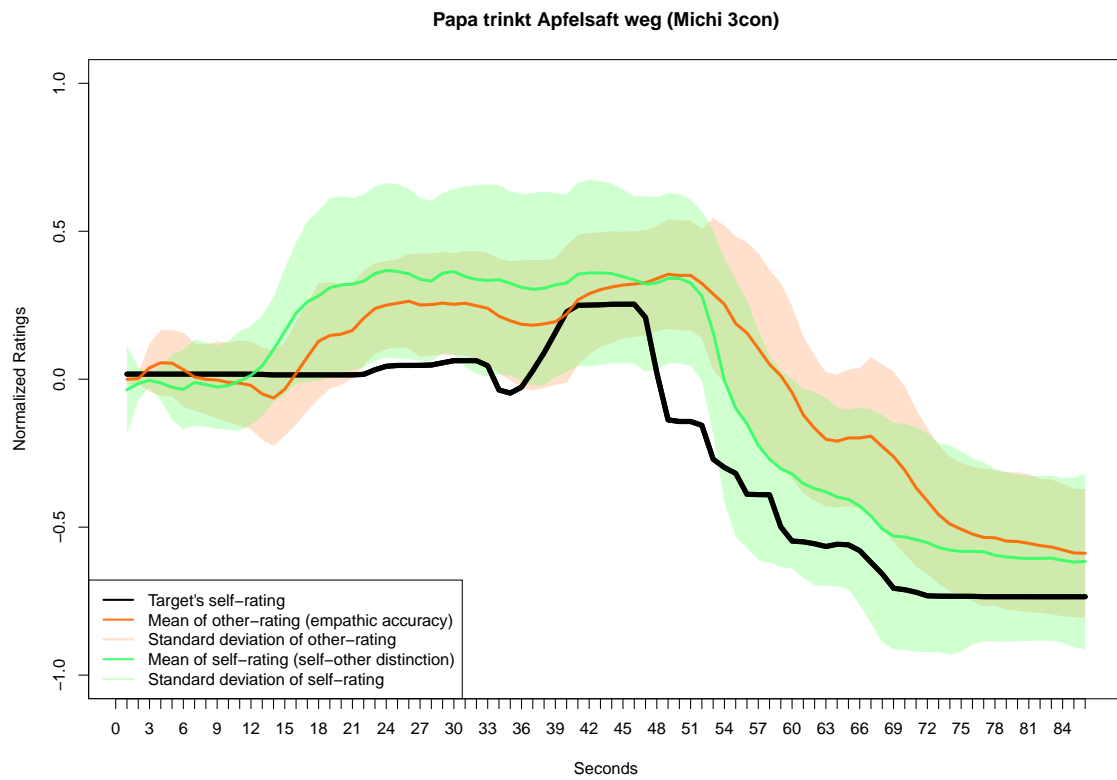
Figure 36: G3 Michi 3con: Michi's self-rating.
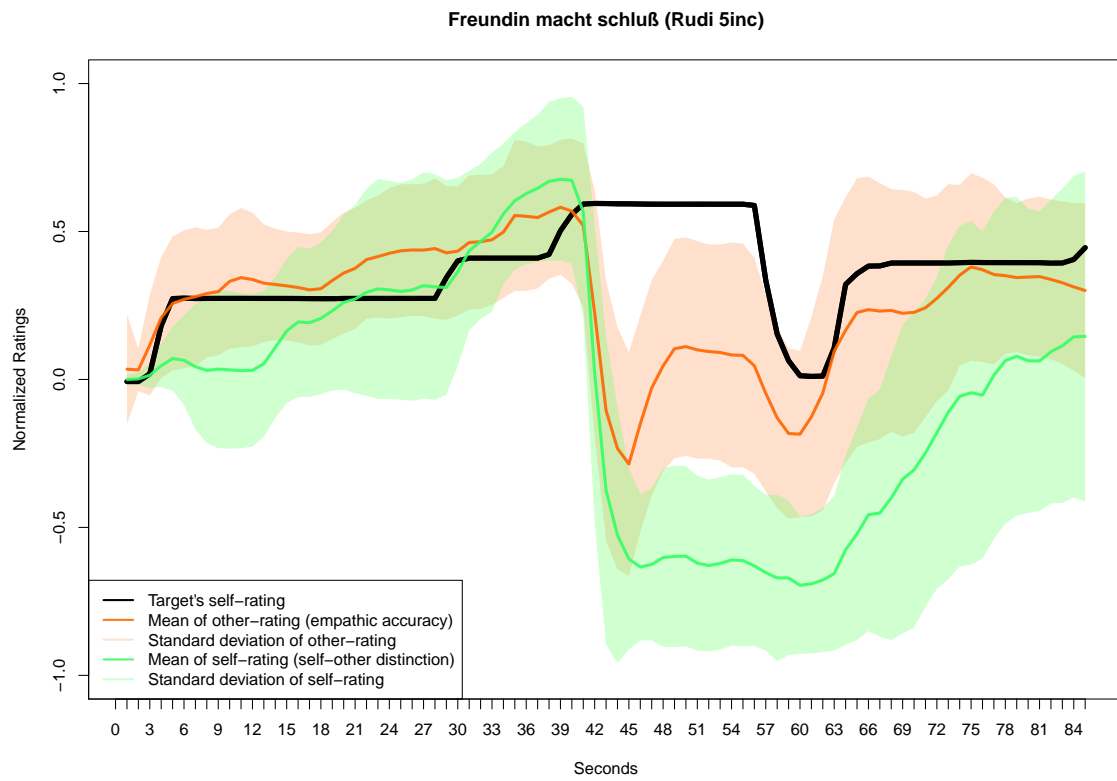
Figure 37: G3 Rudi 5inc: Rudi's self-rating.
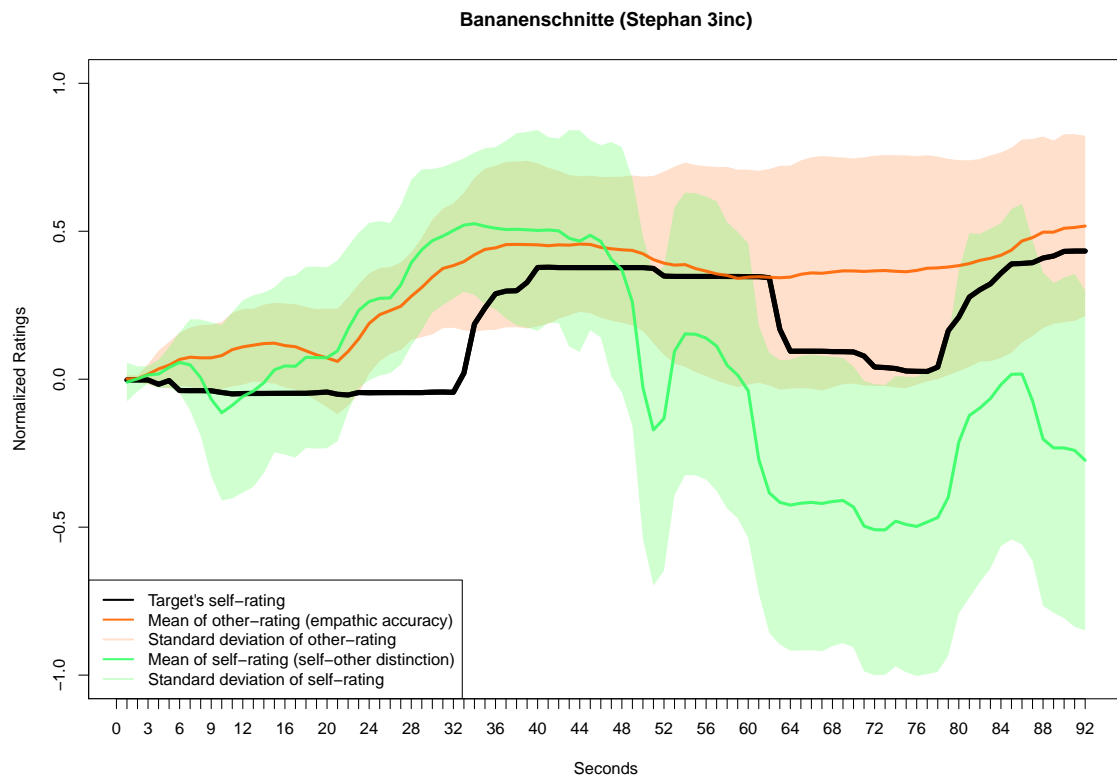
Figure 38: G3 Stephan 3inc: Stephan's self-rating.

Table 23: Pilot Study: Pool of possible congruent videos for the validation study.

Target	"Titel"	2.) Wie attraktiv?	3.) Wie sympathisch?	4.) Wie leicht Gef. nachvollziehen?	5.) Wie sehr mitfühlen?	6.) Wie eindeutig dargestellten Gef.?	Σ aller MW	Σ von 4. (Nachvollziehen) + 5. (Mitfühlen)
Sora_5con	1. Freund	5,14	7,43	8,57	8,29	9	38,43	16,86
Rudi_3con	Kilias Kindsfeier	5,25	7,88	8,25	8,38	8,5	38,26	16,63
Lesly_4con	Aufnahmeprüfung FH	4,5	7,13	8,88	7,75	8,5	36,76	16,63
Rudi_2con	Sponsion	6,14	8	8,43	8	8,71	39,28	16,43
Sora_6con	Chefin kritisiert mich vor Kollegin	6,13	7,25	8,38	7,63	8,5	37,89	16,01
Sora_2con	Tante werden	5,5	7,13	8,25	7,38	8,63	36,89	15,63
Anna_2con	Matura	5,14	6,43	8,43	7,14	8	35,14	15,57
Michi_1con	Gouda - verschimmelt	5,38	7	8,25	7,25	7,75	35,63	15,5
Rudi_5con	Ex-Freundin macht Schluß	6	7,88	8,13	6,88	7,88	36,77	15,01
Michi_3con	Papa trinkt Apfelsaft weg	5,29	6,43	7,86	7,14	7,71	34,43	15
Michi_2con	Gorgonzola	5,57	6,71	7,86	6,86	7,57	34,57	14,72
Rudi_4con	Prüfung nicht geschafft	6,38	7,13	7,88	6,63	8,25	36,27	14,51
Stephan_3con	Bananenschnitte	5,75	6,88	7,88	6,5	8,5	35,51	14,38
Michi_5con	Evas Eltern kennenlernen	5,38	6,25	7,5	6,63	8,63	34,39	14,13
Olcay_4con	Auszug v Daheim	7,29	6,57	6,86	6,43	7,86	35,01	13,29
Sora_4con	Geburtstag - Kroatien	4,86	5,71	7	6	6,57	30,14	13
Michi_4con	Jobzusage ZID	5,57	6,29	7,14	5,71	7	31,71	12,85
Summe		95,27	118,1	135,55	120,6	137,56	607,08	256,15
mean		5,60411765	6,94705882	7,97352941	7,09411765	8,09176471	35,7105882	15,0676471

Table 24: Pilot Study: Pool of possible incongruent videos for the validation study.

Target	"Titel"	2.) Wie attraktiv?	3.) Wie sympathisch?	4.) Wie leicht Gef. nachvollziehen?	5.) Wie sehr mitfühlen?	6.) Wie eindeutig dargestellten Gef.?	Σ aller MW	Σ von 4. (Nachvollziehen) + 5. (Mitfühlen)
Rudi_3inc	Kilias Kindsfeier	6,13	7	7,75	7,25	7,63	35,76	15
Sora_2inc	Tante werden	5,5	6,25	7,5	5,88	8,25	33,38	13,38
Rudolf_1inc	Eigene Wohnung	3,5	6,25	7,5	5,88	7,88	31,01	13,38
Sora_1inc	Grand Canyon	4,88	6,25	7,25	5,38	7,5	31,26	12,63
Stephan_3inc	Bananenschnitte	5,38	6,75	7,38	4,88	8,13	32,52	12,26
Olcay_2inc	Kennenlernen m. Freundes	6,86	5,86	6,43	5,71	6,71	31,57	12,14
Stephan_1inc	Hot Stone Massage	6,25	6	7,63	4,38	6,63	30,89	12,01
Rudi_5inc	Ex-Freundin macht Schluß	5,57	6,29	6	5,43	6	29,29	11,43
Sora_4inc	Geburtstag - Kroatien	5,43	6,29	5,57	5,29	5,57	28,15	10,86
Rudi_2inc	Sponsion	5	7,14	5,57	5,14	6,29	29,14	10,71
Michi_1inc	Gouda - verschimmelt	5,71	6,71	3,57	3	5,86	24,85	6,57
Olcay_5inc	Tod vom Opa	5,57	2,86	3,43	2,29	5,71	19,86	5,72
Summen		65,78	73,65	75,58	60,51	82,16	357,68	136,09
mean		5,48166667	6,1375	6,29833333	5,0425	6,84666667	29,8066667	11,3408333

Table 25: Pilot Study: Cites of the participants (TP).

Target / Video	TP	Zitat
To Michi in general	TP12	"Anfangs war er mir eher wenig sympathisch. Aber je mehr Videos ich von Michi gesehen habe, desto mehr stieg die Sympathie."
	TP19	"Je emotionaler, desto mehr zwinkert er mit den Augen."
To Michi 1con	TP14	"Doch, ich hab das schon geglaubt, aber eklig ists."
To Michi 1inc	TP03	"Sehr gut - man ekelt sich, bevor er sagt es schmecke ihm."
	TP09	"Schon möglich, daß es wirklich so war, aber schräg, daß man auf Schimmel stehen kann..."
To Olcay 5inc	TP11	"So unsympathisch! Ob sie das wirklich so erlebt hat, ist schon fraglich."
	TP13	"Das ist einfach zu heftig…"
	TP18	"Total schlimm, da Olcay ehrenamtlich arbeitet, dann aber eiskalt – unsympathisch! Aber ihr Aussehen ist sehr hübsch. Anfangs habe ichs nicht geglaubt, da es nicht zampasst und dann hab ichs bisi nachvollziehen können – 'anscheinend gibt's so Leute'. – Ich habe dann lang überlegen müssen, also 'find ich sie jetzt attraktiv oder nicht?'. Zuerst dachte ich mir, die kriegt ne 8 bei Attraktivität [*Attraktivität* wurde in einer 9-stufigen Likert-Skala eingestuft, wobei 9 *sehr* darstellte, Anm.]. Aber als sie dann so eiskalt vom Tod ihres Opas erzählte, musste ich lang überlegen, ob ich sie überhaupt attraktiv finde…"
	TP19	(erst beim nachfragen) "Ich hab das gar nicht so schlimm empfunden, sondern sehr ehrlich. Es geht sicher einigen so und sie sagen es nur nicht. So war es sehr ehrlich und sie gibt's zu. Und zudem gibt es ja auch Kulturen, wo Sterben nicht so negativ gesehen wird."
To Rudi in general	TP22	"Rudi spricht im Dialekt – das wirkt weniger geschauspielert."
	TP12	"Rudi ist der coolste überhaupt, weil er ist so naturnah".
To Rudi 3inc	TP03	"Und sind sie wieder zamgekommen?"
	TP09	"Diese Geschichte hat in mir eine Abwehrreaktion ausgelöst. In weiterer Folge habe ich Rudis [weiteren, Anm.] Vidoes dann 'mit Vorsicht genossen'."
To Sora in general	TP22	"Generell schwerer einzuschätzen, da sie so neutral im Ausdruck war."
To Stephan 3inc	TP12	"Glaubhaft und amüsant und super rübergebracht. Ich habs ihm voll geglaubt. Aber ich find ihn an sich unsympathisch."
	TP19	"Sehr amüsant; am meisten amüsant von allen Videos."
In general	TP17	(Erst beim nachfragen, auf die Frage: "Meinst Du irgendein Video war vielleicht nicht ganz wahr?") "Ich bin mir unsicher, ob alle wahr waren. Aber glaubhaft waren sie. Und dieser Faktor ist egal. Mitfühlen kann man sowieso."

Table 26: Validation study: Correlation of questionnaires and $|SOD|$.

Column groups: **BVAQ-B** = Emotionalizing … AllSum; **EC** = Positive / Negative / Total_Mean; **SPF** = Empathic Concern … Empathy (= EC+PT+FS); **Post-Video Questionnaire** = 1. interessant … 6. ähnliche Situation. The *Hypotheses* column refers to the pos. correlation between TP_{other}, TP_{self} & self-other distinction.

TP_{other} − TP_{self} / Self-Other Distinction	Hypotheses	timeperiod	Emotionalizing	Fantasizing	Identifying	Analyzing	Verbalizing	I. Affective Factor	II. Cognitive Factor	AllSum	Positive	Negative	Total_Mean	Empathic Concern	Fantasy	Perspective Taking	Personal Distress	Empathy (= EC+PT+FS)	Reading the Mind in the Eyes	Finger Lifting Task	shelve Task	1. interessant	2. attraktiv	3. sympathisch	4. distanziert	5. abgeschweift	6. ähnliche Situation
(hypothesis)			-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	?	+	-	+	?	?	?	+	+	?
G1 Lesly 4con	con / pos	1.			-.45*											.47*				-.48*		.45*	.45*			-.51*	
		2.	.45*	.45*												.57**		.47*		-.45*			.45*	.46*			
		3.																									
G1 Sora 6con	con / neg	1.																	-.53*								
		2.																							-.48*		
		3.			-.45*																				-.54*	-.51*	
G1 Rudi 3inc	inc / pos	1.								.47*																	
		2.								.44*	-.47*																
		3.																									
G1 Michi 1inc	inc / neg	1.	.50*														-.59**				-.46*	-.48*		-.46*			
		2.					.45*																				
		3.																									
G2 Rudi 3con	con / pos	1.													-.60**								-.65**				
		2.				.47*						-.49*	-.49*														
		3.	.54*									-.49*	-.49*					-.45*							.51*	.57**	
G2 Michi 1con	con / neg	1.																									
		2.																									
		3.																									
G2 Sora 2inc	inc / pos	1.										.45*								.46*							
		2.																		.49*							
		3.																									
G2 Olcay 5inc	inc / neg	1.	-.46*																	.44*				-.52*	.44*		
		2.						-.45*												.56**				-.51*	.52*		
		3.																									
G3 Sora 5con	con / pos	1.	.46*											-.46*													
		2.														-.47*											
		3.																									
G3 Michi 3con	con / neg	1.												-.50*												.56*	
		2.						-.53*								-.53		-.54*								.62**	
		3.																									
G3 Rudi 5inc	inc / neg	1.												.47*													
		2.				-.48*				-.60**														.50*		.56*	
		3.				-.48*				-.50*						-.55*											
G3 Stephan 3inc	inc / neg	1.																	.54*		.52*						-.62**
		2.												-.58**				-.55*									-.52*
		3.																									

* p < .05 (two-sided)
** p < .001 (two-sided)

Table 27: Validation study: Correlation of questionnaires and |"*EA*"|.

TP_other - GroupMean_other "Empathic Accuracy"	timeperiod	BVAQ-B Emotionalizing	Fantasizing	Identifying	Analyzing	Verbalizing	I. Affective Factor	II. Cognitive Factor	AllSum	EC Positive	Negative	Total_Mean	SPF Empathic Concern	Fantasy	Perspective Taking	Personal Distress	Empathy (= EC+PT+FS)	Reading the Mind in the Eyes	Finger Lifting Task	Shelve Task	PVQ 1. interessant	2. attraktiv	3. sympathisch	4. distanziert	5. abgeschweift	6. ähnliche Situation
Hypotheses (neg. correlation between TP_other~Group_other & empathic accuracy, hence hypotheses are vice verse)		+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	+	+	-
G1 Lesly 4con con pos	1. / 2. / 3.																	-.50*								
G1 Sora 6con con neg	1. / 2. / 3.	.60**			-.48*		.54*	-.45*	.45*	.53*					.47*			-.61**								
G1 Rudi 3inc inc pos	1. / 2. / 3.													-.54*												.48*
G1 Michi 1inc inc neg	1. / 2. / 3.																	.46*					-.55** / -.51*	.65**	.60**	-.45*
G2 Rudi 3con con pos	1. / 2. / 3.																		-.45* / -.53*							
G2 Michi 1con con neg	1. / 2. / 3.																									.49*
G2 Sora 2inc inc pos	1. / 2. / 3.																									
G2 Olcay 5inc inc neg	1. / 2. / 3.															-.47*										
G3 Sora 5con con pos	1. / 2. / 3.		.60**				.48*									.50*										
G3 Michi 3con con neg	1. / 2. / 3.	.52*	.52*				.67***	.52*	.52*							.74***			.49*		.47*	.47*				
G3 Rudi 5inc inc neg	1. / 2. / 3.	.72**	.72**				.57*									.55*						.51*				
G3 Stephan 3inc inc neg	1. / 2. / 3.		-.48*		-.50*			-.49*																		

* p < .05 (two-sided)
** p < .001 (two-sided)

Table 28: Validation study: Results of the Post Video Questionnaire (PVQ).

Gruppe	Video	*interessant*	*attraktiv*	*sympathisch*	*distanzieren*	*abgeschweift*	*ähnliche Sit.*
G1	Lesly 4con - Aufnahmepr. FH	6,04	5,28	6,61	2,76	2,52	2,52
	Sora 6con - Chefin kritisiert mich vor Kollegin	6,61	6,04	7	3	2,19	2,23
	Rudi 3inc - Kilias Kindsfeier	6,71	6	7,71	2,71	2	1,76
	Michi 1inc - Gouda verschimmelt	4,52	3,47	5,38	5	2,33	1,9
G2	Rudi 3con - Kilias Kindsfeier	7,38	6,38	8	1,66	1,9	1,52
	Michi 1con - Gouda verschimmelt	4,61	3,38	6,66	4,52	2	2,8
	Sora 2inc - Tante werden	5,23	4,47	5,04	4,38	2,38	1,38
	Olcay 5inc - Tod vom Opa	4,57	6,19	3	6,38	2,33	1,61
G3	Sora 5con - 1. Freund	5,71	4,8	7	2,04	2,04	3,14
	Michi 3con - Papa trinkt Apfelsaft weg	4,19	3,28	5,23	4,19	2,19	1,8
	Rudi 5inc - Exfreundin macht Schluß	6,42	5,47	6,57	5,19	2,19	2,04
	Stephan 3inc - Bananenschnitte	4,9	4,76	6	5,04	2,09	1,9
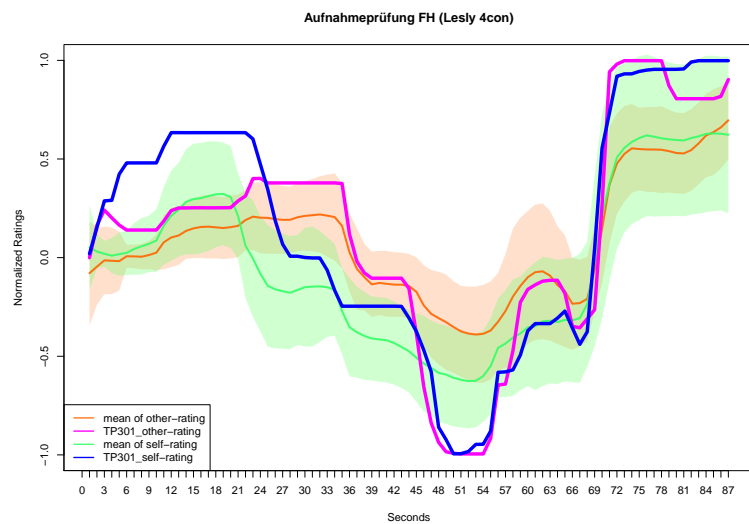
Figure 39: G1 Lesly 4con: Ratings by test-person 301.
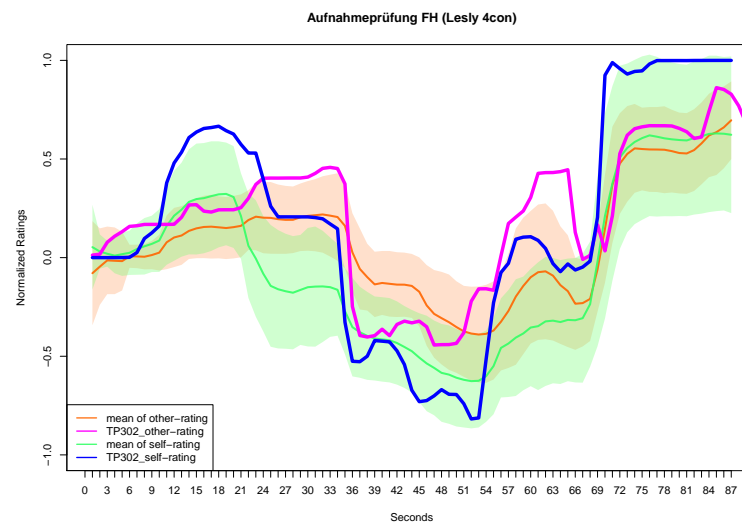
Figure 40: G1 Lesly 4con: Ratings by test-person 302.
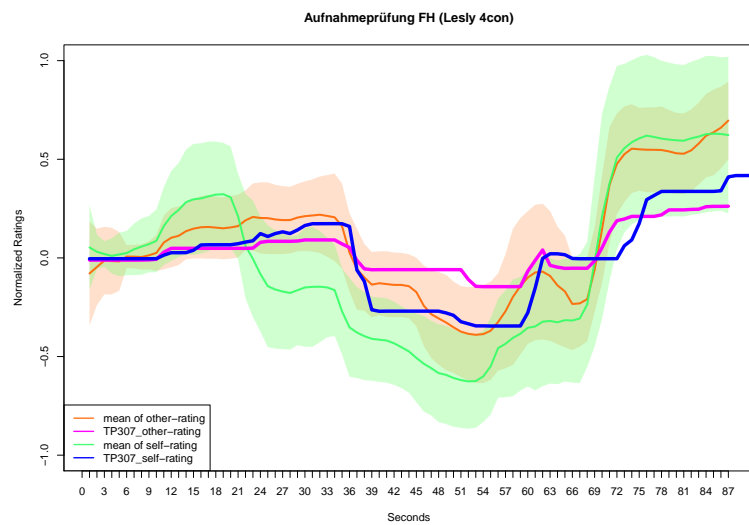
Figure 41: G1 Lesly 4con: Ratings by test-person 307.
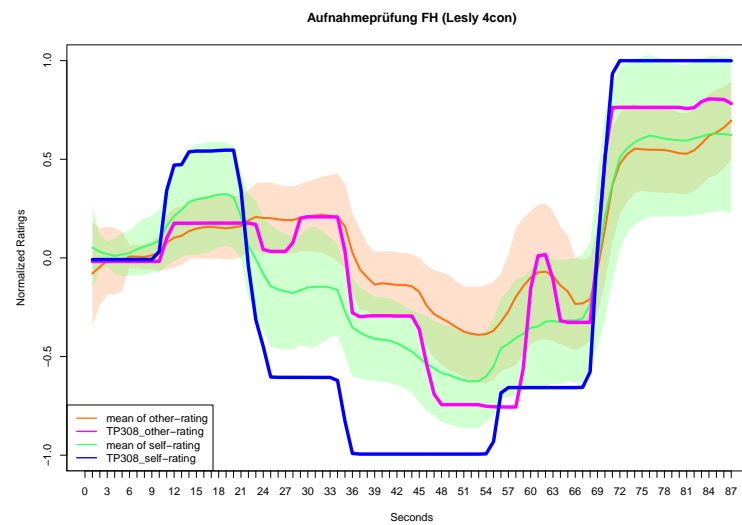
Figure 42: G1 Lesly 4con: Ratings by test-person 308.
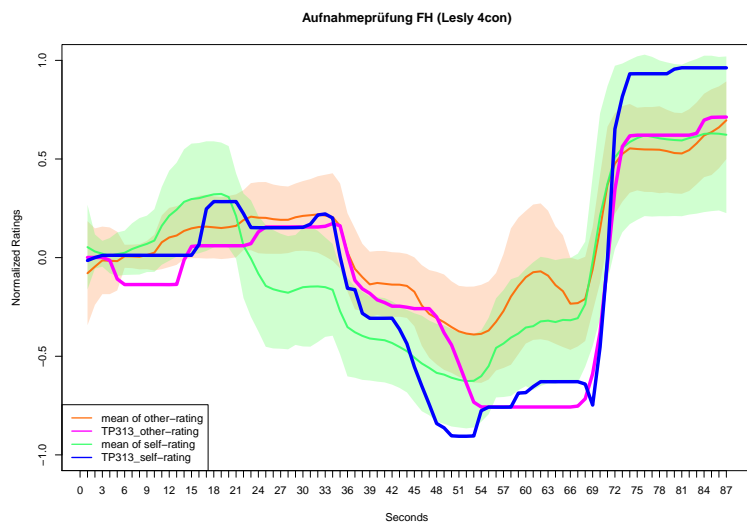
Figure 43: G1 Lesly 4con: Ratings by test-person 313.
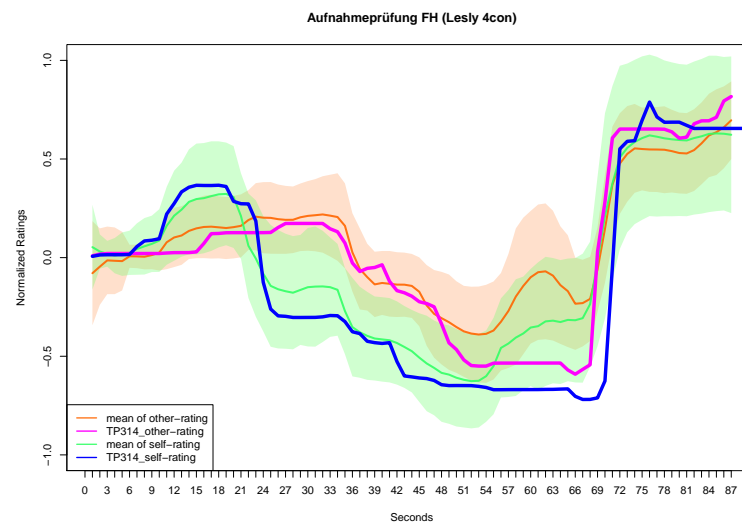
Figure 44: G1 Lesly 4con: Ratings by test-person 314.
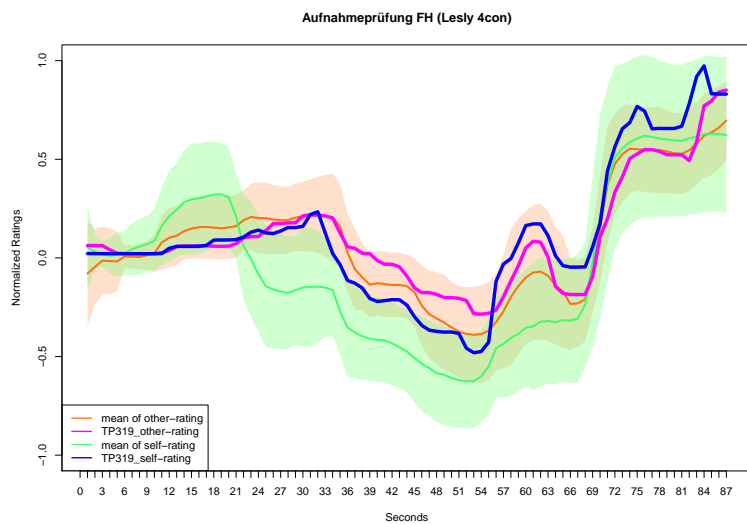
Figure 45: G1 Lesly 4con: Ratings by test-person 319.
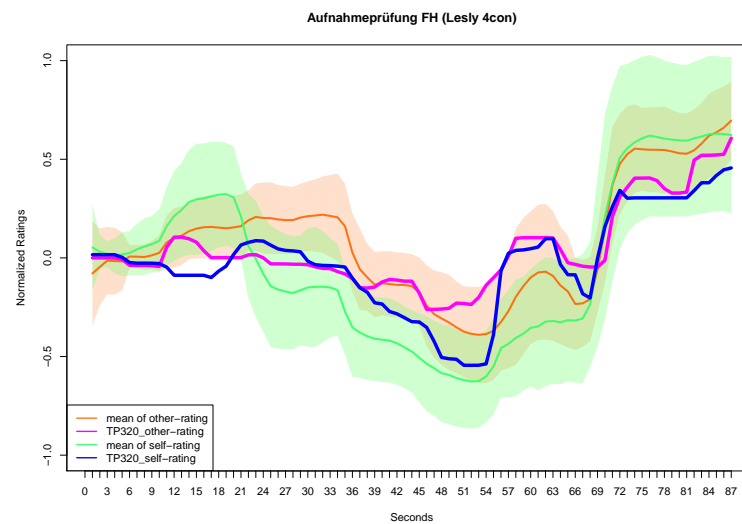
Figure 46: G1 Lesly 4con: Ratings by test-person 320.
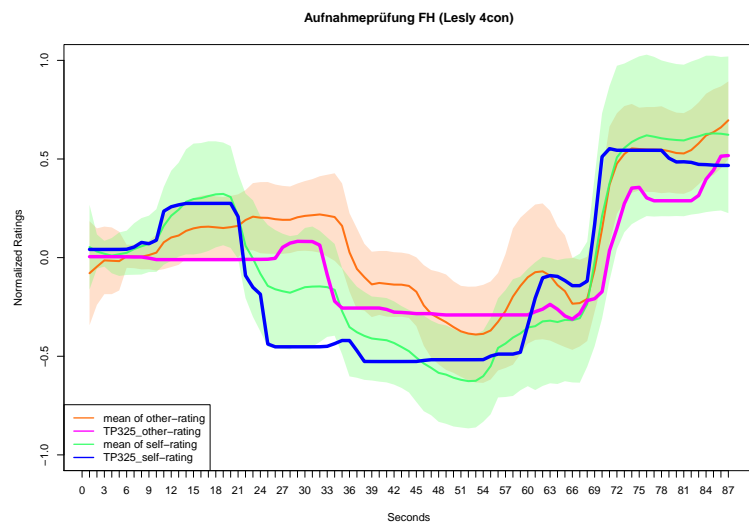
Figure 47: G1 Lesly 4con: Ratings by test-person 325.
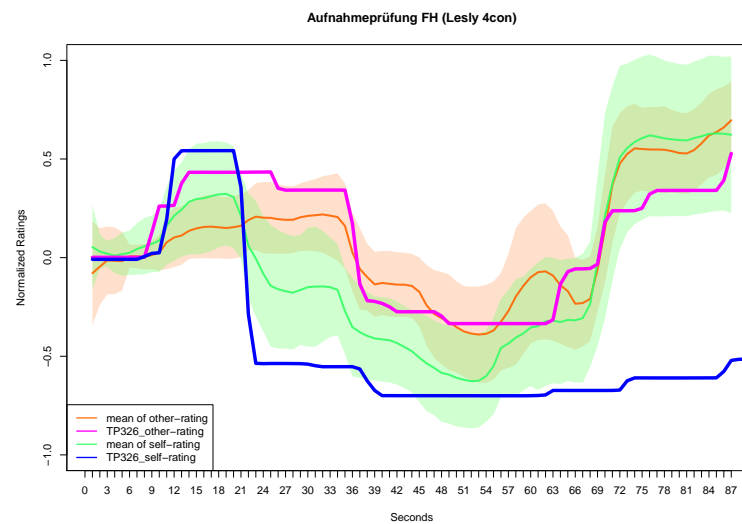
Figure 48: G1 Lesly 4con: Ratings by test-person 326.
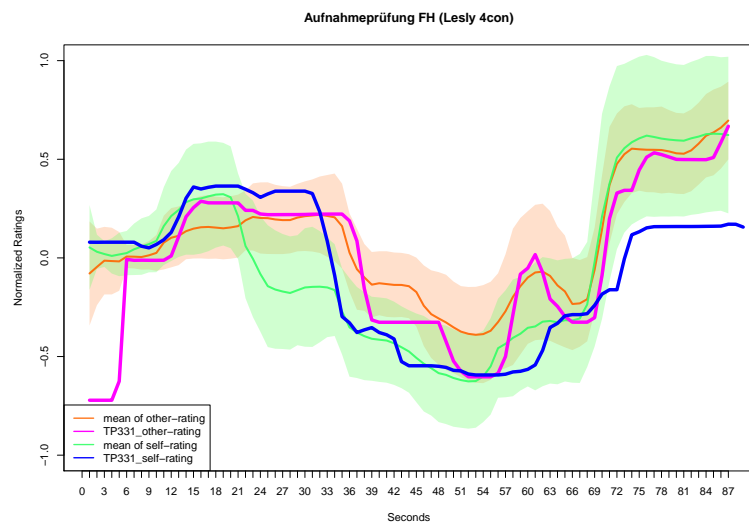
Figure 49: G1 Lesly 4con: Ratings by test-person 331.
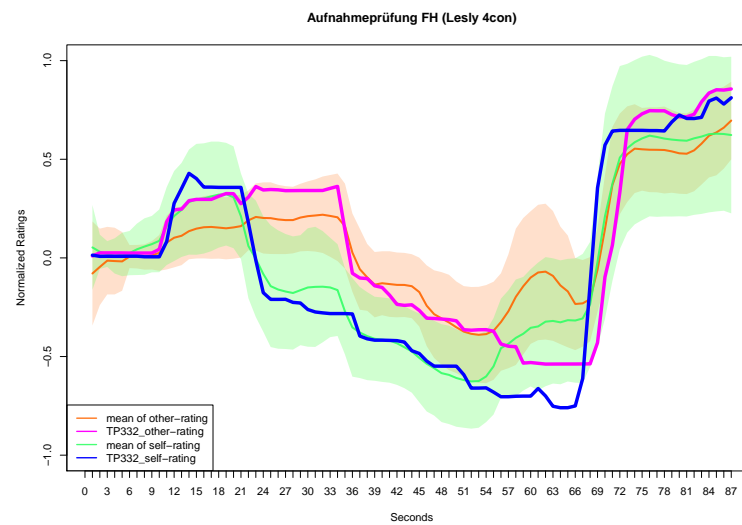
Figure 50: G1 Lesly 4con: Ratings by test-person 332.
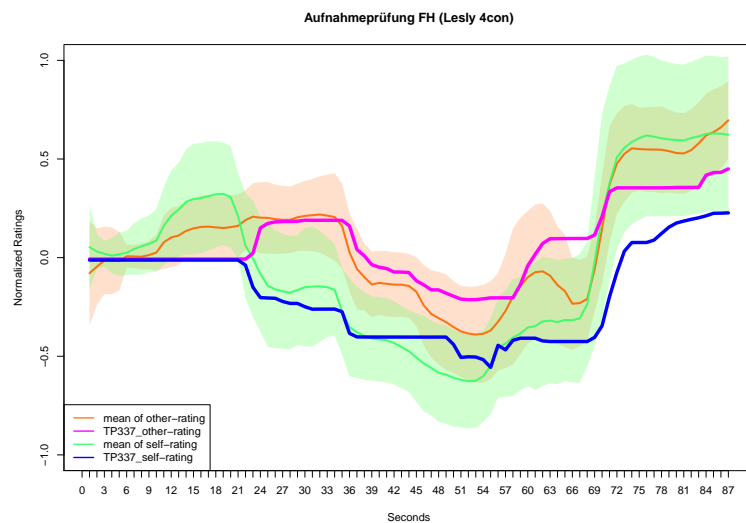
Figure 51: G1 Lesly 4con: Ratings by test-person 337.
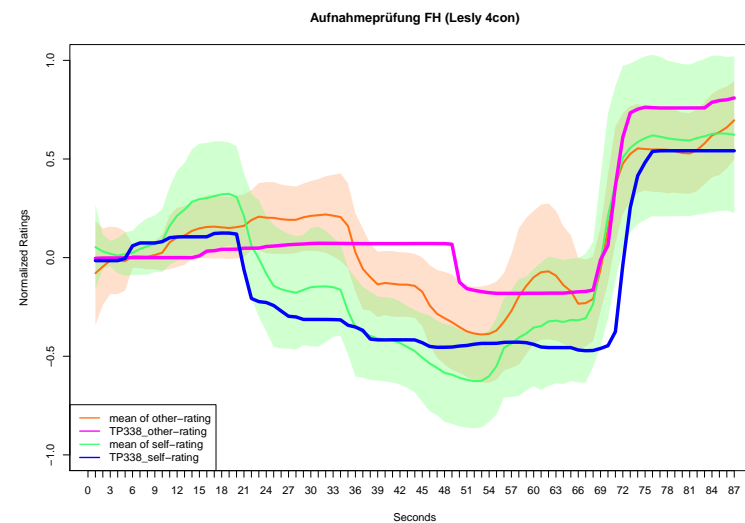
Figure 52: G1 Lesly 4con: Ratings by test-person 338.
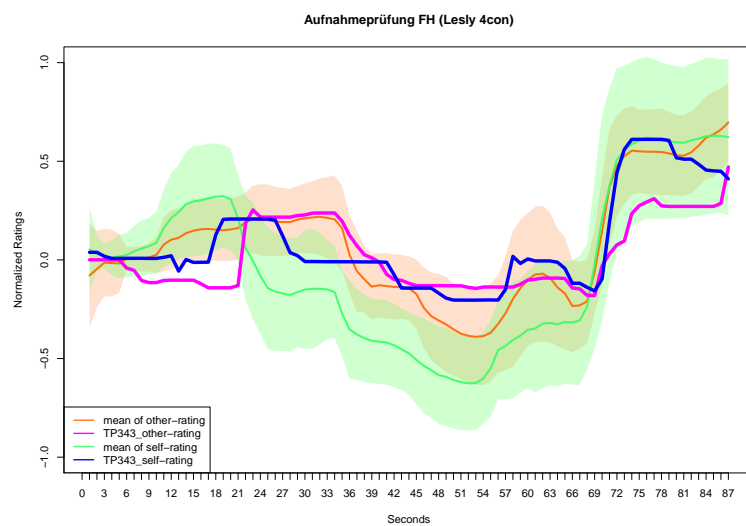
Figure 53: G1 Lesly 4con: Ratings by test-person 343.
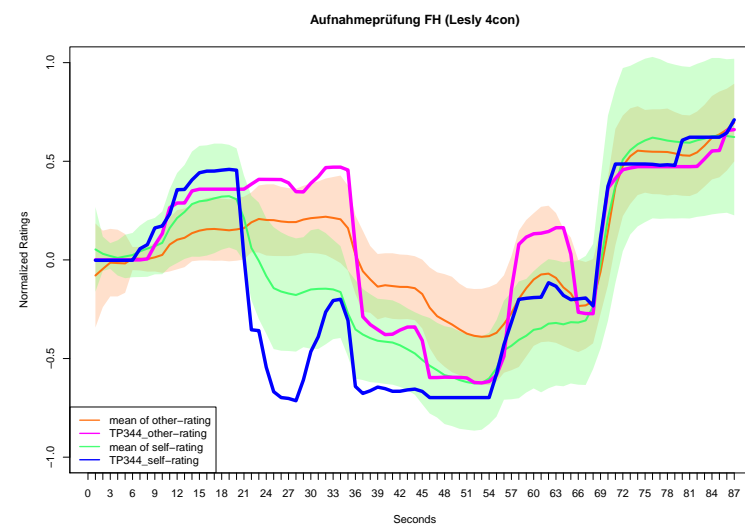
Figure 54: G1 Lesly 4con: Ratings by test-person 344.
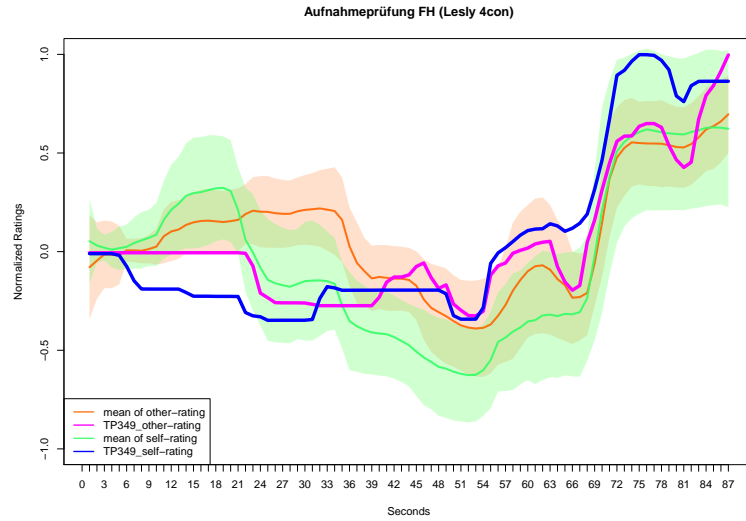
Figure 55: G1 Lesly 4con: Ratings by test-person 349.
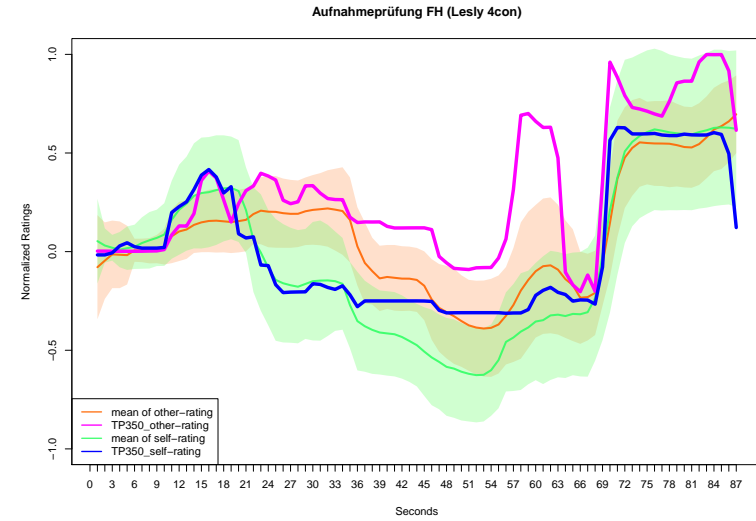
Figure 56: G1 Lesly 4con: Ratings by test-person 350.
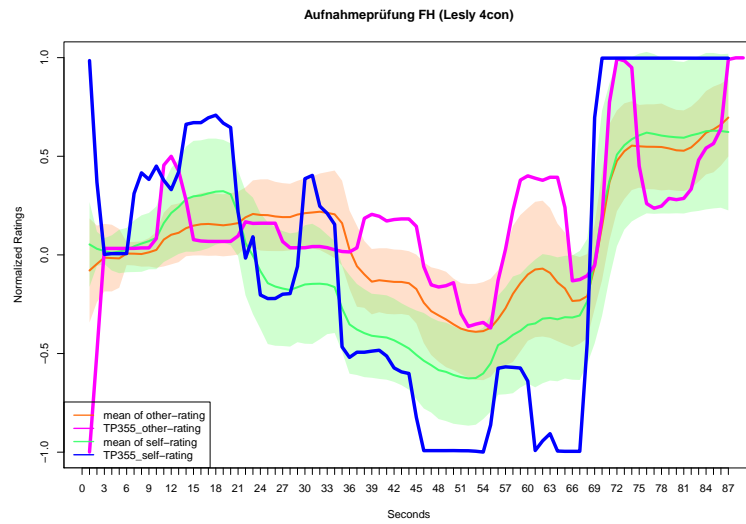
Figure 57: G1 Lesly 4con: Ratings by test-person 355.
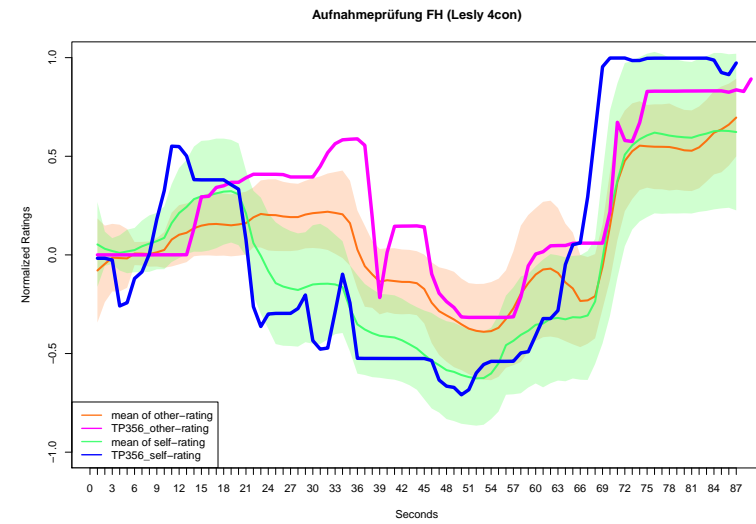
Figure 58: G1 Lesly 4con: Ratings by test-person 356.
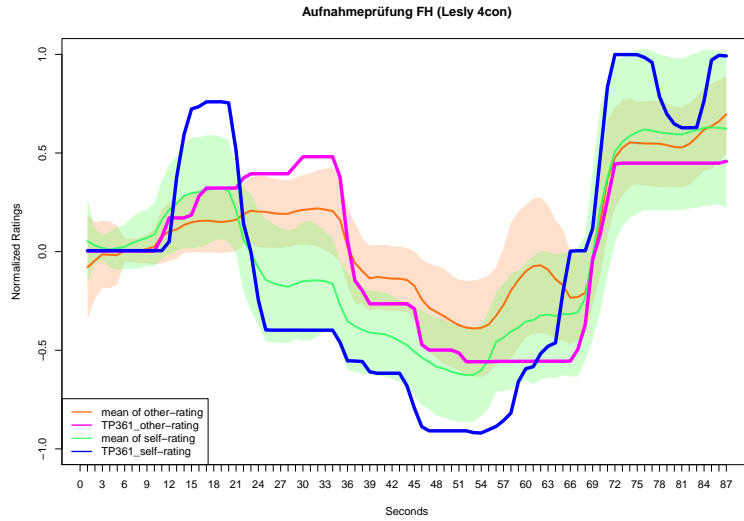
Figure 59: G1 Lesly 4con: Ratings by test-person 361.
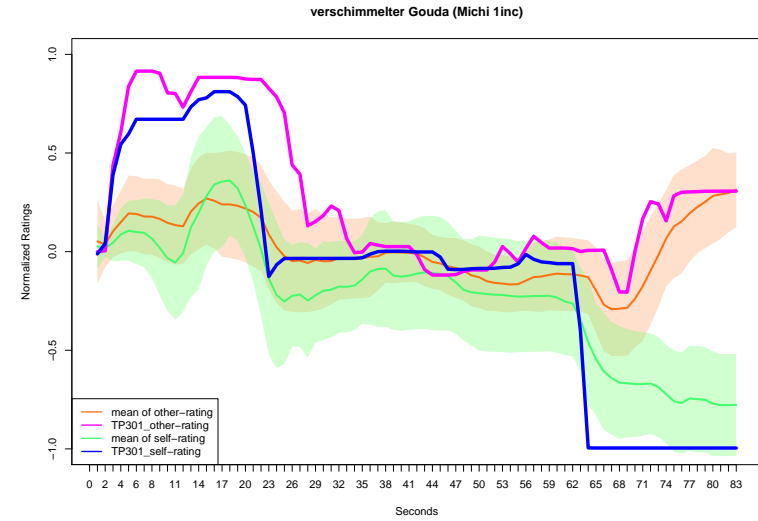
Figure 60: G1 Michi 1inc: Ratings by test-person 301.
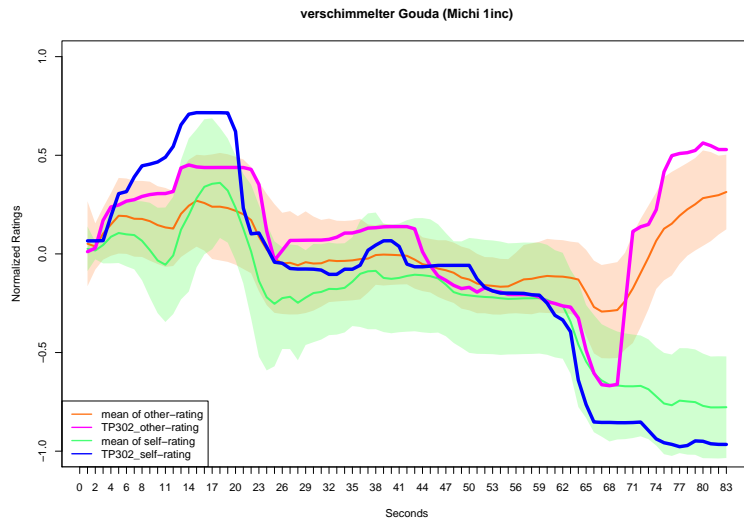
Figure 61: G1 Michi 1inc: Ratings by test-person 302.
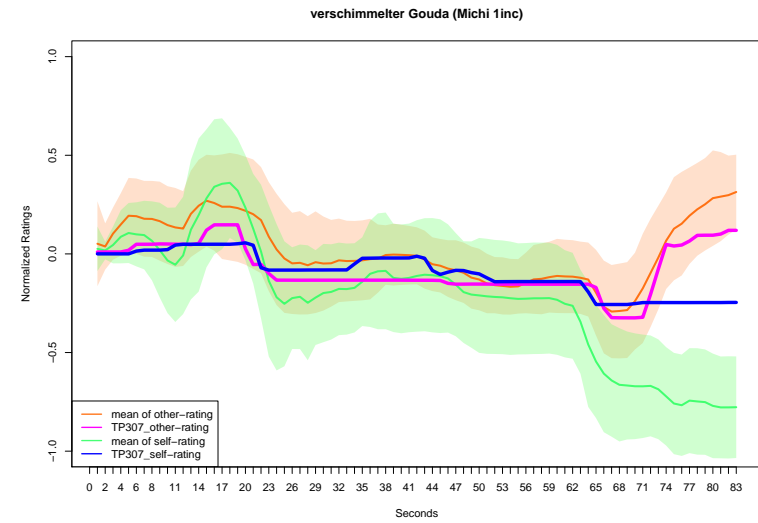
Figure 62: G1 Michi 1inc: Ratings by test-person 307.
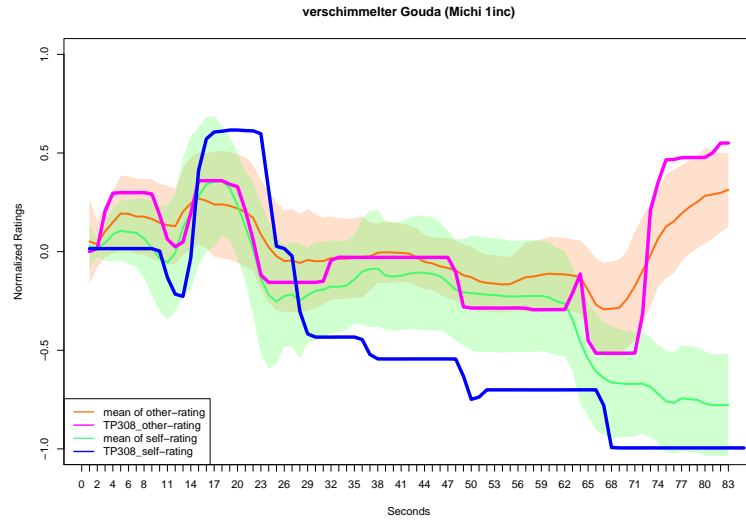
Figure 63: G1 Michi 1inc: Ratings by test-person 308.
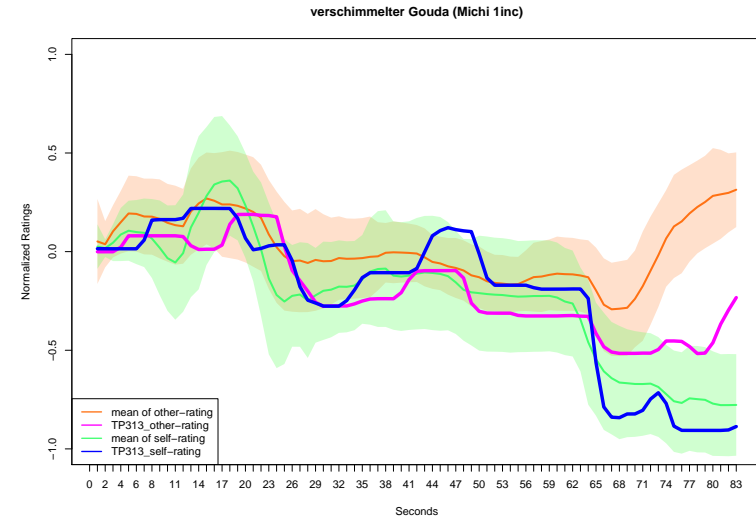
Figure 64: G1 Michi 1inc: Ratings by test-person 313.
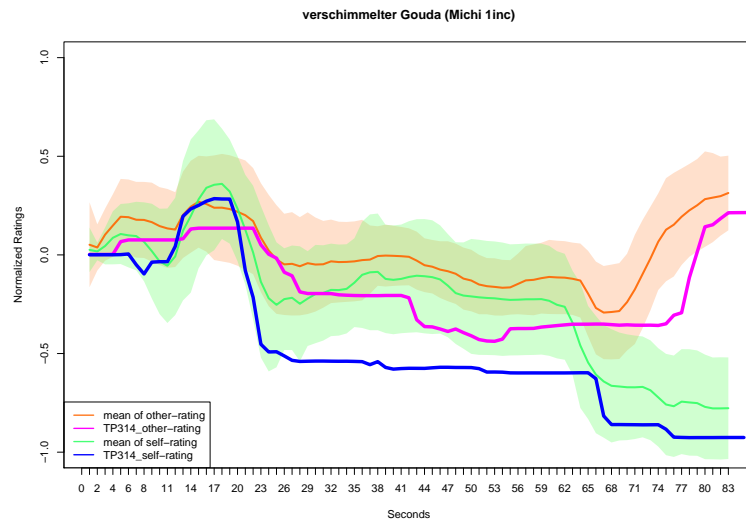
Figure 65: G1 Michi 1inc: Ratings by test-person 314.
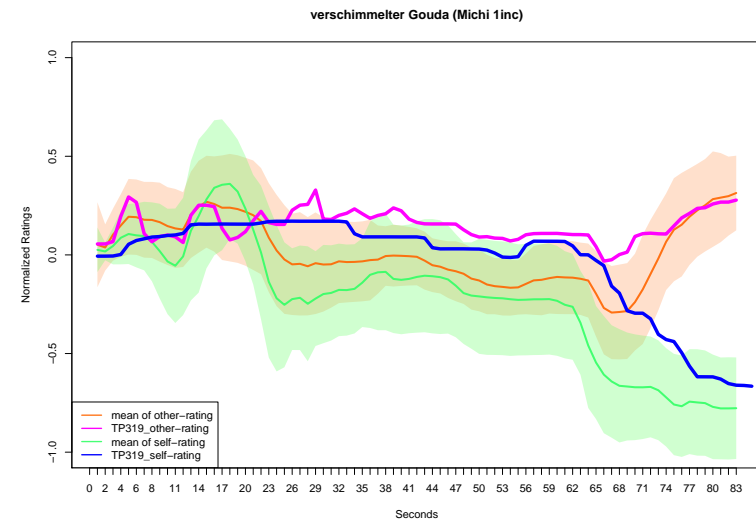
Figure 66: G1 Michi 1inc: Ratings by test-person 319.
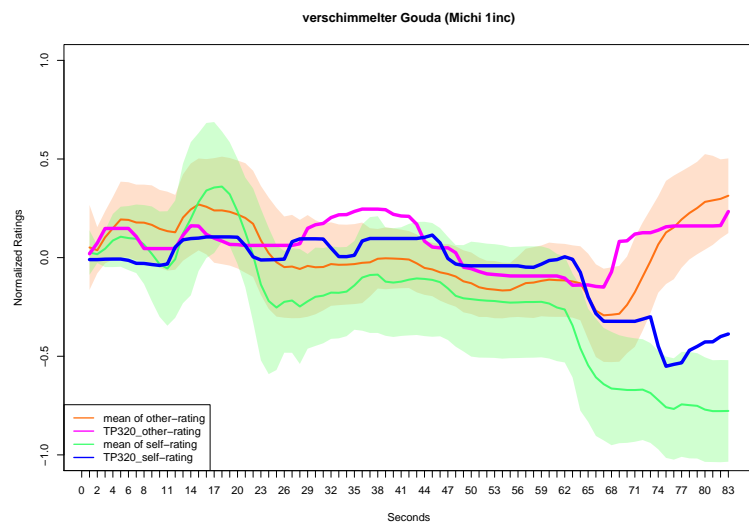
Figure 67: G1 Michi 1inc: Ratings by test-person 320.
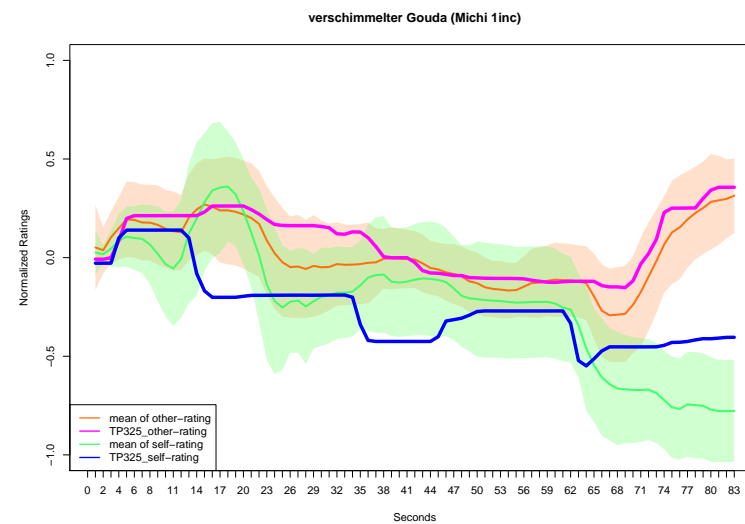
Figure 68: G1 Michi 1inc: Ratings by test-person 325.
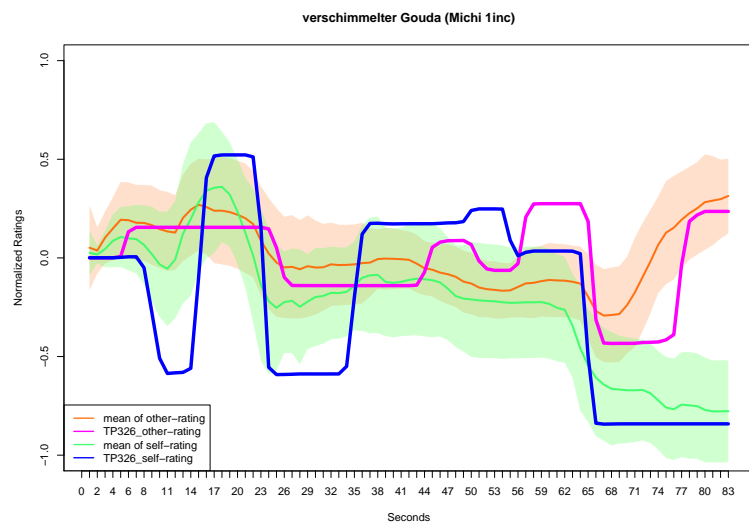
Figure 69: G1 Michi 1inc: Ratings by test-person 326.
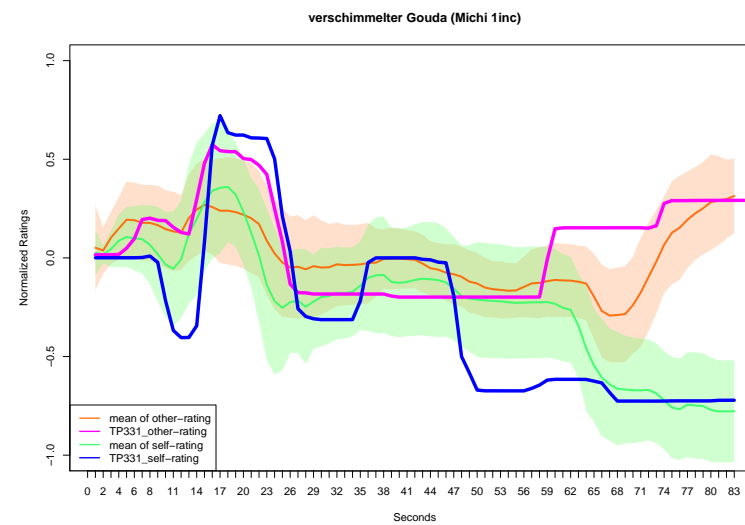
Figure 70: G1 Michi 1inc: Ratings by test-person 331.
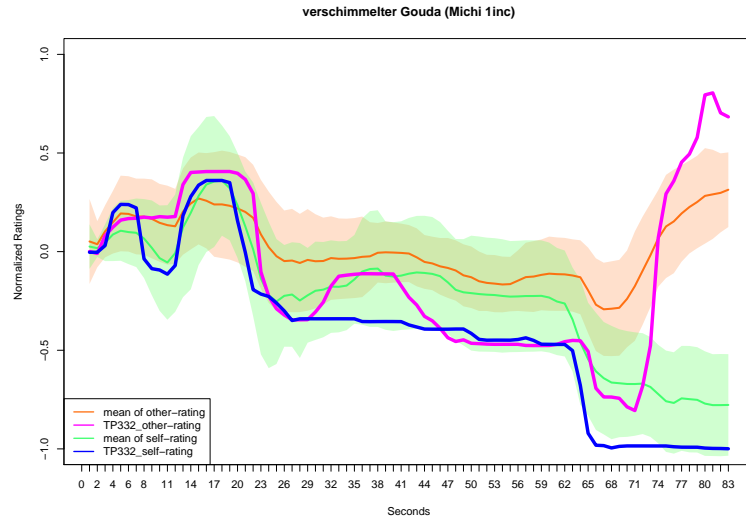
Figure 71: G1 Michi 1inc: Ratings by test-person 332.
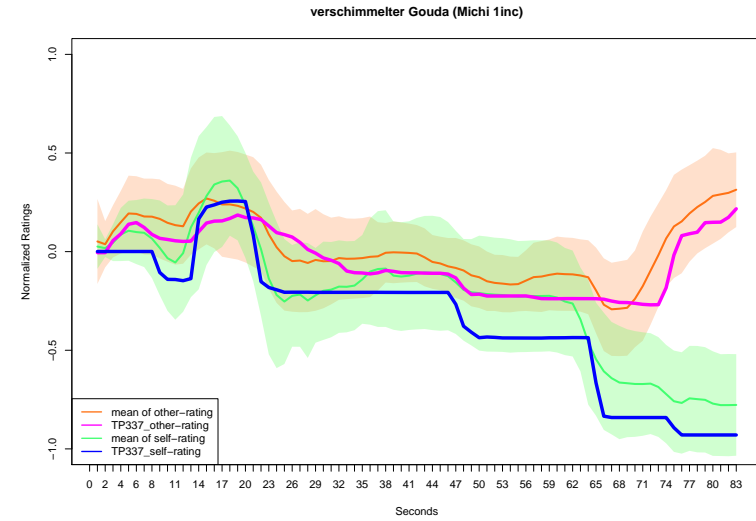
Figure 72: G1 Michi 1inc: Ratings by test-person 337.
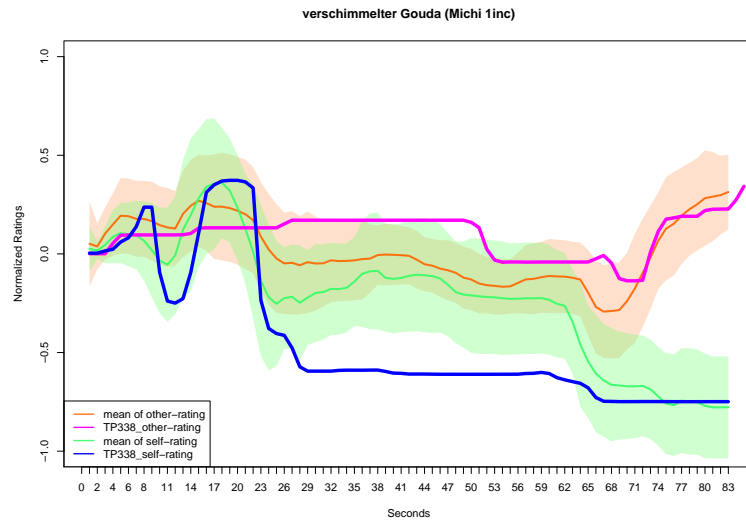
Figure 73: G1 Michi 1inc: Ratings by test-person 338.
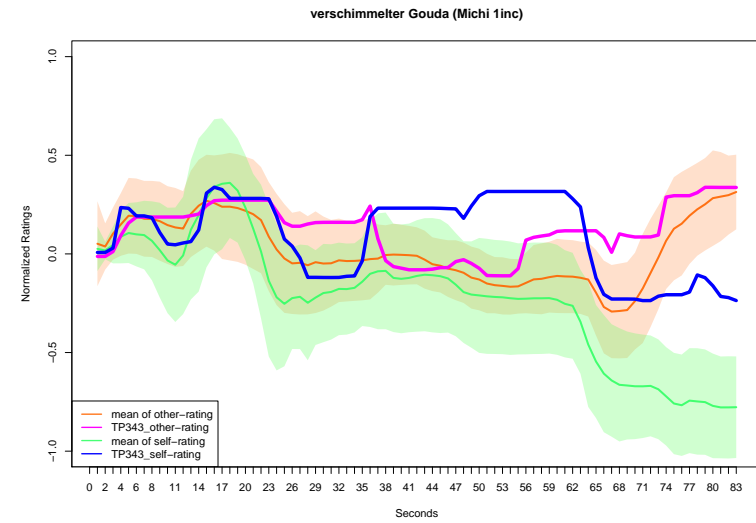
Figure 74: G1 Michi 1inc: Ratings by test-person 343.
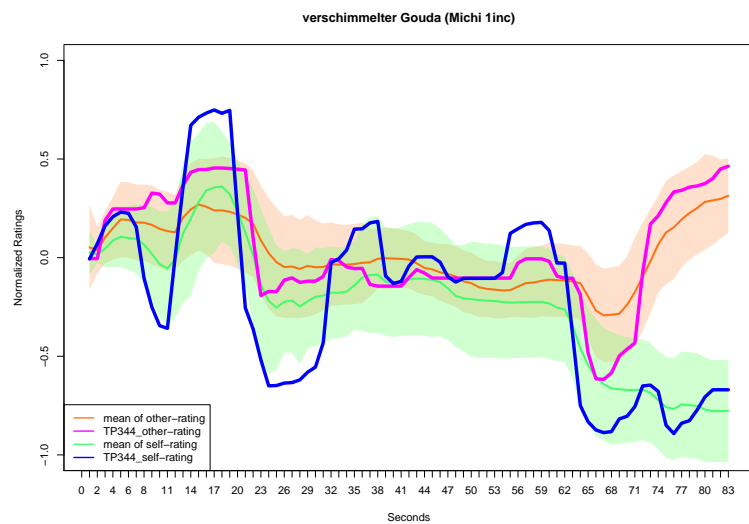
Figure 75: G1 Michi 1inc: Ratings by test-person 344.
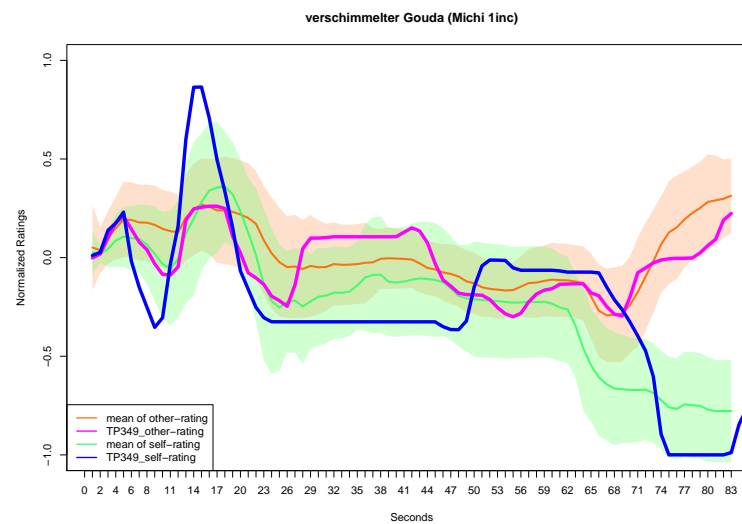
Figure 76: G1 Michi 1inc: Ratings by test-person 349.
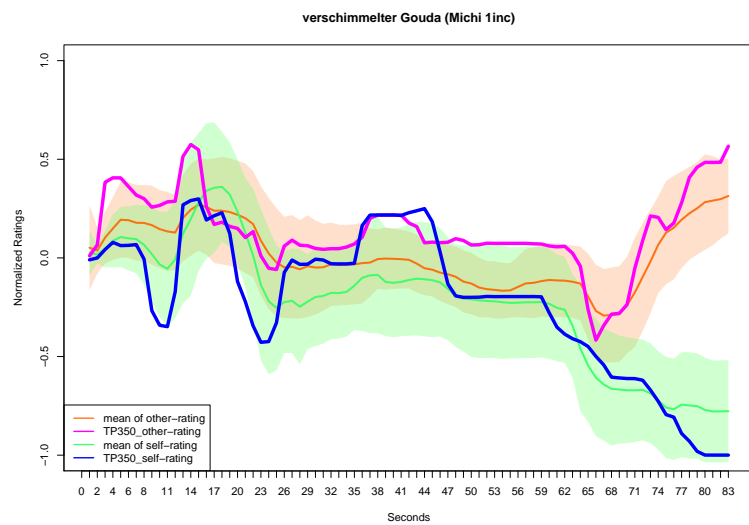
Figure 77: G1 Michi 1inc: Ratings by test-person 350.
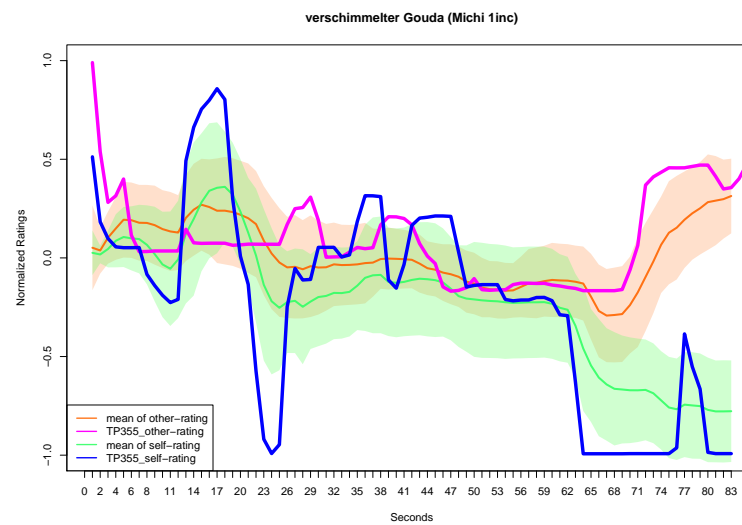
Figure 78: G1 Michi 1inc: Ratings by test-person 355.
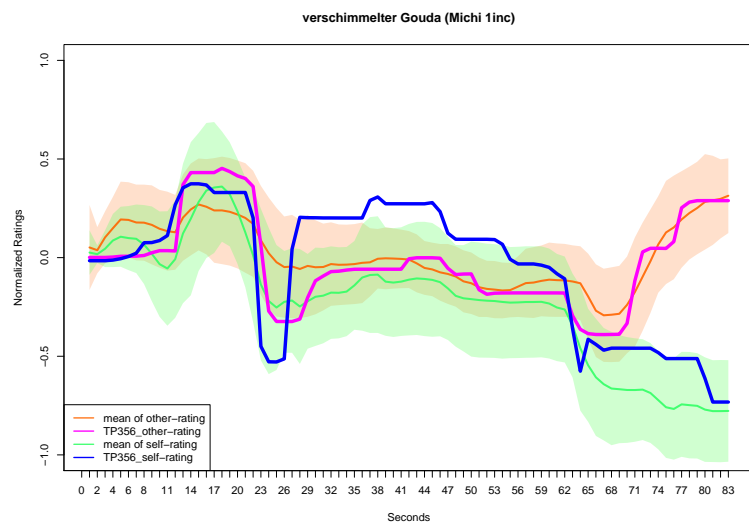
Figure 79: G1 Michi 1inc: Ratings by test-person 356.
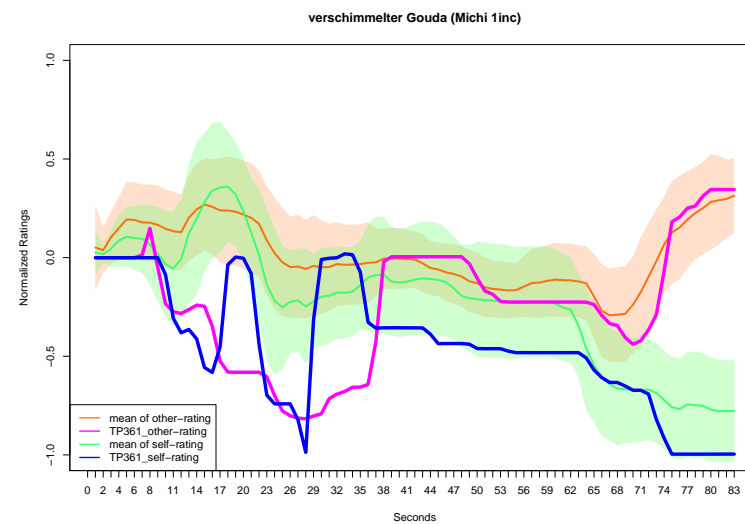
Figure 80: G1 Michi 1inc: Ratings by test-person 361.
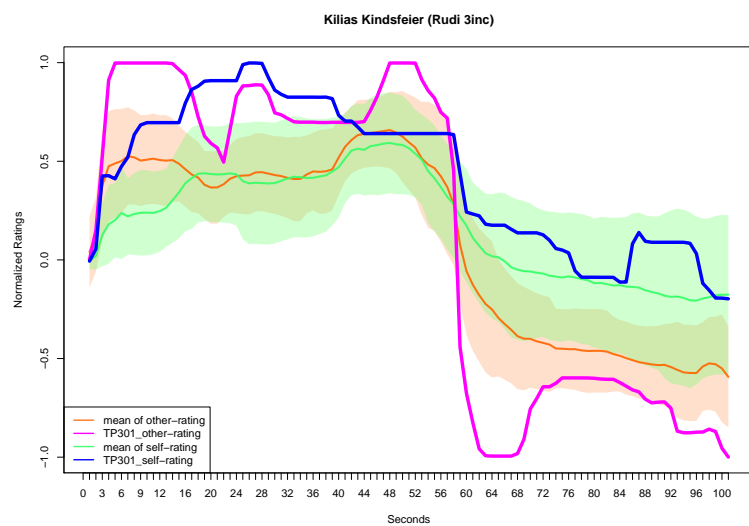
Figure 81: G1 Rudi 3inc: Ratings by test-person 301.
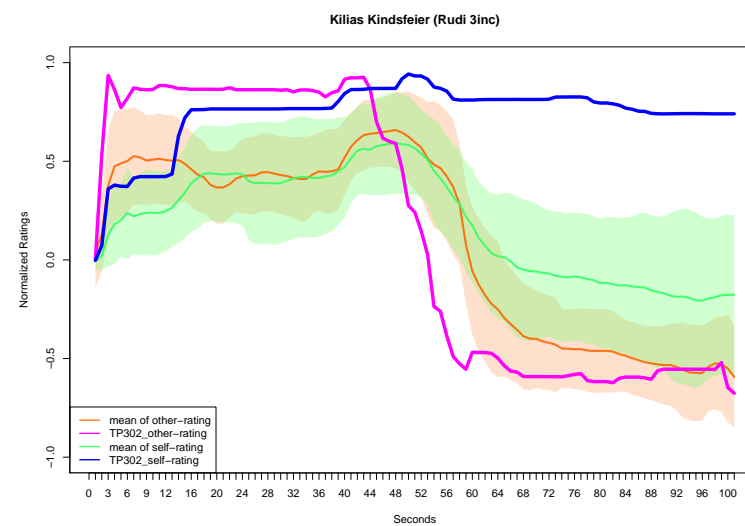
Figure 82: G1 Rudi 3inc: Ratings by test-person 302.
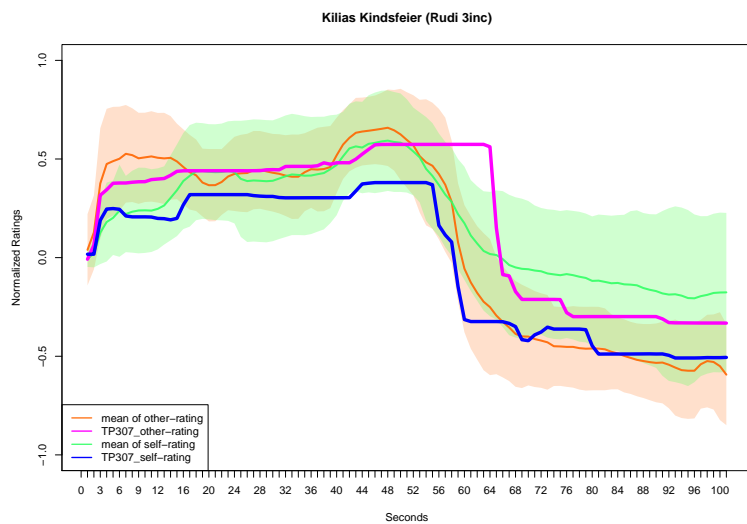
Figure 83: G1 Rudi 3inc: Ratings by test-person 307.
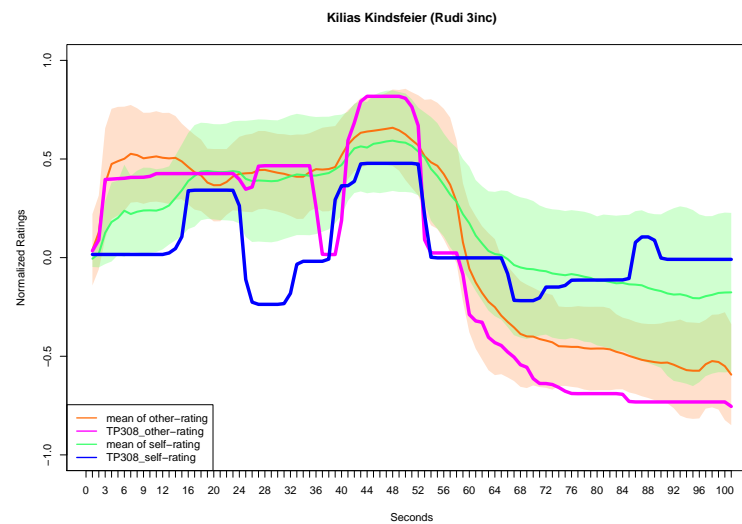
Figure 84: G1 Rudi 3inc: Ratings by test-person 308.
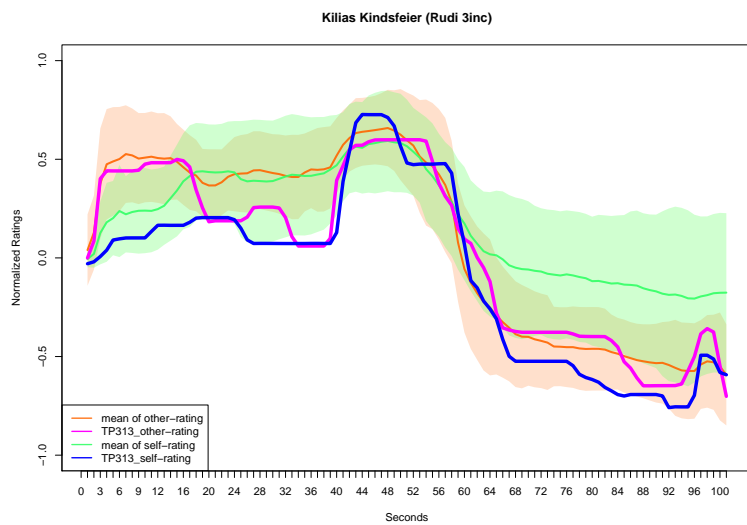
Figure 85: G1 Rudi 3inc: Ratings by test-person 313.
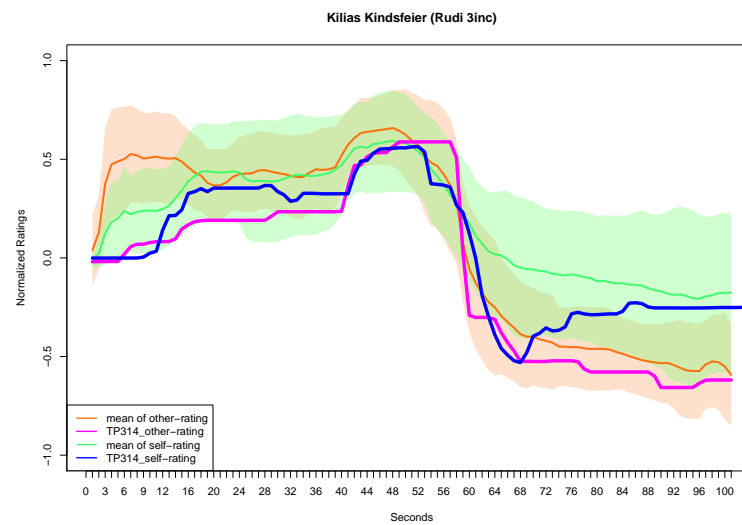
Figure 86: G1 Rudi 3inc: Ratings by test-person 314.
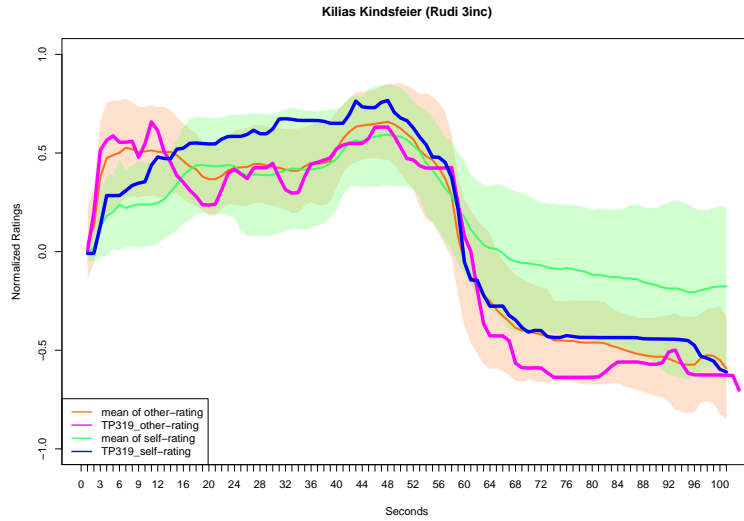
Figure 87: G1 Rudi 3inc: Ratings by test-person 319.
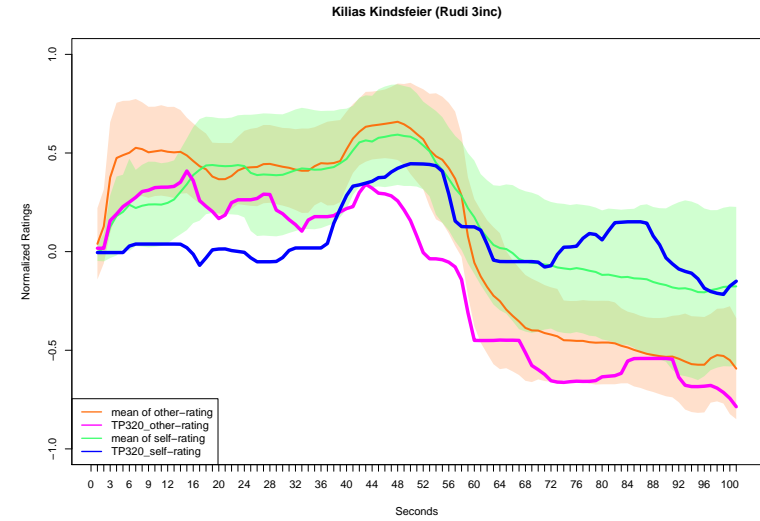
Figure 88: G1 Rudi 3inc: Ratings by test-person 320.
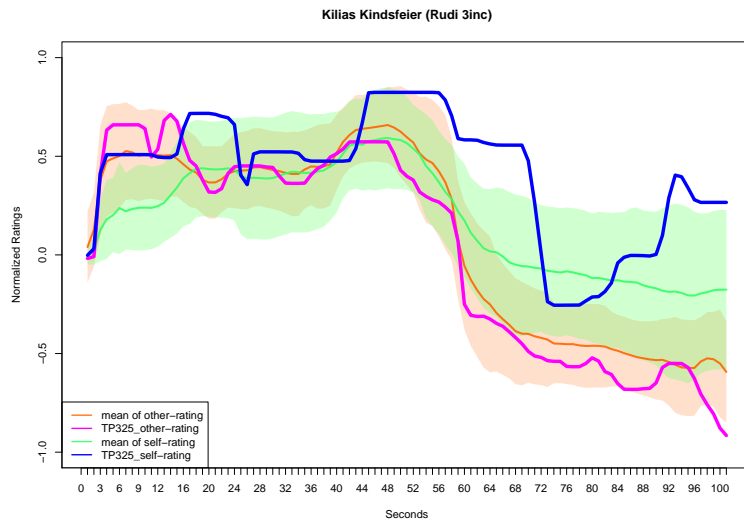
Figure 89: G1 Rudi 3inc: Ratings by test-person 325.
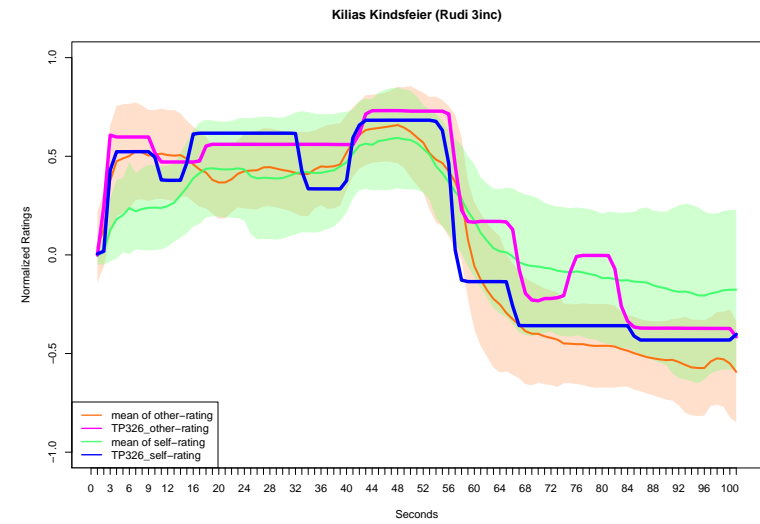
Figure 90: G1 Rudi 3inc: Ratings by test-person 326.
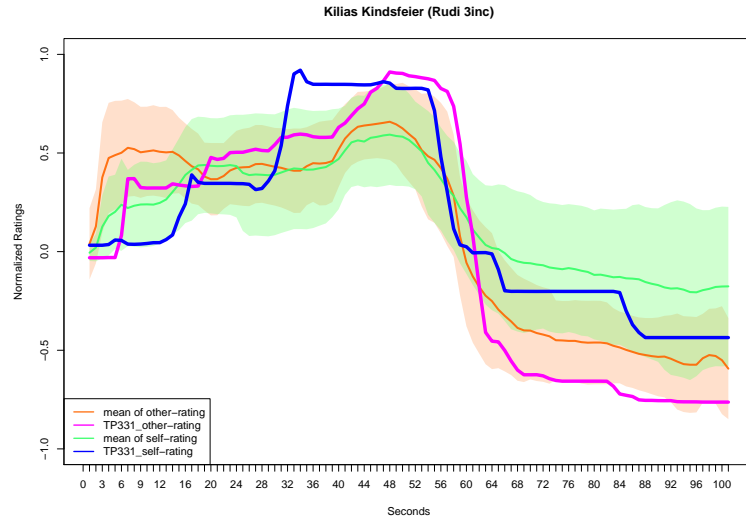
Figure 91: G1 Rudi 3inc: Ratings by test-person 331.
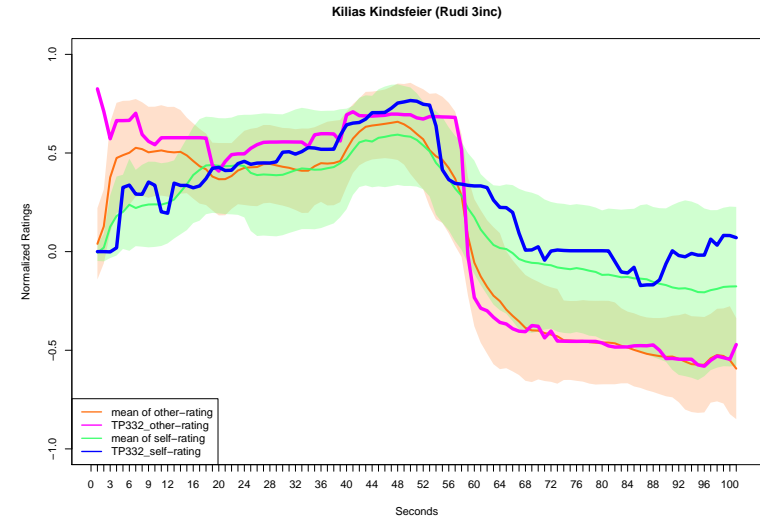
Figure 92: G1 Rudi 3inc: Ratings by test-person 332.
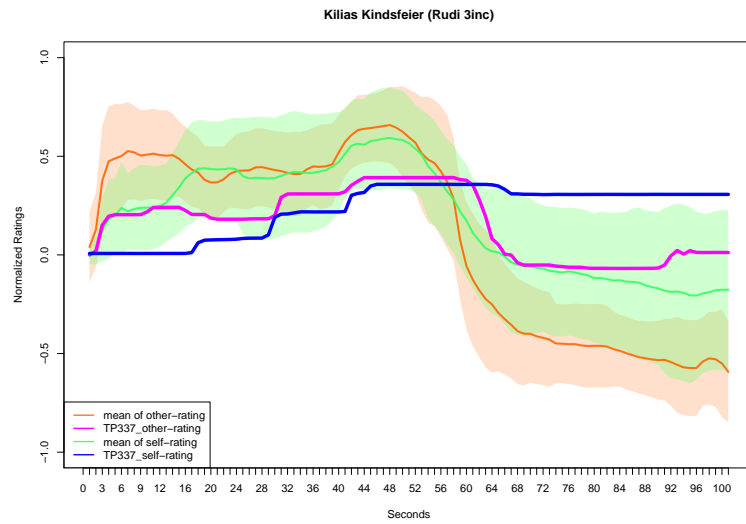
Figure 93: G1 Rudi 3inc: Ratings by test-person 337.
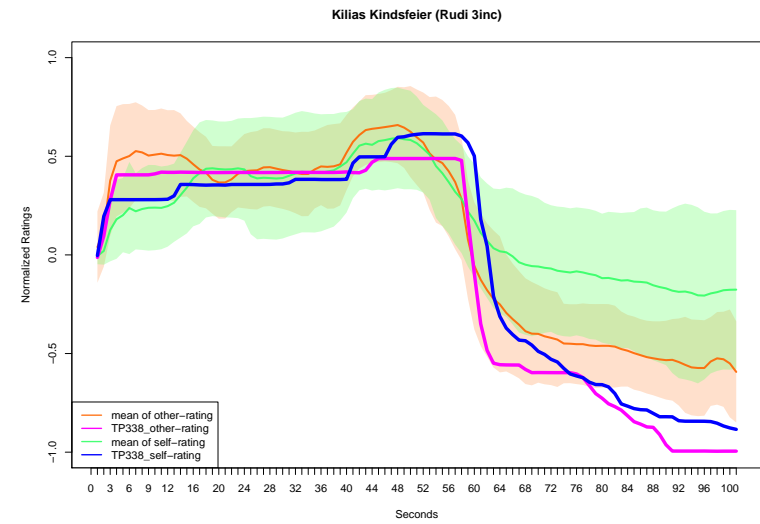
Figure 94: G1 Rudi 3inc: Ratings by test-person 338.
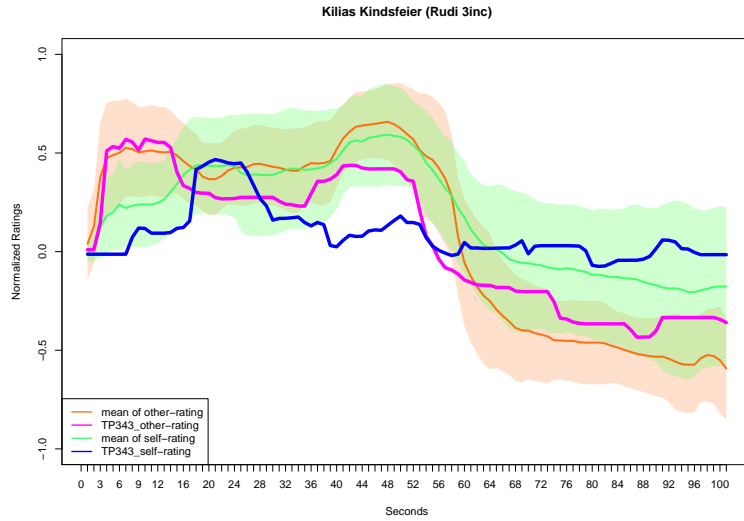
Figure 95: G1 Rudi 3inc: Ratings by test-person 343.
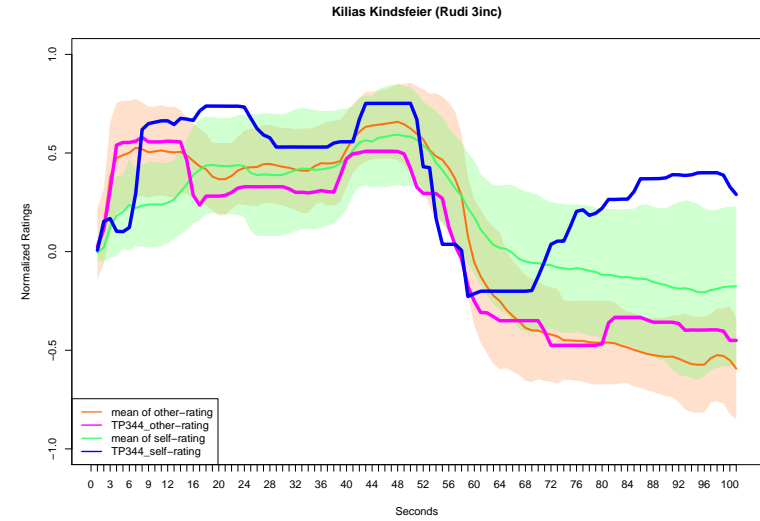
Figure 96: G1 Rudi 3inc: Ratings by test-person 344.
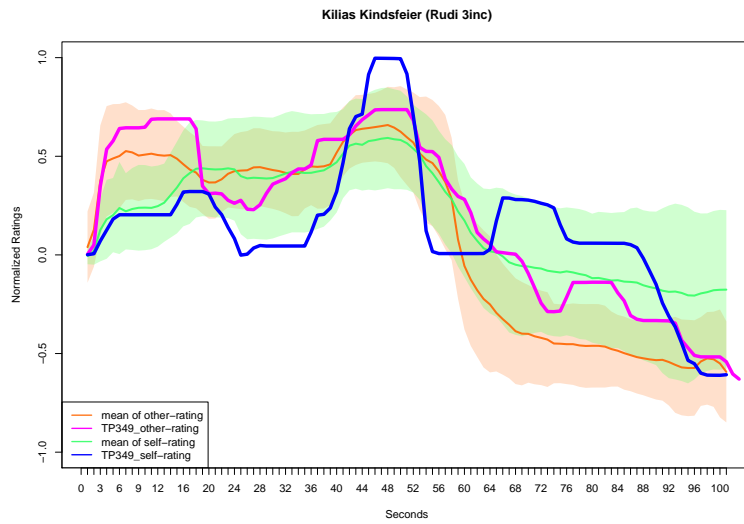
Figure 97: G1 Rudi 3inc: Ratings by test-person 349.
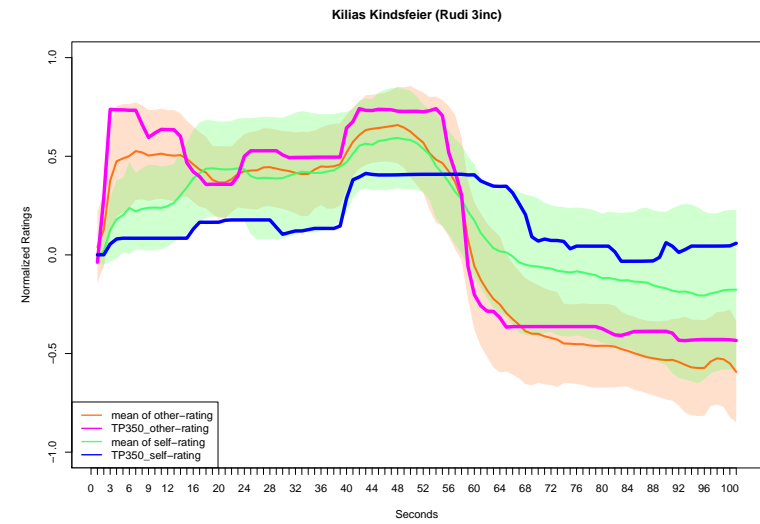
Figure 98: G1 Rudi 3inc: Ratings by test-person 350.
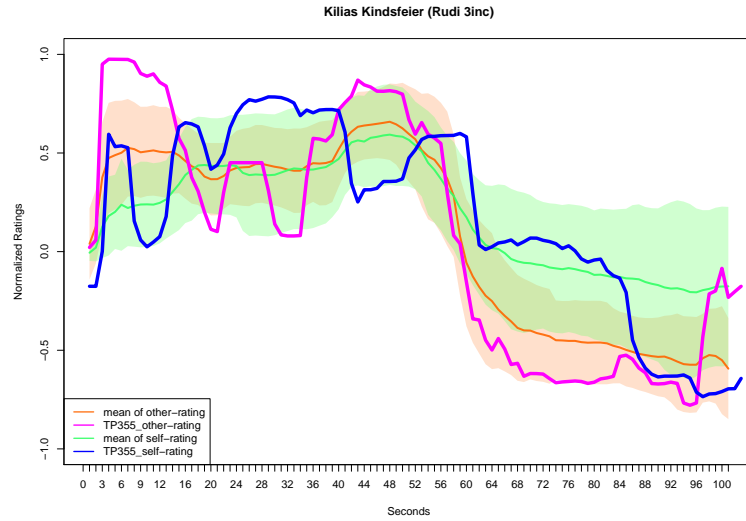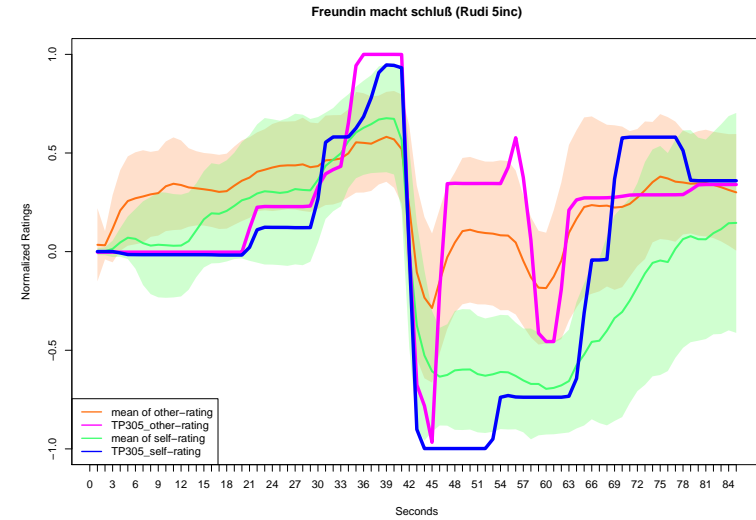
Figure 99: G1 Rudi 3inc: Ratings by test-person 355.
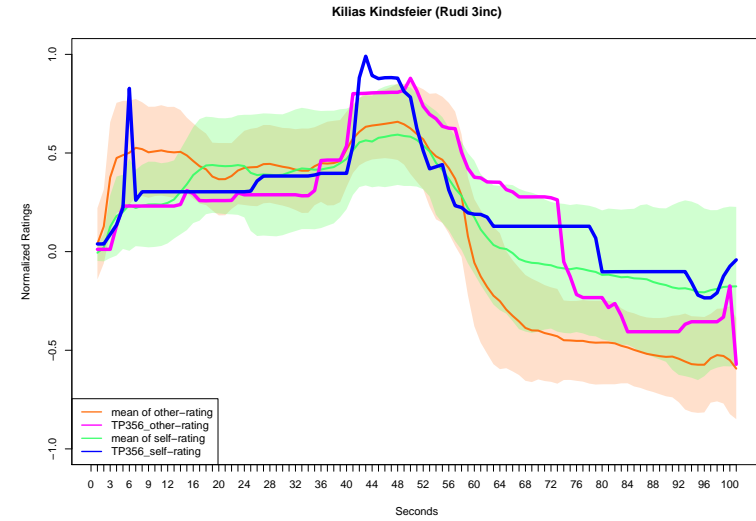
Figure 100: G1 Rudi 3inc: Ratings by test-person 356.
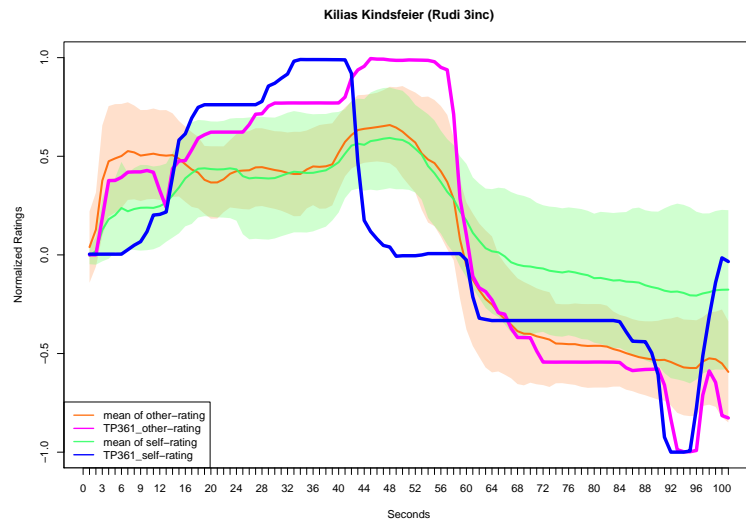
Figure 101: G1 Rudi 3inc: Ratings by test-person 361.
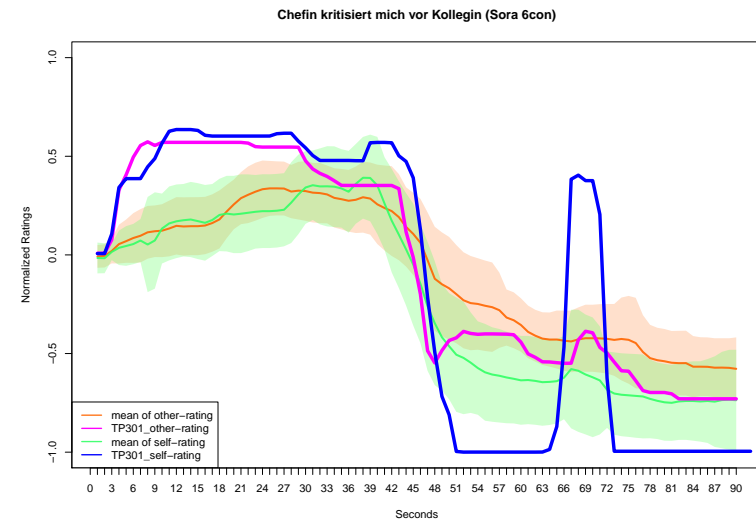
Figure 102: G1 Sora 6con: Ratings by test-person 301.
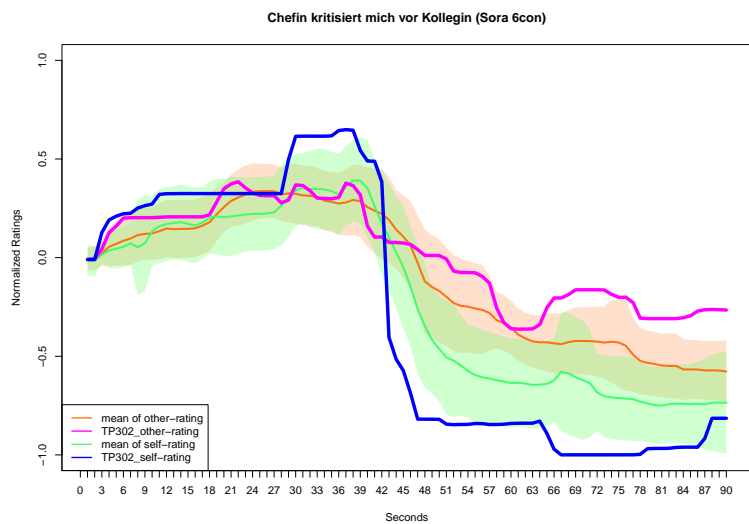
Figure 103: G1 Sora 6con: Ratings by test-person 302.
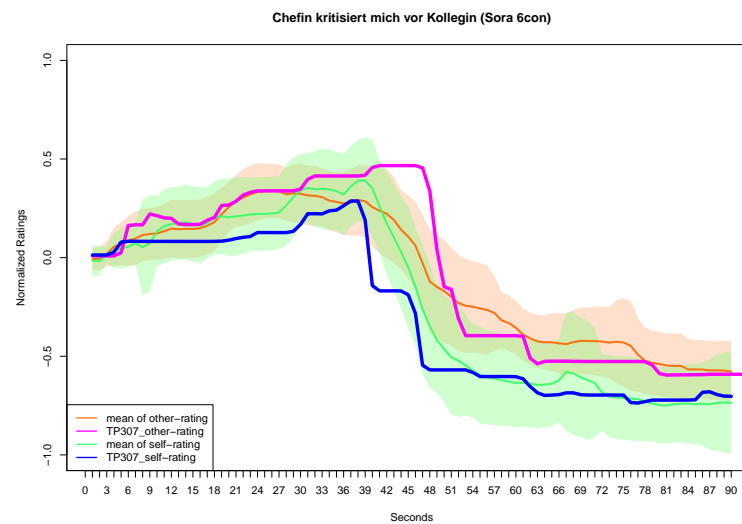
Figure 104: G1 Sora 6con: Ratings by test-person 307.
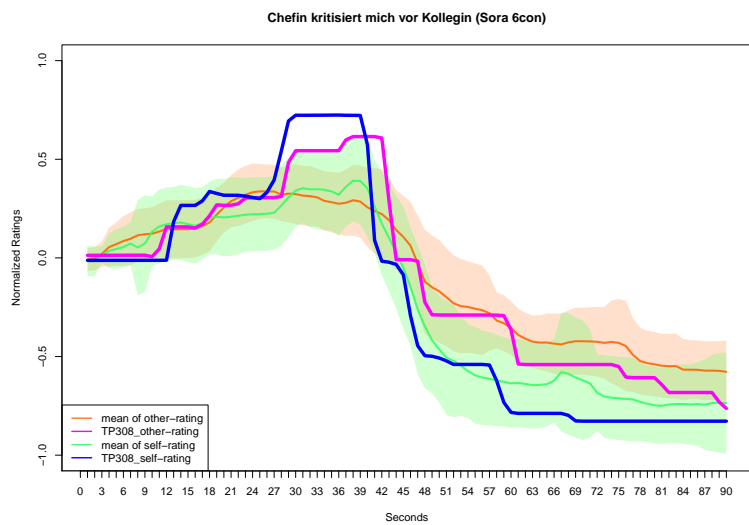
Figure 105: G1 Sora 6con: Ratings by test-person 308.
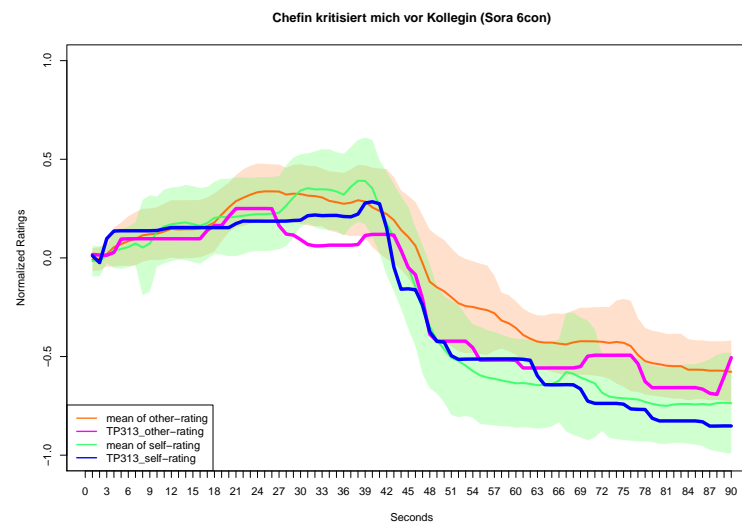
Figure 106: G1 Sora 6con: Ratings by test-person 313.
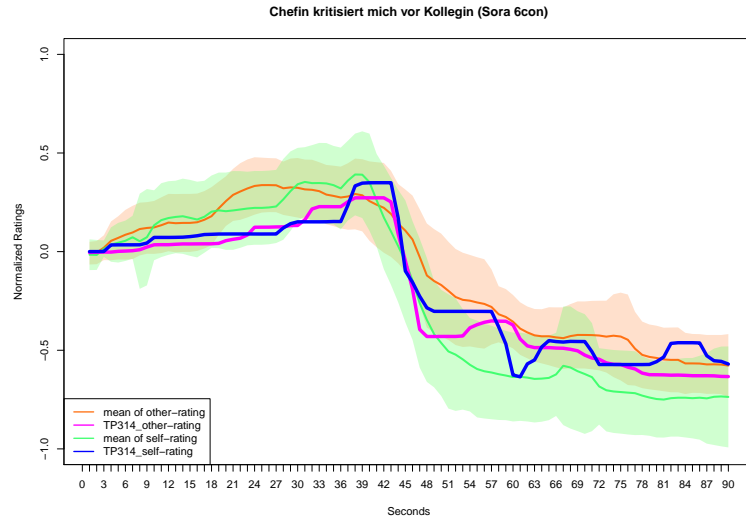
Figure 107: G1 Sora 6con: Ratings by test-person 314.
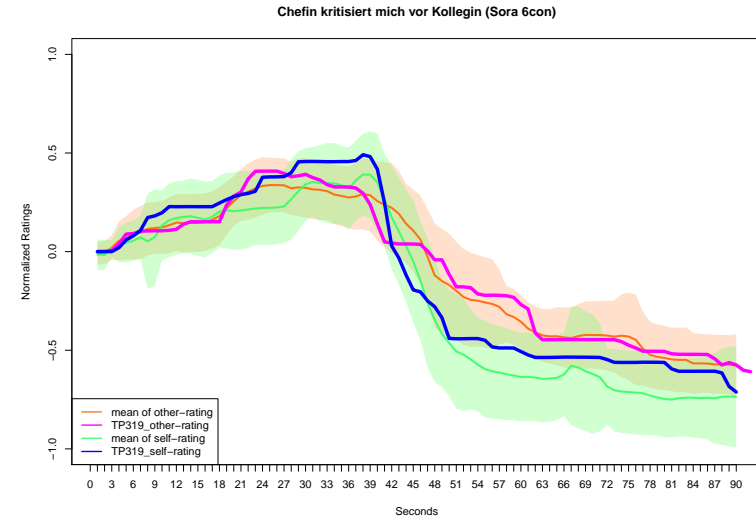
Figure 108: G1 Sora 6con: Ratings by test-person 319.
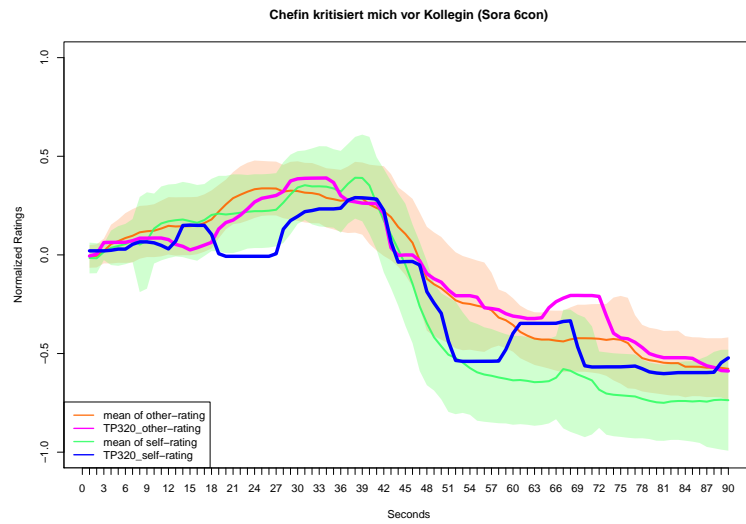
Figure 109: G1 Sora 6con: Ratings by test-person 320.
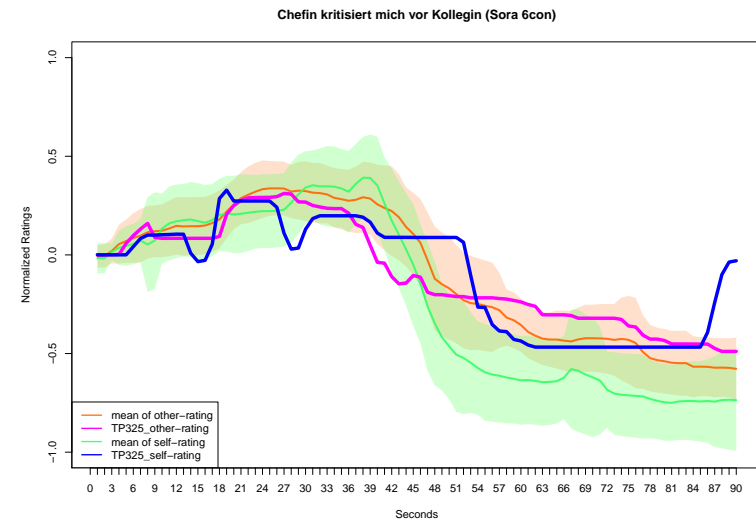
Figure 110: G1 Sora 6con: Ratings by test-person 325.
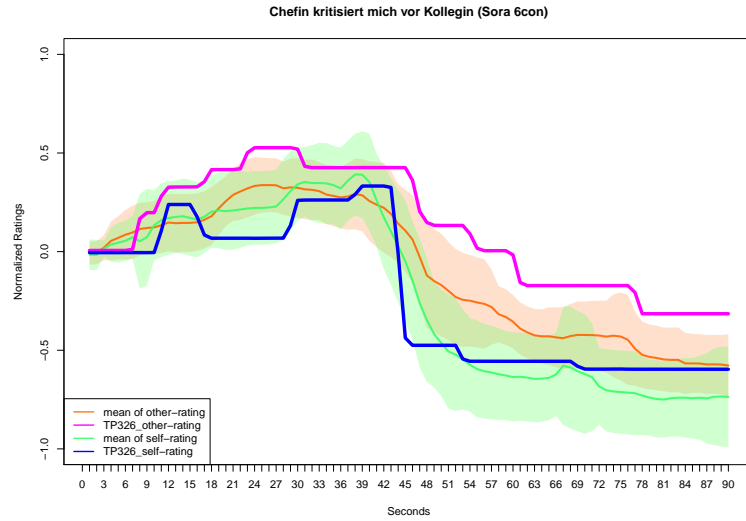
Figure 111: G1 Sora 6con: Ratings by test-person 326.
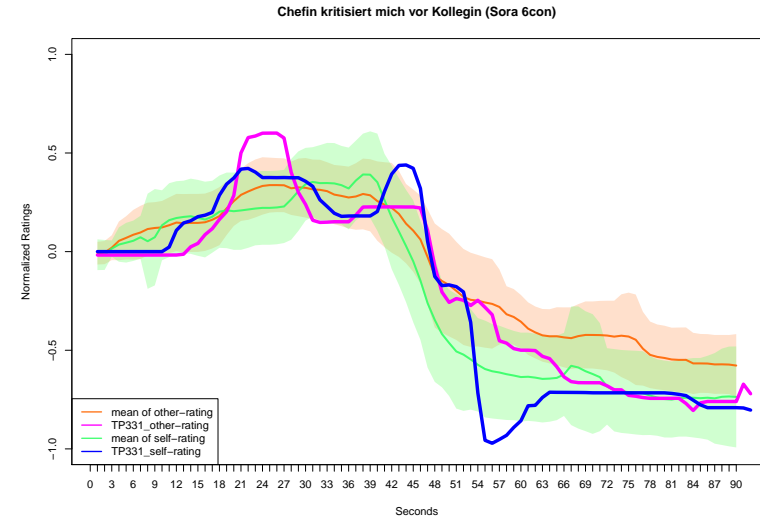
Figure 112: G1 Sora 6con: Ratings by test-person 331.
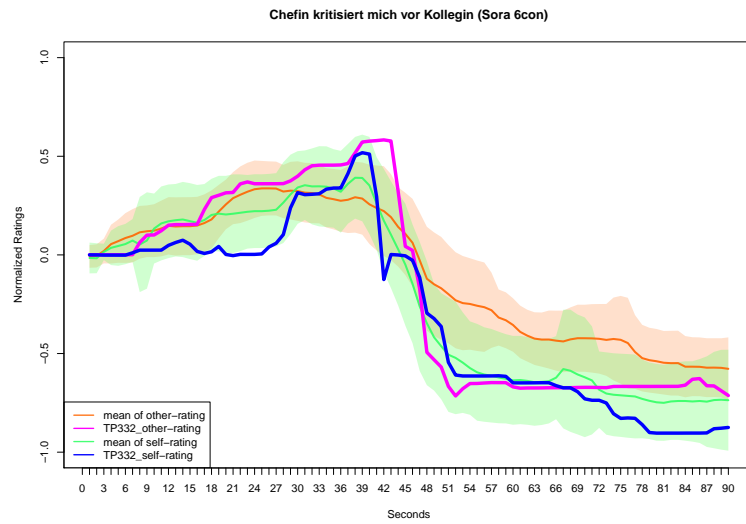
Figure 113: G1 Sora 6con: Ratings by test-person 332.
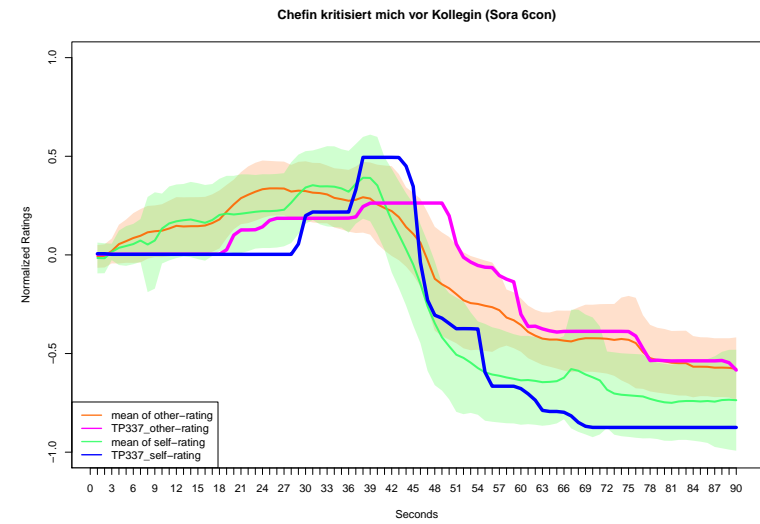
Figure 114: G1 Sora 6con: Ratings by test-person 337.
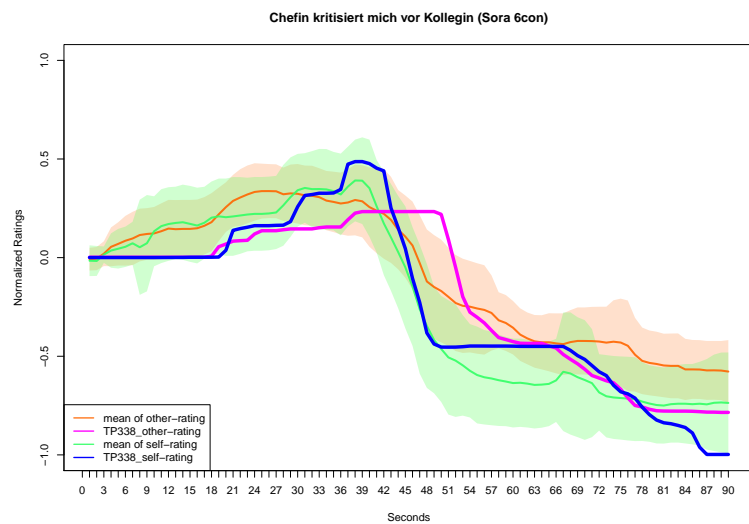
Figure 115: G1 Sora 6con: Ratings by test-person 338.
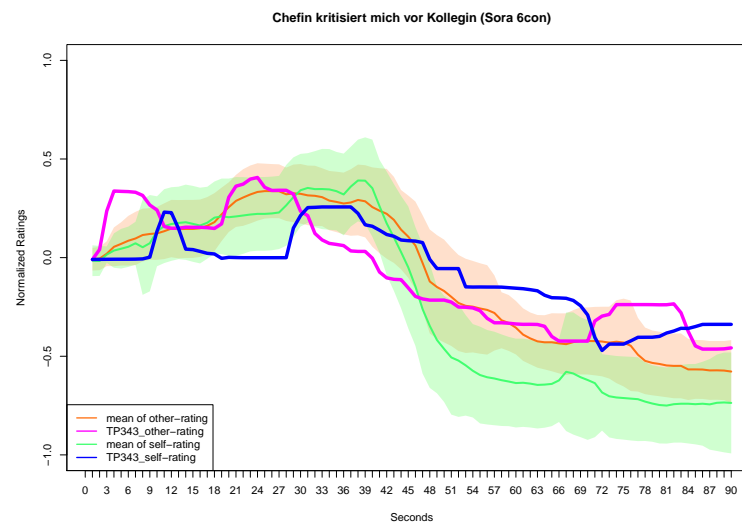
Figure 116: G1 Sora 6con: Ratings by test-person 343.
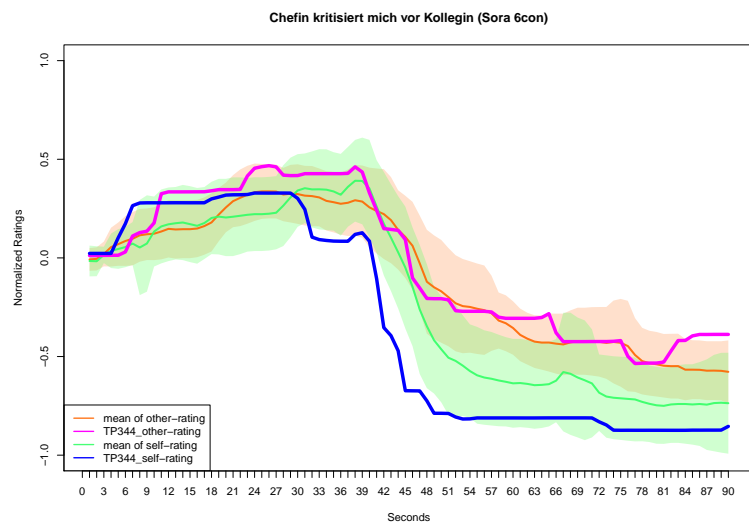
Figure 117: G1 Sora 6con: Ratings by test-person 344.
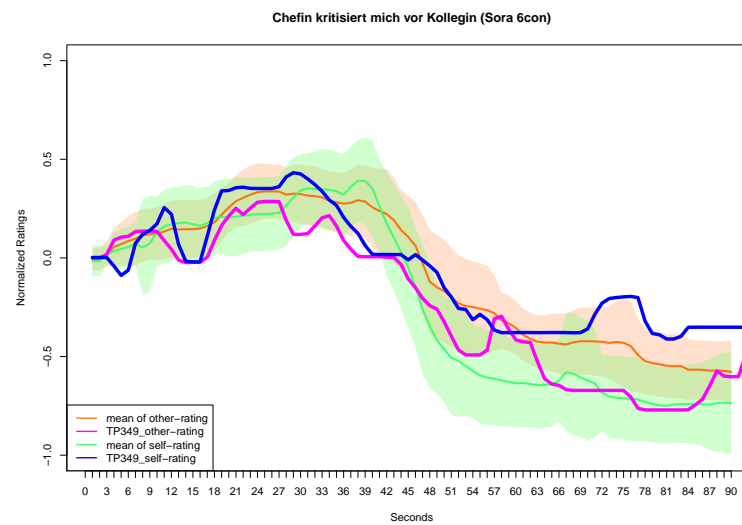
Figure 118: G1 Sora 6con: Ratings by test-person 349.
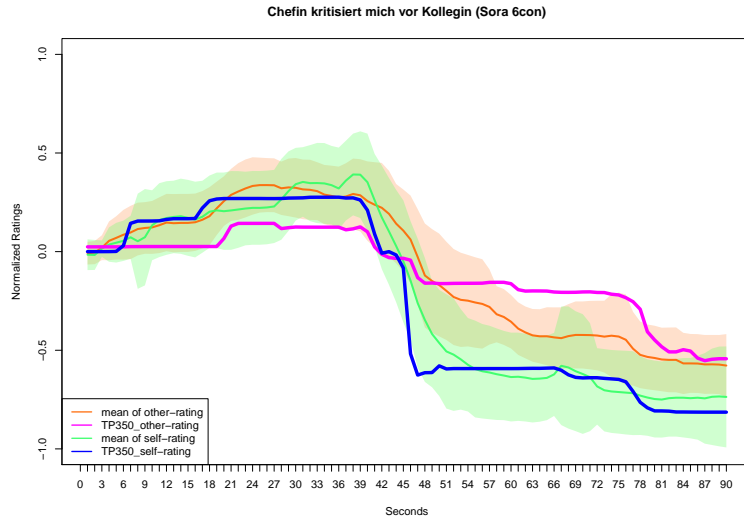
Figure 119: G1 Sora 6con: Ratings by test-person 350.
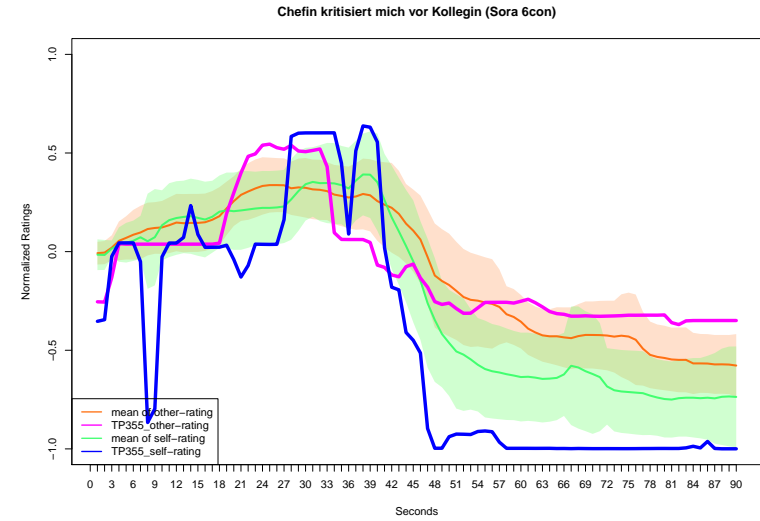
Figure 120: G1 Sora 6con: Ratings by test-person 355.
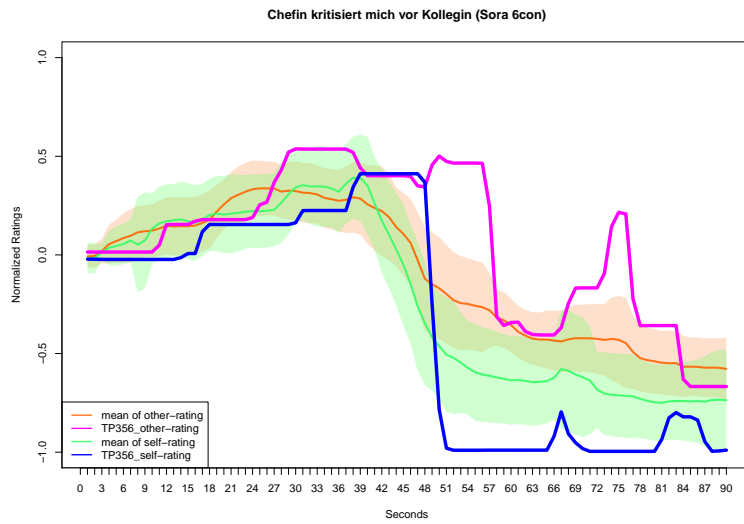
Figure 121: G1 Sora 6con: Ratings by test-person 356.
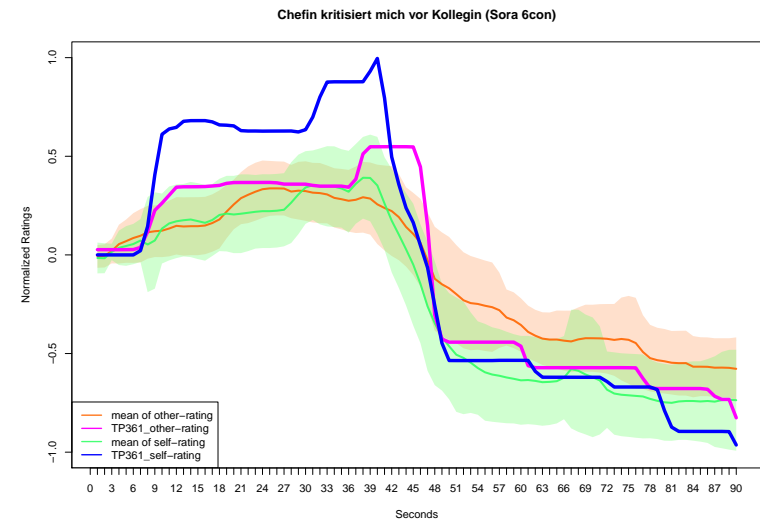
Figure 122: G1 Sora 6con: Ratings by test-person 361.
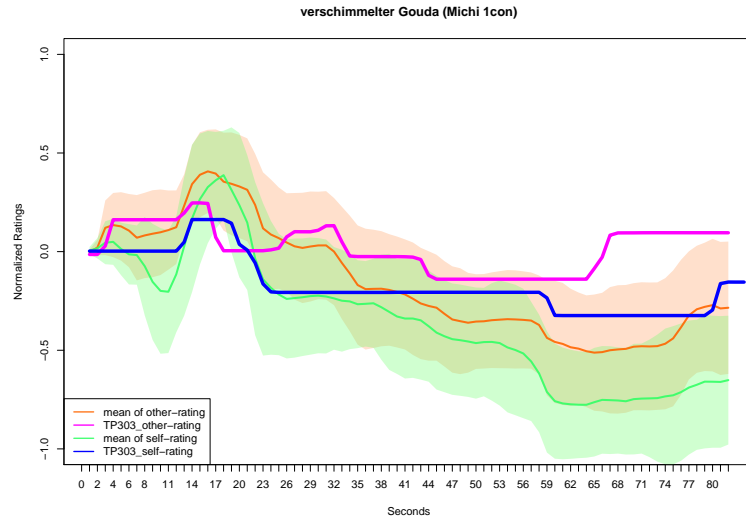
Figure 123: G2 Michi 1con: Ratings by test-person 303.
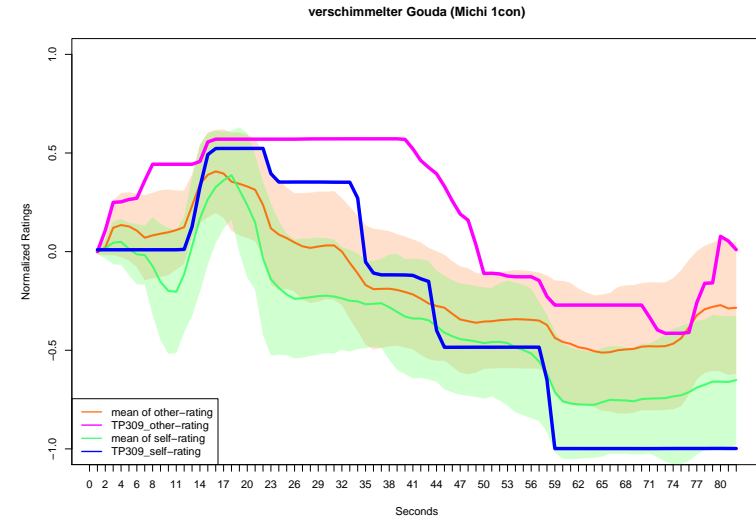
Figure 124: G2 Michi 1con: Ratings by test-person 309.
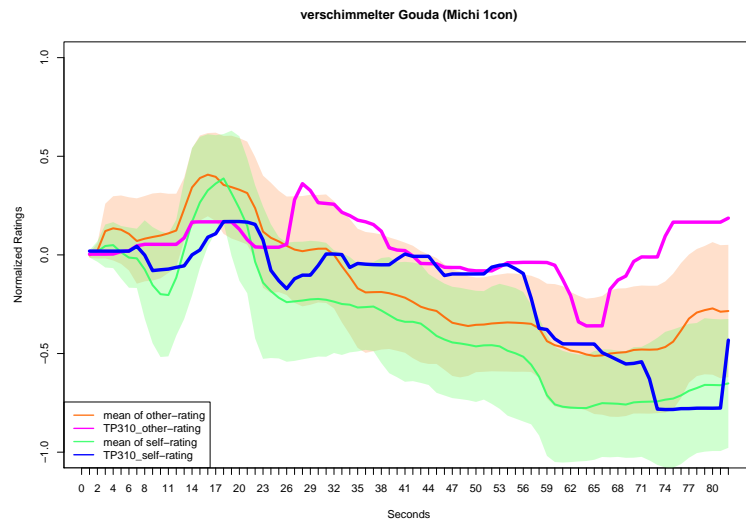
Figure 125: G2 Michi 1con: Ratings by test-person 310.
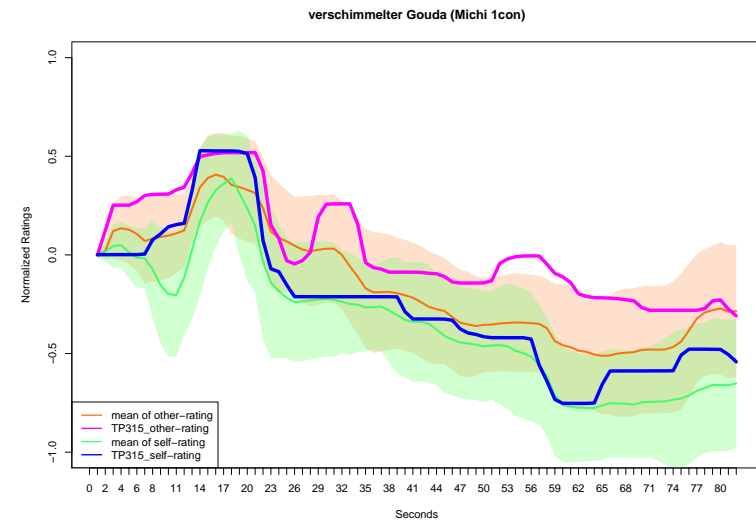
Figure 126: G2 Michi 1con: Ratings by test-person 315.
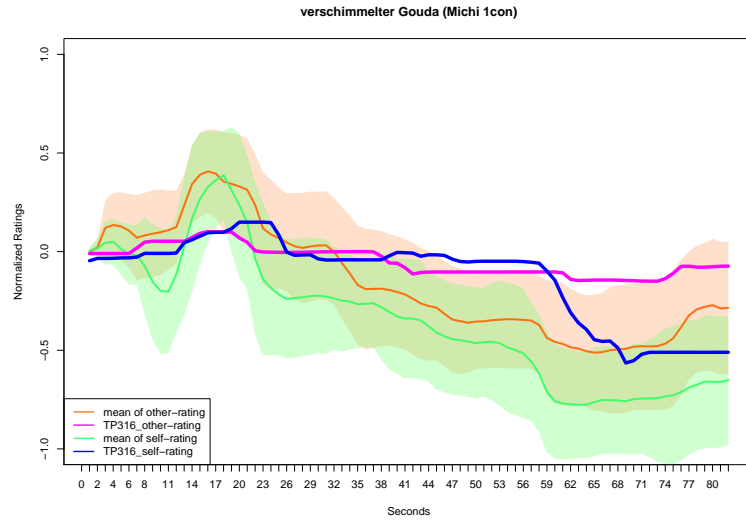
Figure 127: G2 Michi 1con: Ratings by test-person 316.
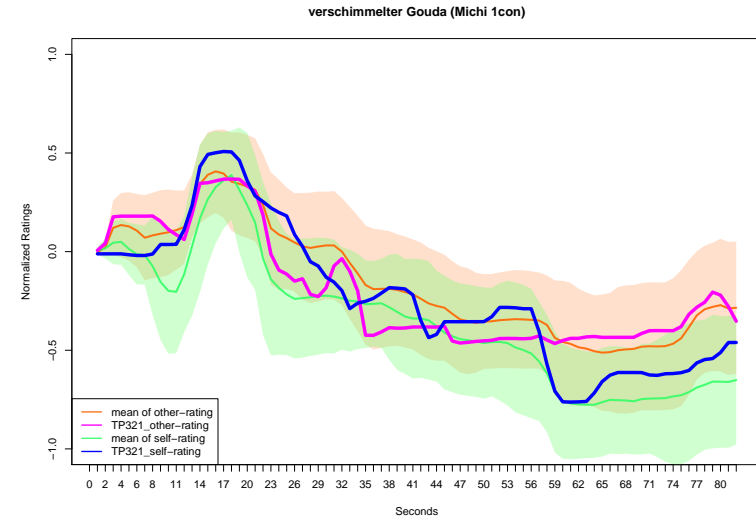
Figure 128: G2 Michi 1con: Ratings by test-person 321.
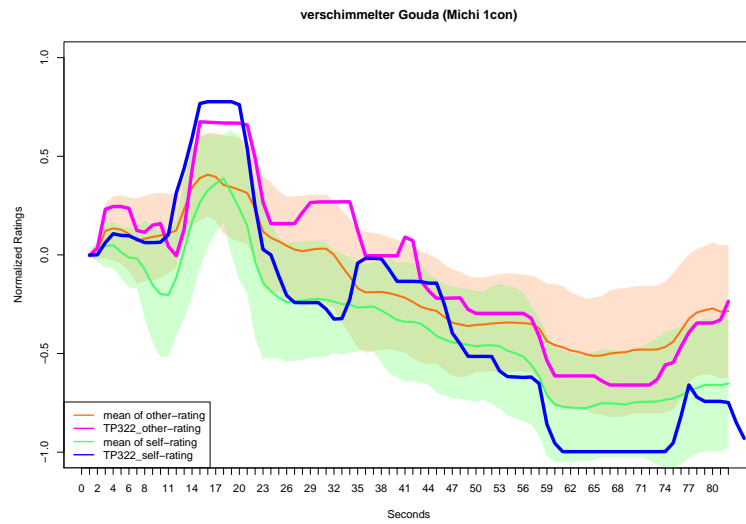
Figure 129: G2 Michi 1con: Ratings by test-person 322.
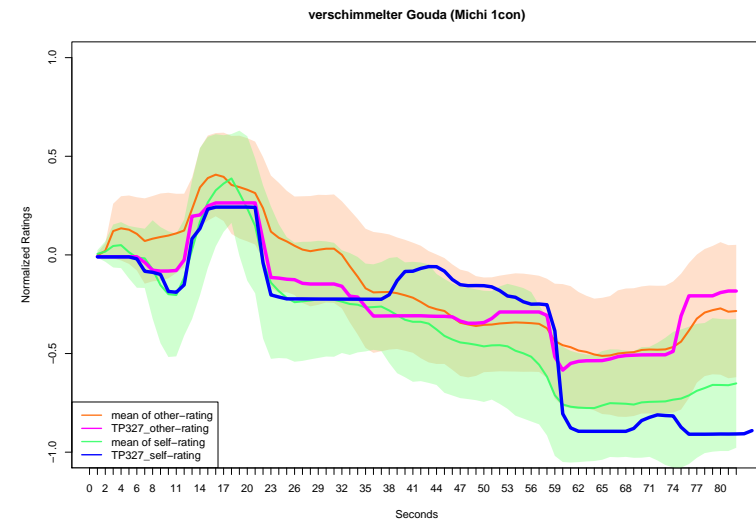
Figure 130: G2 Michi 1con: Ratings by test-person 327.
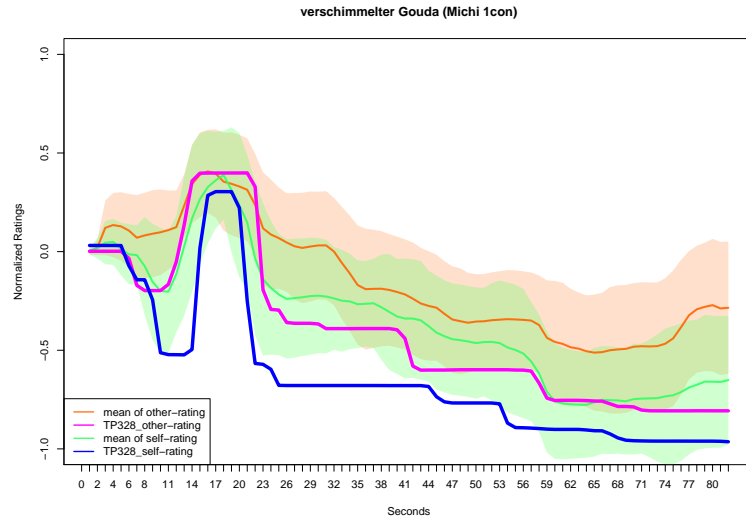
Figure 131: G2 Michi 1con: Ratings by test-person 328.
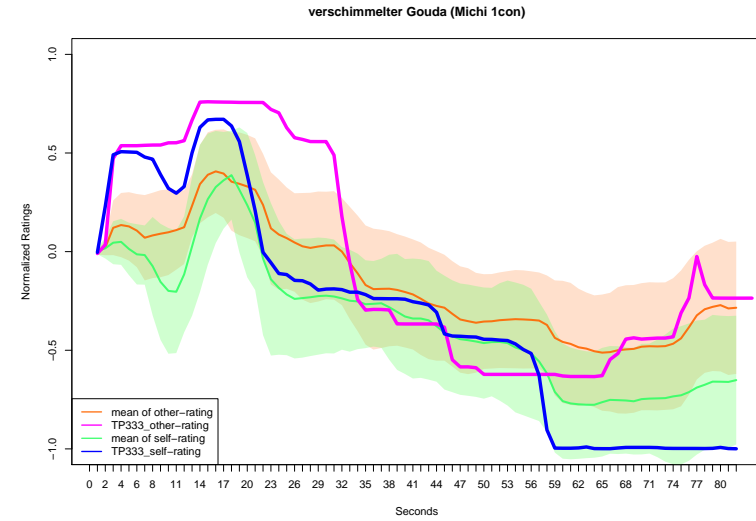
Figure 132: G2 Michi 1con: Ratings by test-person 333.
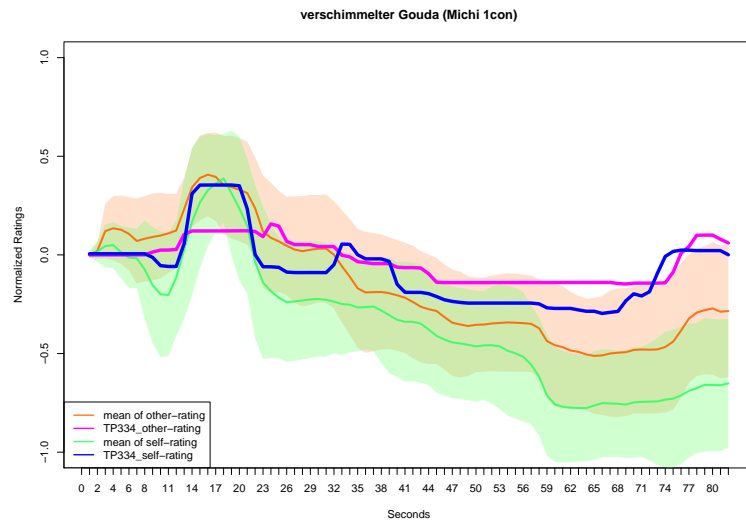
Figure 133: G2 Michi 1con: Ratings by test-person 334.
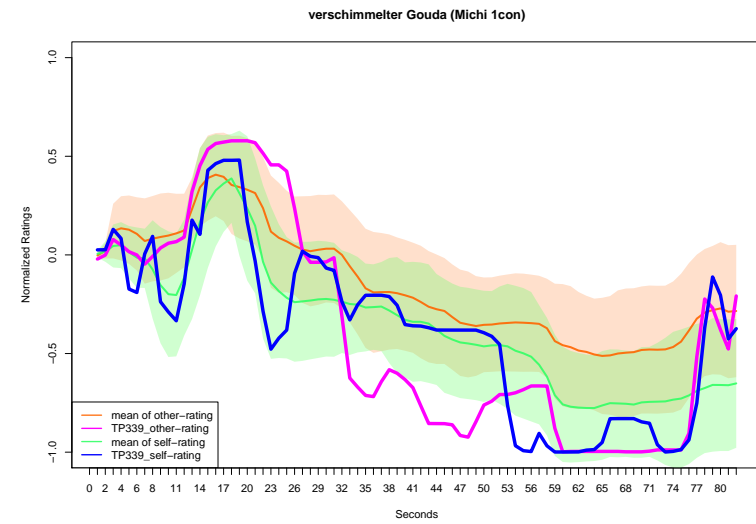
Figure 134: G2 Michi 1con: Ratings by test-person 339.
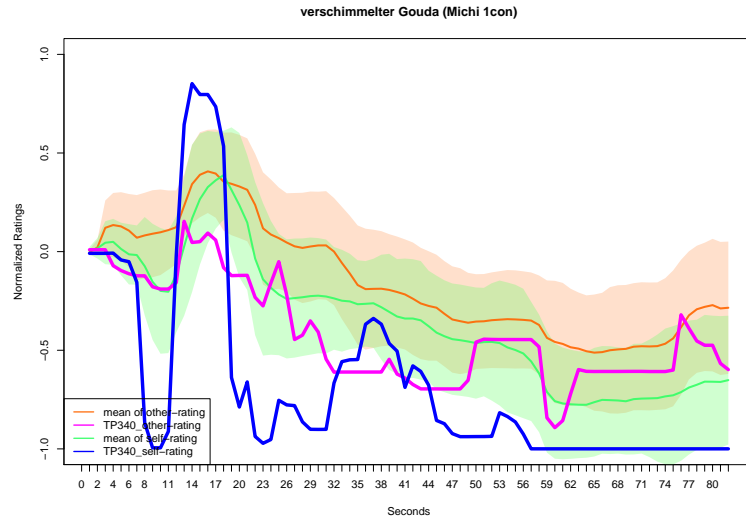
Figure 135: G2 Michi 1con: Ratings by test-person 340.
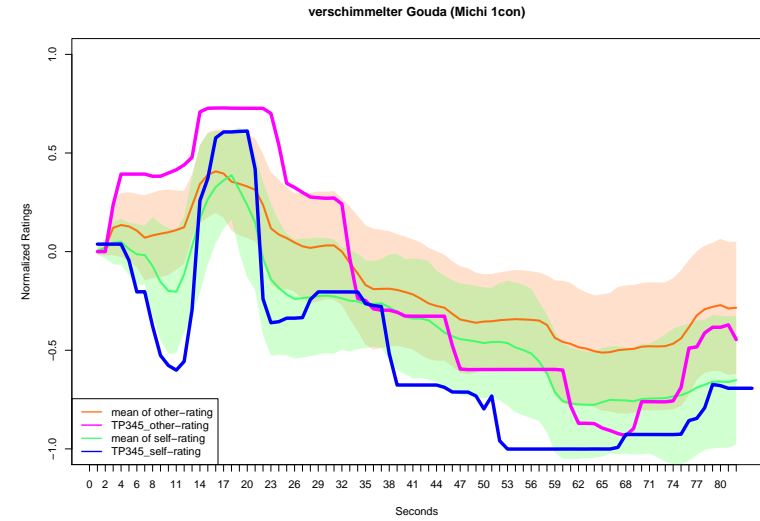
Figure 136: G2 Michi 1con: Ratings by test-person 345.
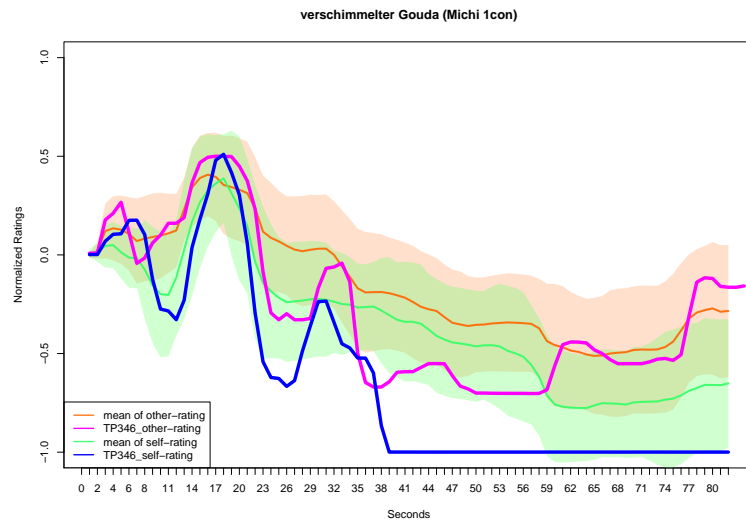
Figure 137: G2 Michi 1con: Ratings by test-person 346.
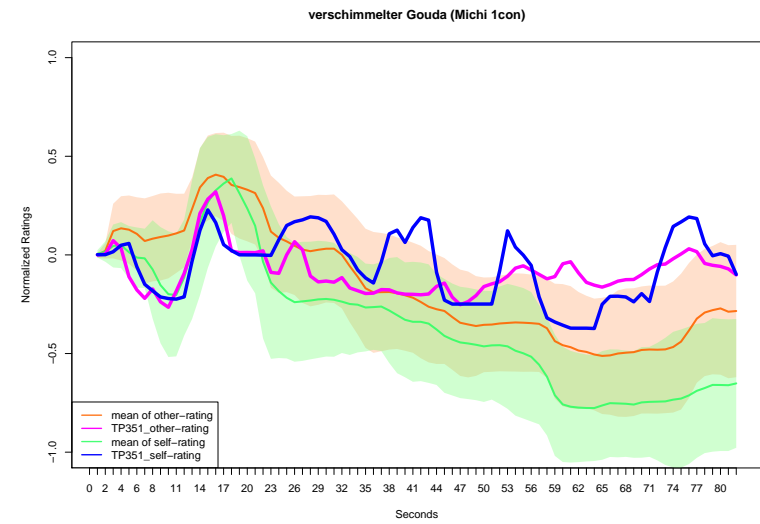
Figure 138: G2 Michi 1con: Ratings by test-person 351.
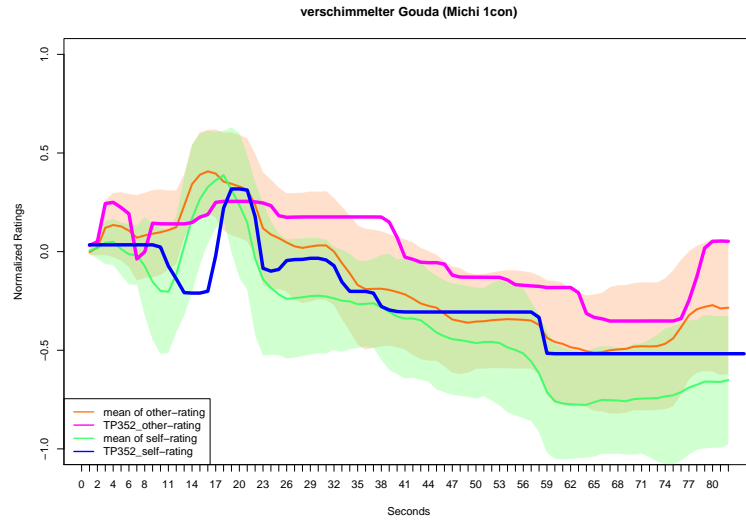
Figure 139: G2 Michi 1con: Ratings by test-person 352.
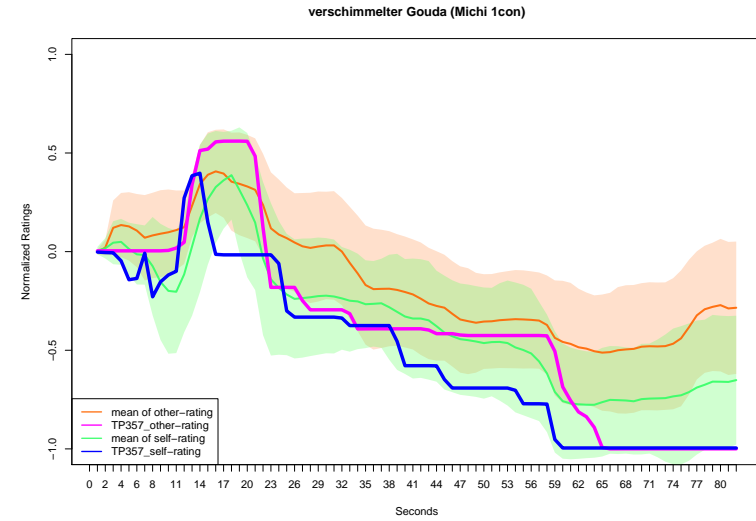
Figure 140: G2 Michi 1con: Ratings by test-person 357.
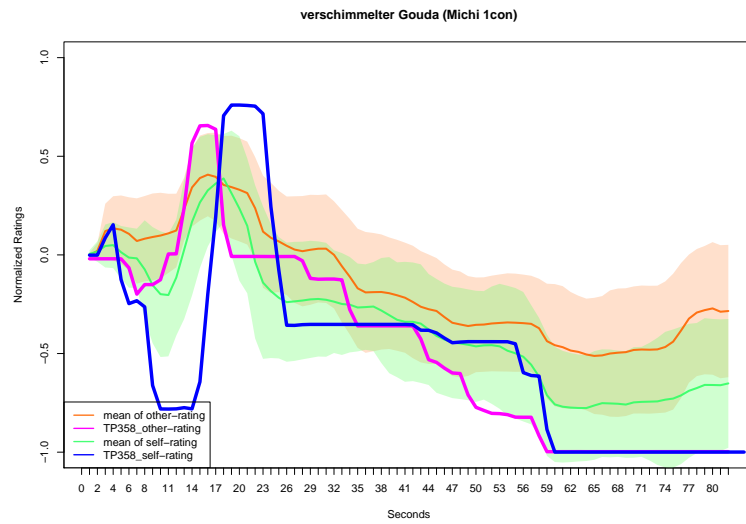
Figure 141: G2 Michi 1con: Ratings by test-person 358.
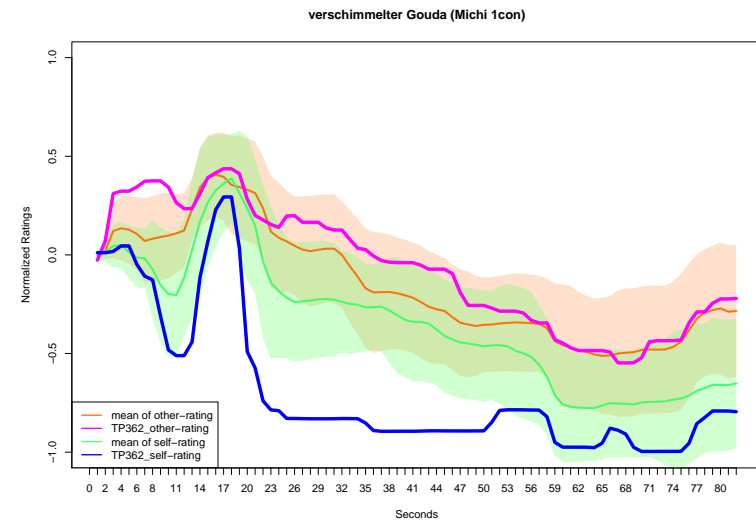
Figure 142: G2 Michi 1con: Ratings by test-person 362.
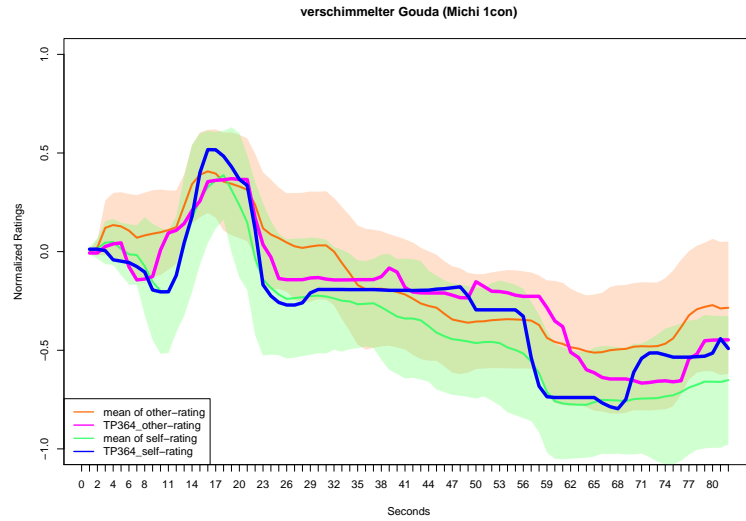
Figure 143: G2 Michi 1con: Ratings by test-person 364.
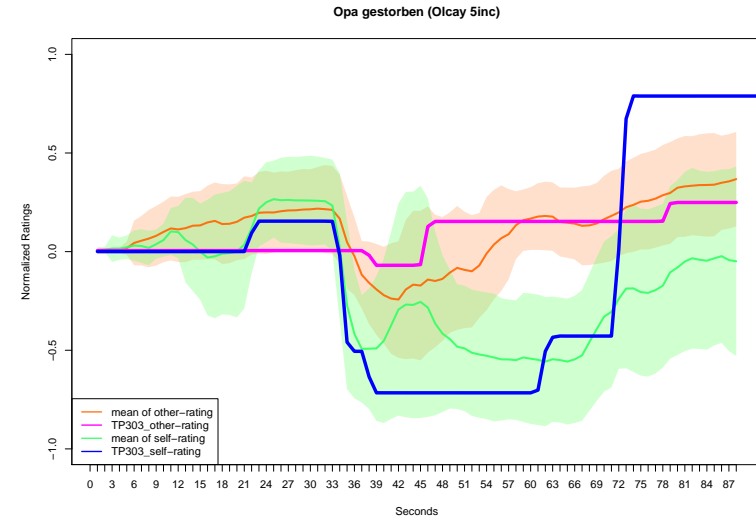
Figure 144: G2 Olcay 5inc: Ratings by test-person 303.
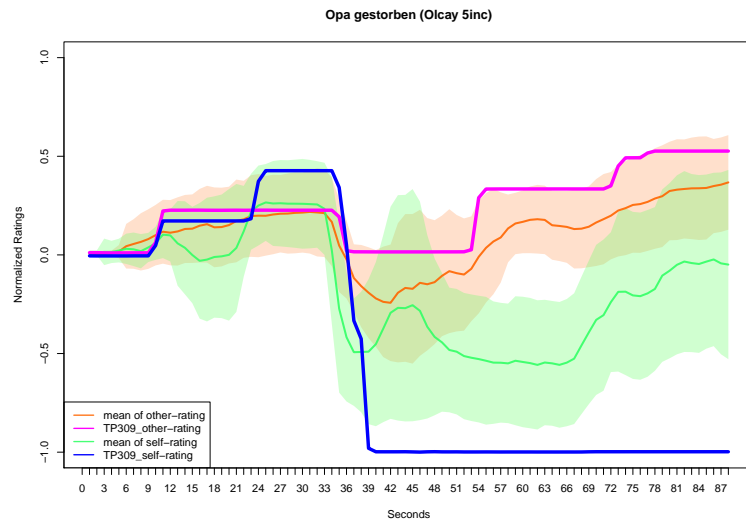
Figure 145: G2 Olcay 5inc: Ratings by test-person 309.
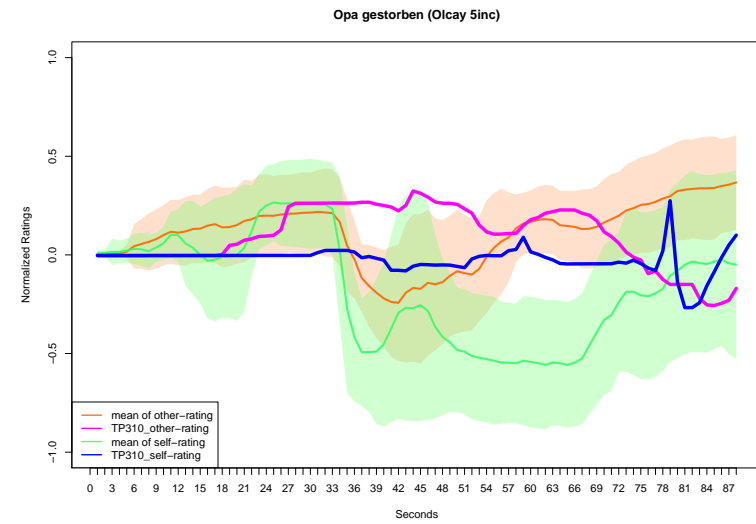
Figure 146: G2 Olcay 5inc: Ratings by test-person 310.
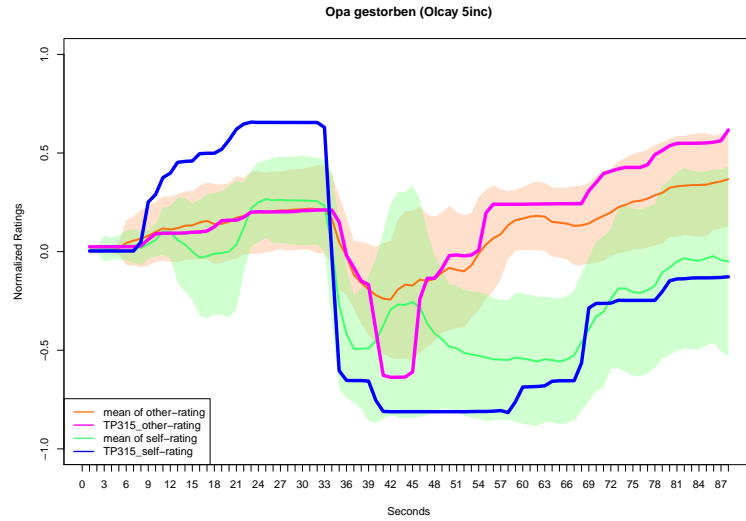
Figure 147: G2 Olcay 5inc: Ratings by test-person 315.
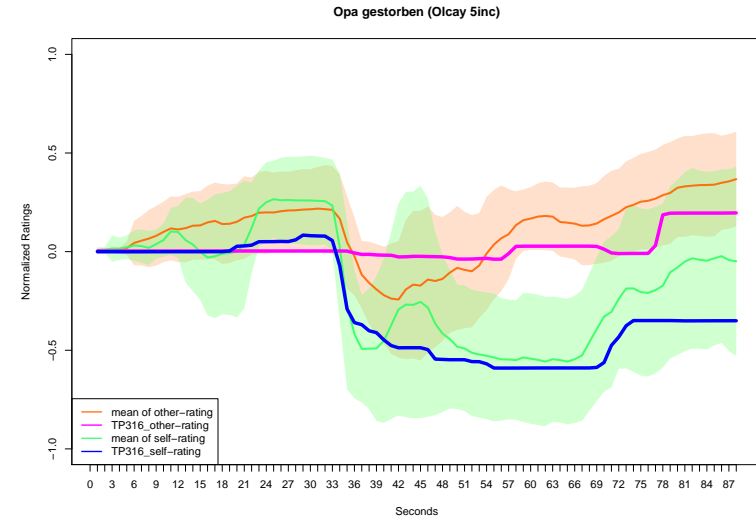
Figure 148: G2 Olcay 5inc: Ratings by test-person 316.
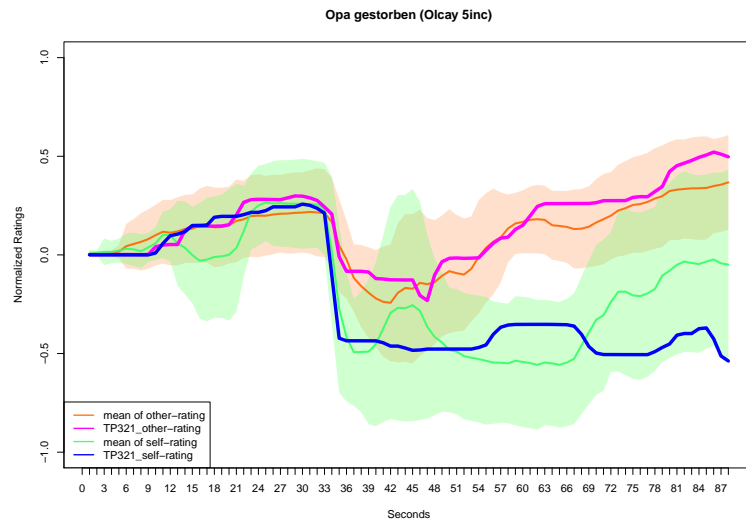
Figure 149: G2 Olcay 5inc: Ratings by test-person 321.
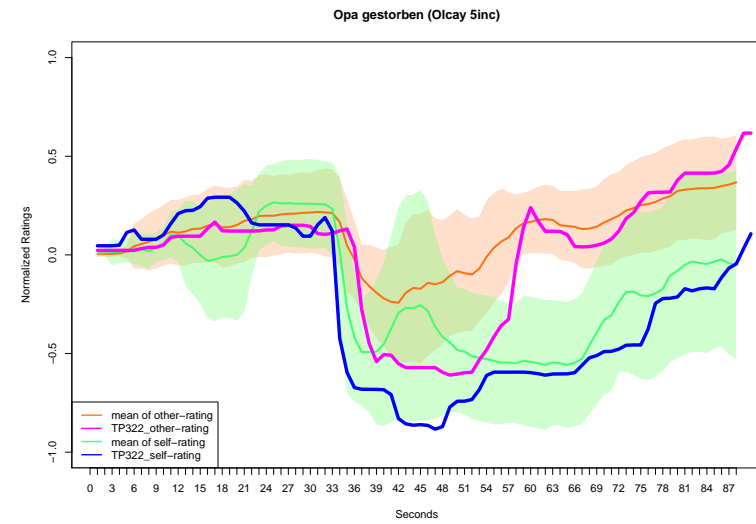
Figure 150: G2 Olcay 5inc: Ratings by test-person 322.
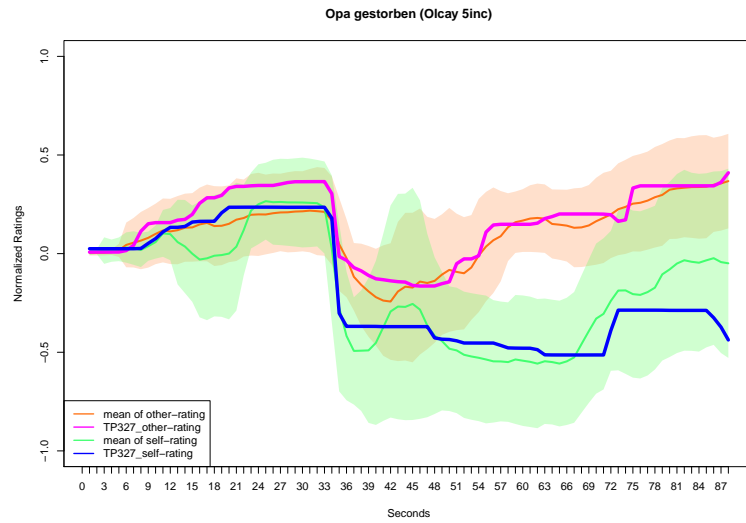
Figure 151: G2 Olcay 5inc: Ratings by test-person 327.
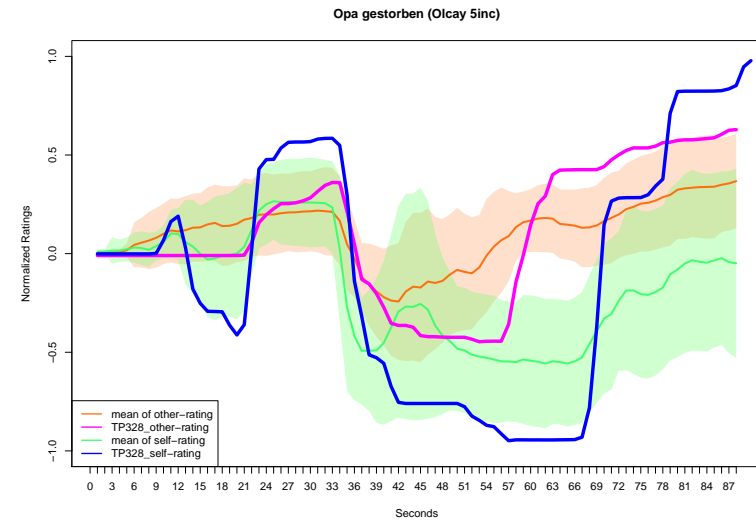
Figure 152: G2 Olcay 5inc: Ratings by test-person 328.
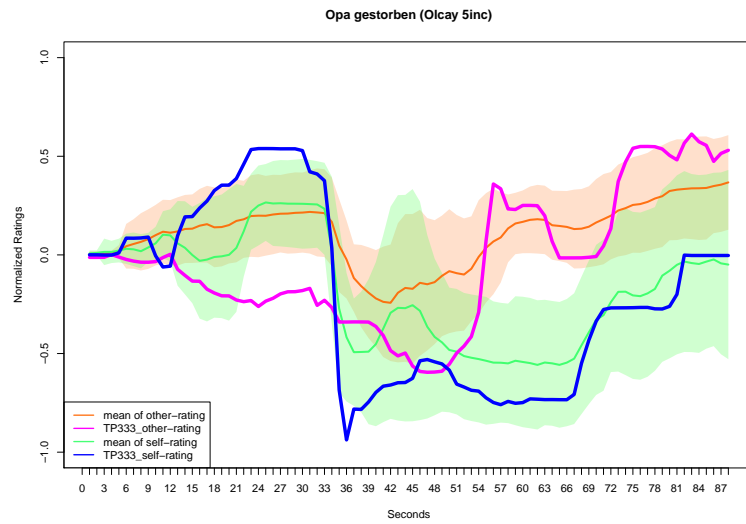
Figure 153: G2 Olcay 5inc: Ratings by test-person 333.
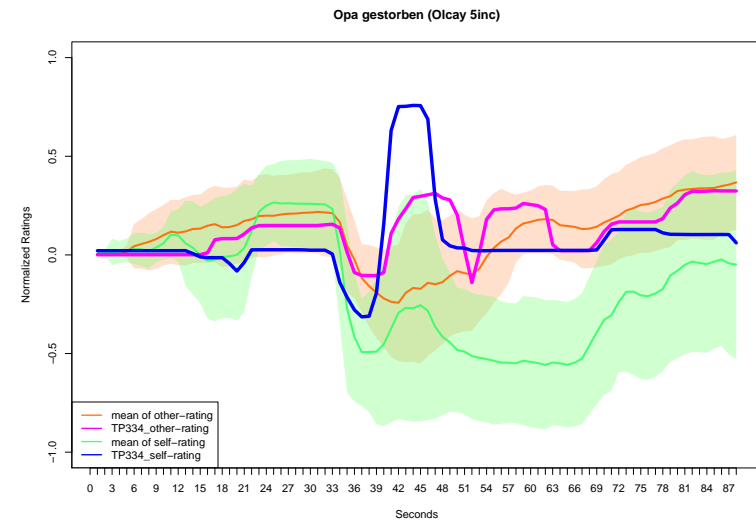
Figure 154: G2 Olcay 5inc: Ratings by test-person 334.
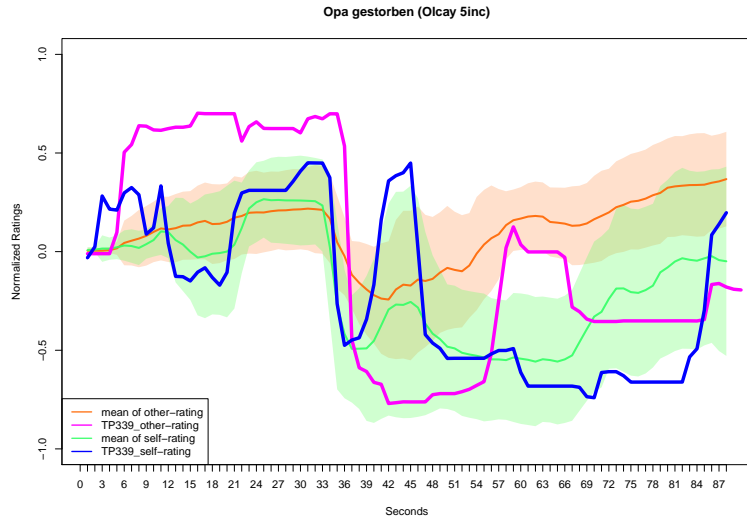
Figure 155: G2 Olcay 5inc: Ratings by test-person 339.
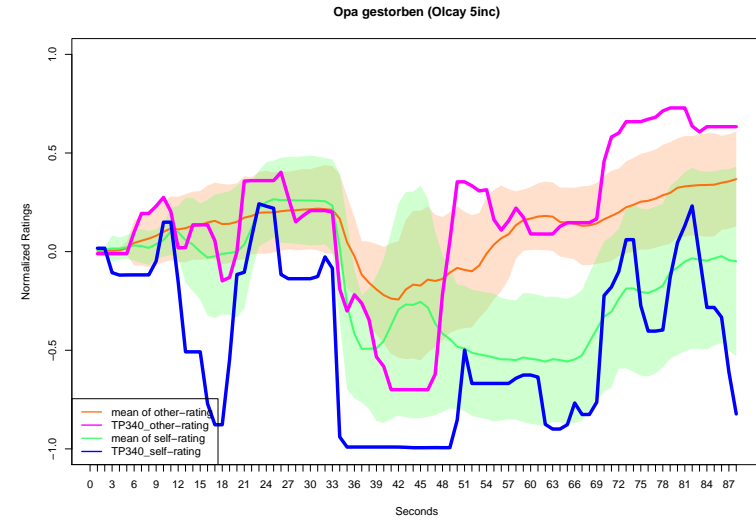
Figure 156: G2 Olcay 5inc: Ratings by test-person 340.
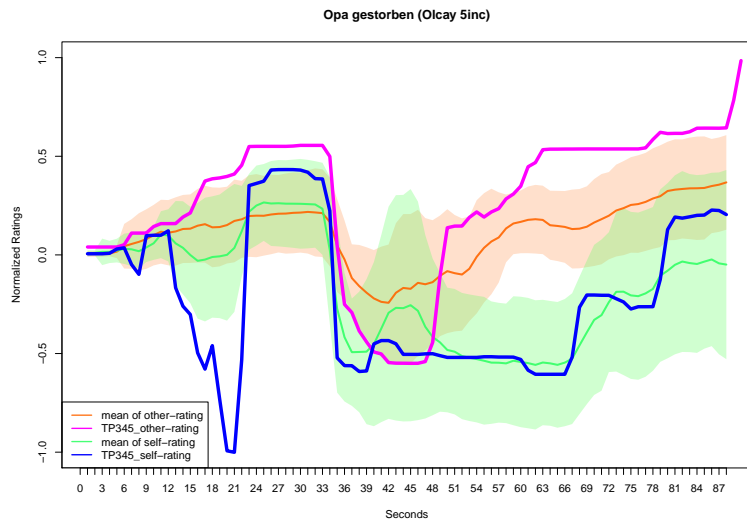
Figure 157: G2 Olcay 5inc: Ratings by test-person 345.
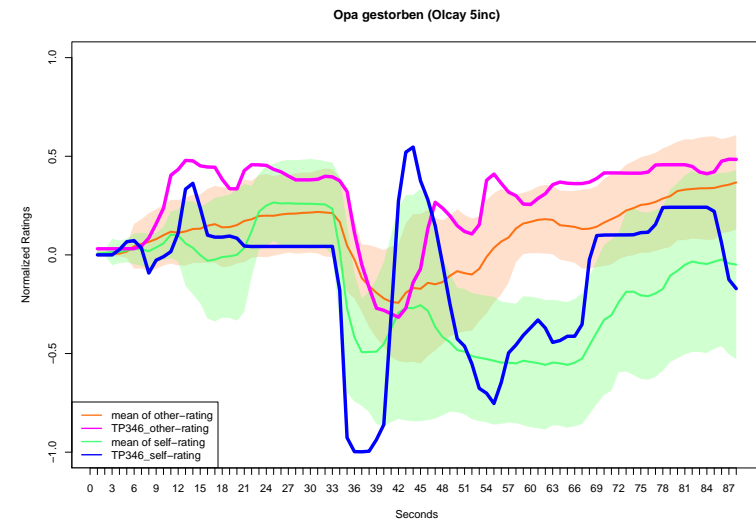
Figure 158: G2 Olcay 5inc: Ratings by test-person 346.
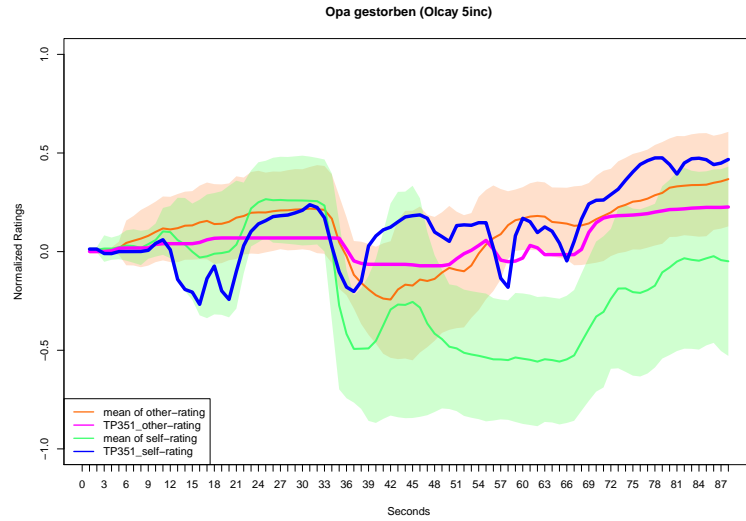
Figure 159: G2 Olcay 5inc: Ratings by test-person 351.
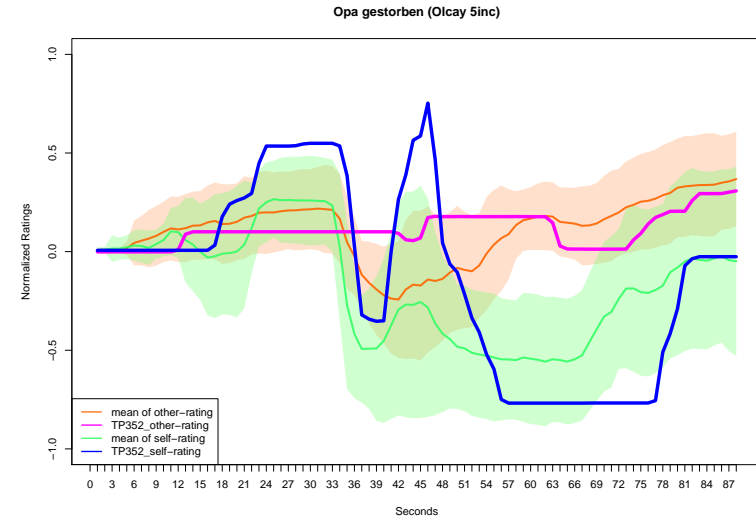
Figure 160: G2 Olcay 5inc: Ratings by test-person 352.
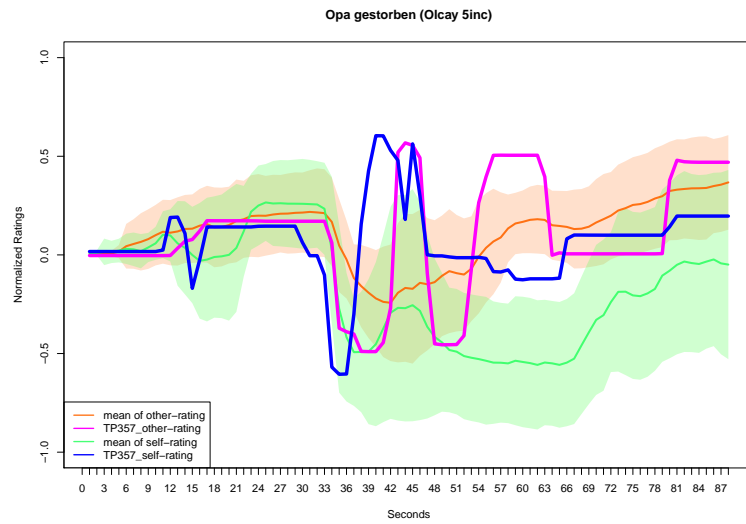
Figure 161: G2 Olcay 5inc: Ratings by test-person 357.
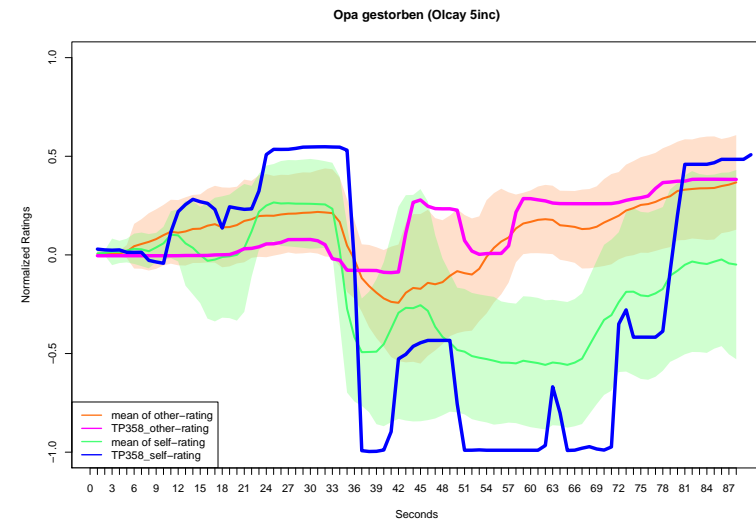
Figure 162: G2 Olcay 5inc: Ratings by test-person 358.
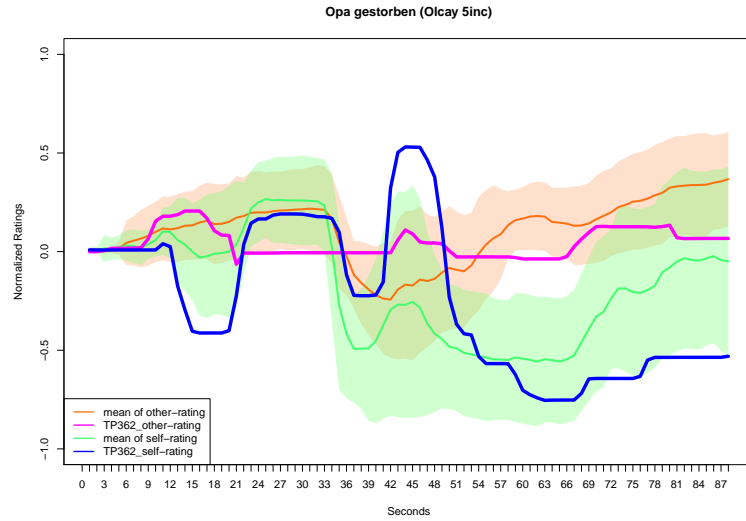
Figure 163: G2 Olcay 5inc: Ratings by test-person 362.
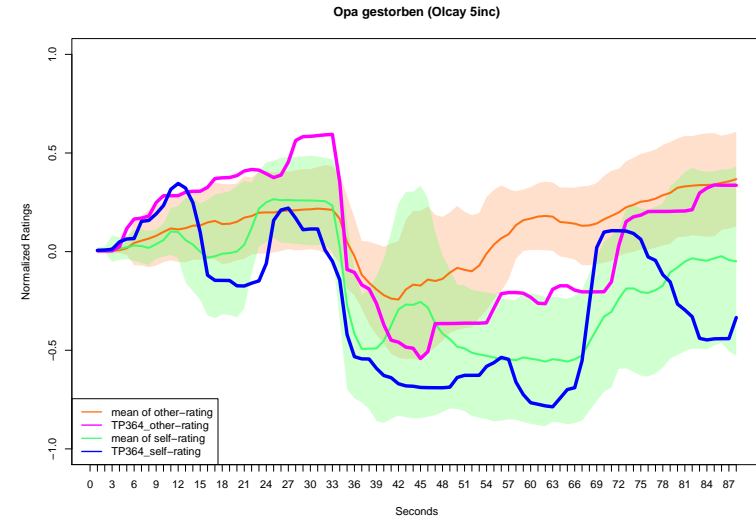
Figure 164: G2 Olcay 5inc: Ratings by test-person 364.
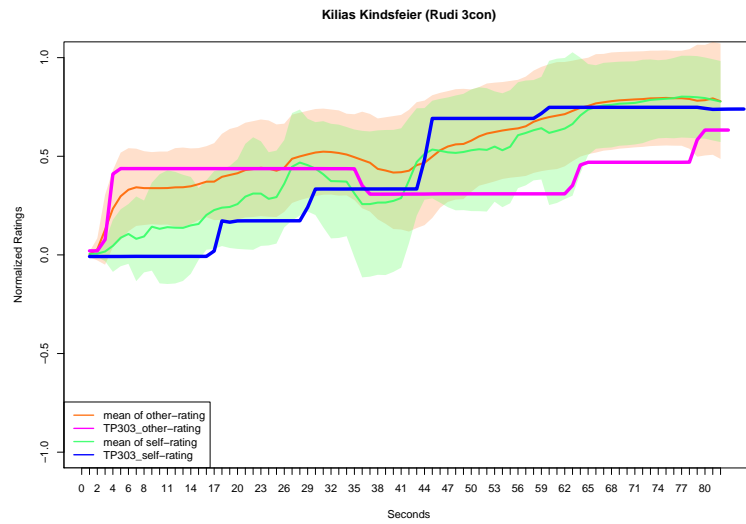
Figure 165: G2 Rudi 3con: Ratings by test-person 303.
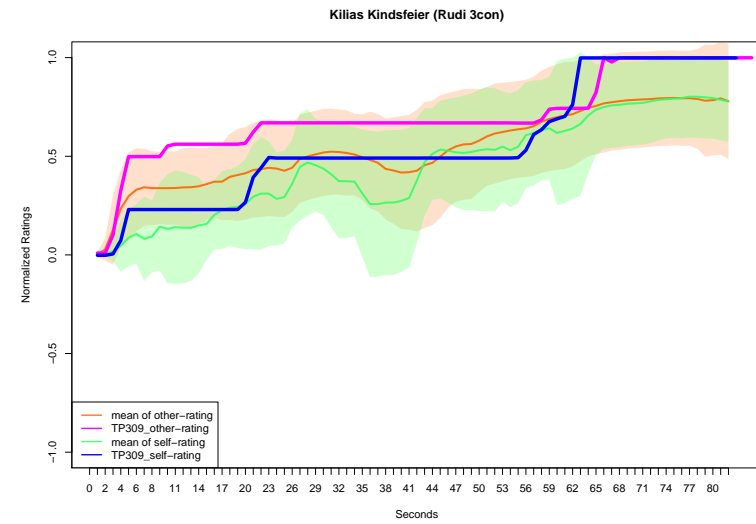
Figure 166: G2 Rudi 3con: Ratings by test-person 309.
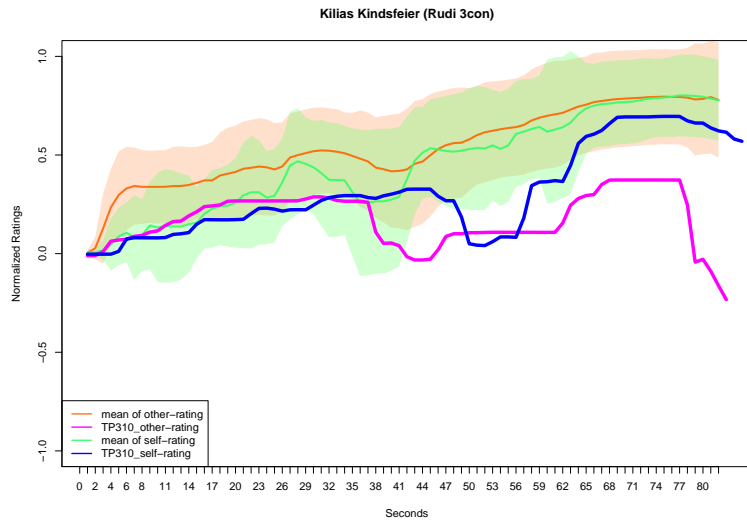
Figure 167: G2 Rudi 3con: Ratings by test-person 310.
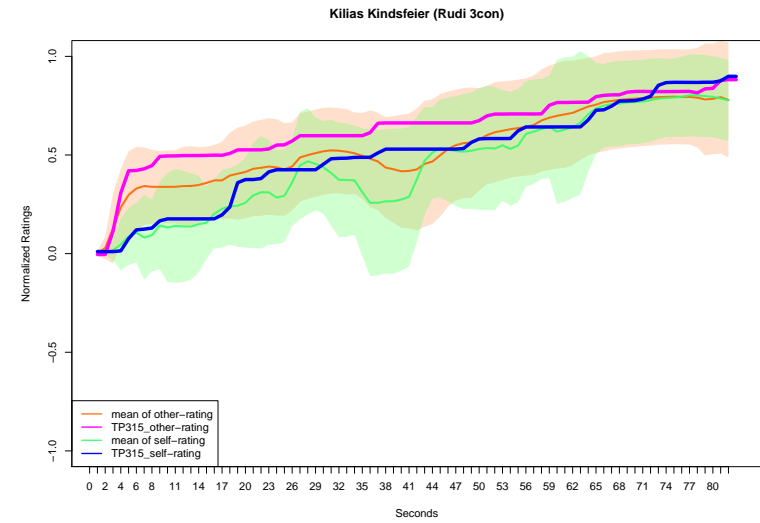
Figure 168: G2 Rudi 3con: Ratings by test-person 315.
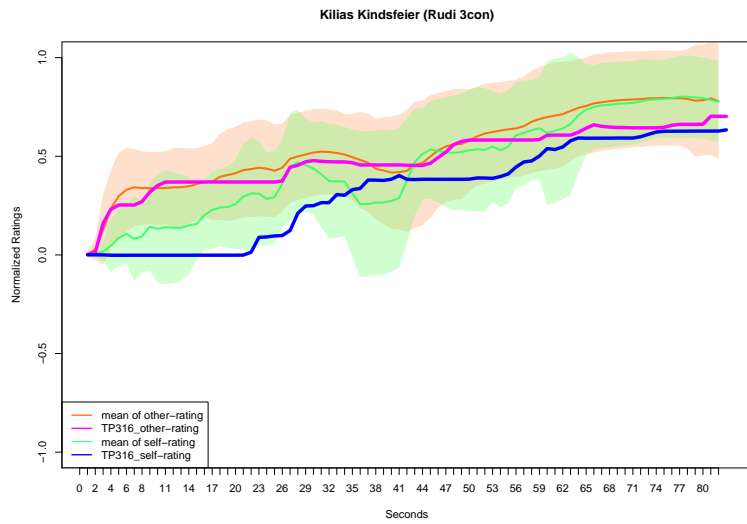
Figure 169: G2 Rudi 3con: Ratings by test-person 316.
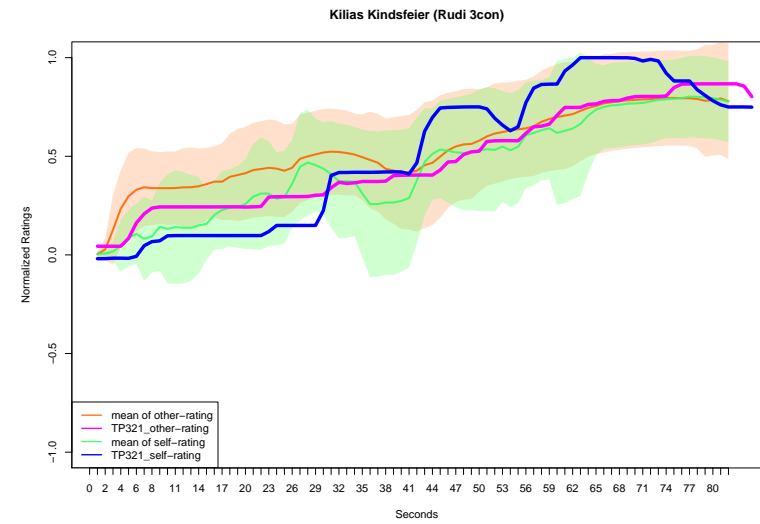
Figure 170: G2 Rudi 3con: Ratings by test-person 321.
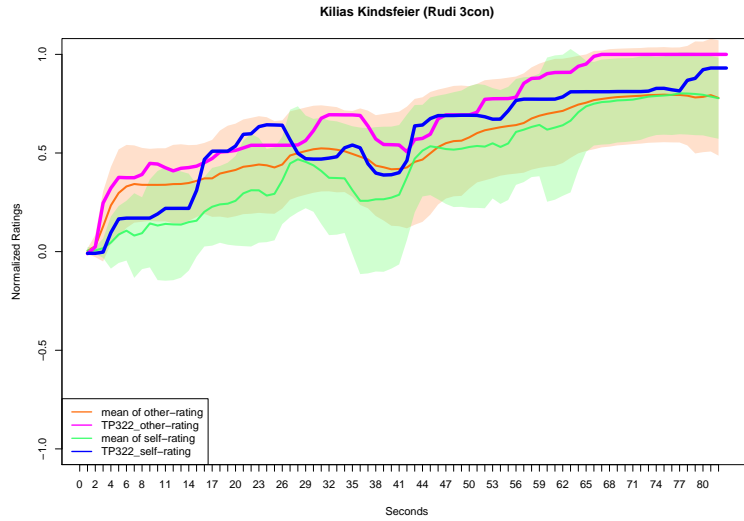
Figure 171: G2 Rudi 3con: Ratings by test-person 322.
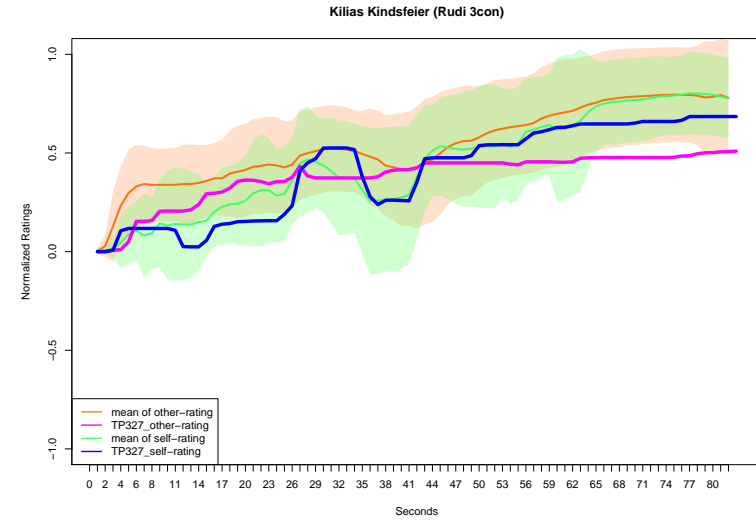
Figure 172: G2 Rudi 3con: Ratings by test-person 327.
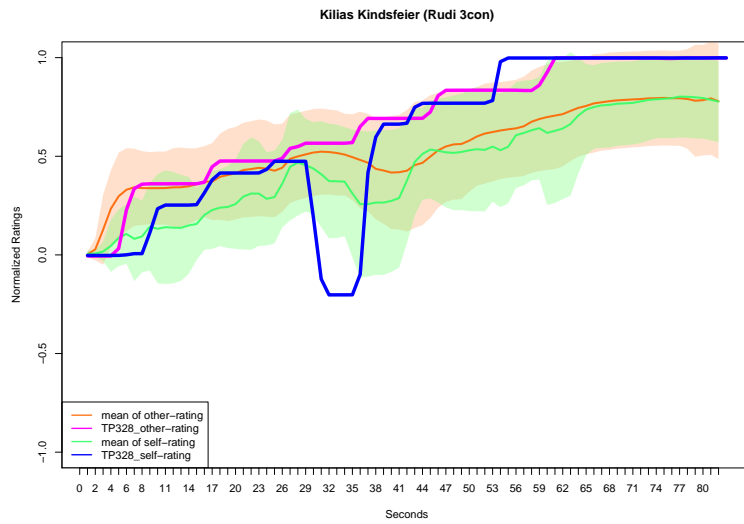
Figure 173: G2 Rudi 3con: Ratings by test-person 328.
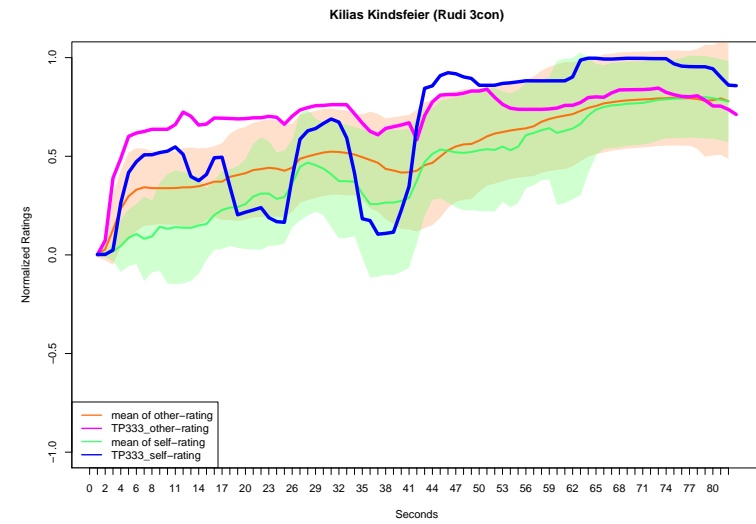
Figure 174: G2 Rudi 3con: Ratings by test-person 333.
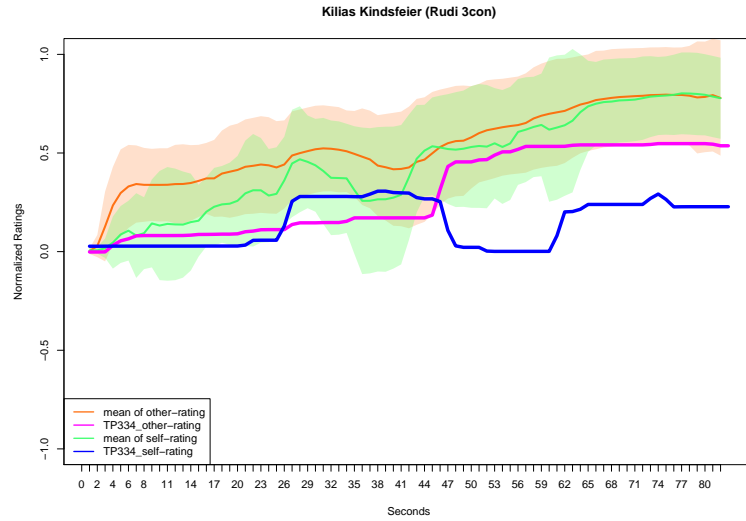
Figure 175: G2 Rudi 3con: Ratings by test-person 334.
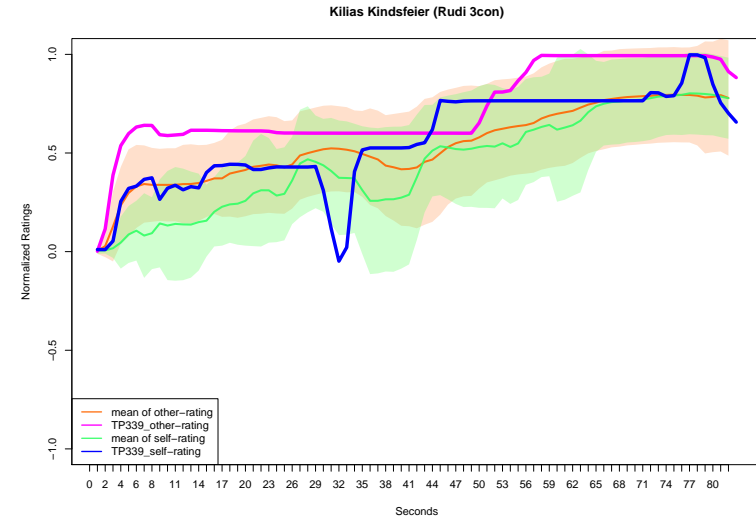
Figure 176: G2 Rudi 3con: Ratings by test-person 339.
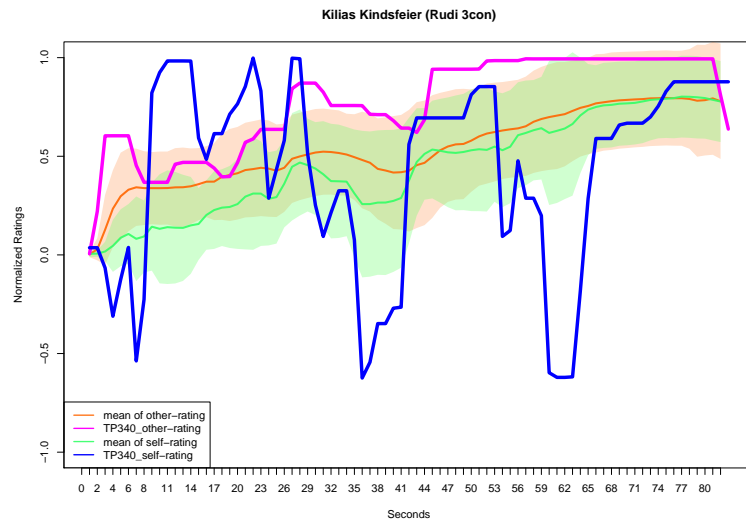
Figure 177: G2 Rudi 3con: Ratings by test-person 340.
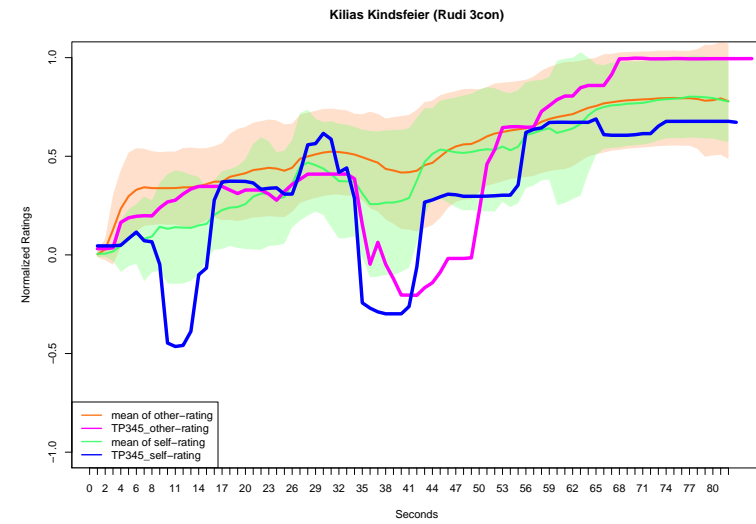
Figure 178: G2 Rudi 3con: Ratings by test-person 345.
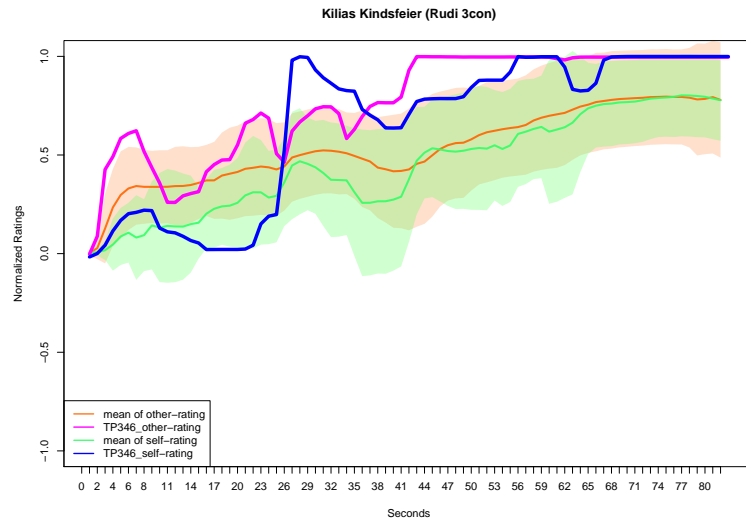
Figure 179: G2 Rudi 3con: Ratings by test-person 346.
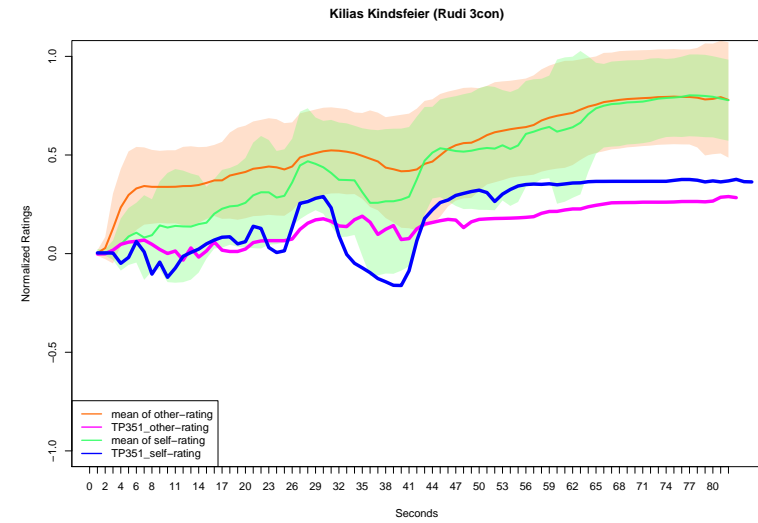
Figure 180: G2 Rudi 3con: Ratings by test-person 351.
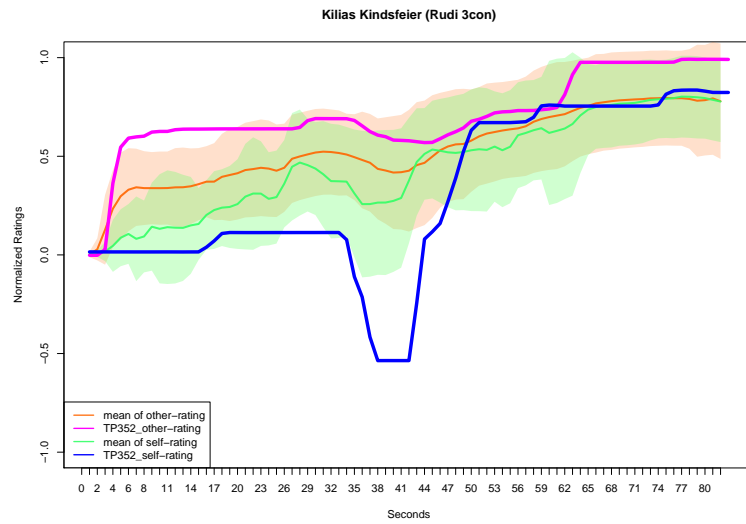
Figure 181: G2 Rudi 3con: Ratings by test-person 352.
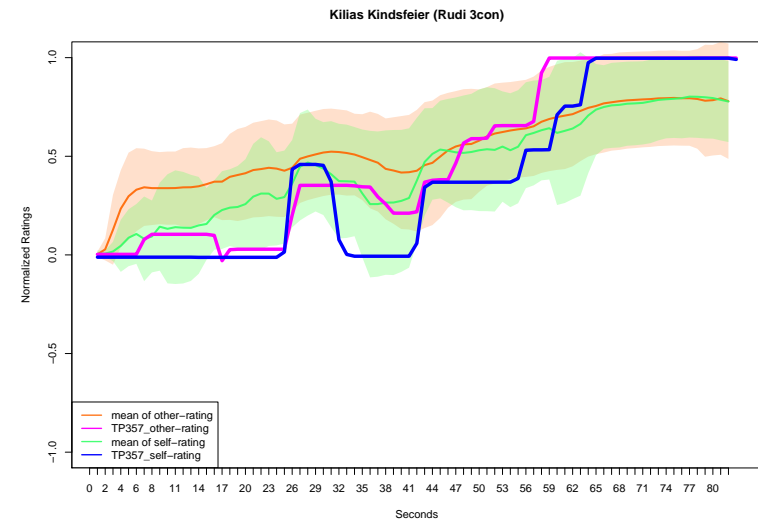
Figure 182: G2 Rudi 3con: Ratings by test-person 357.
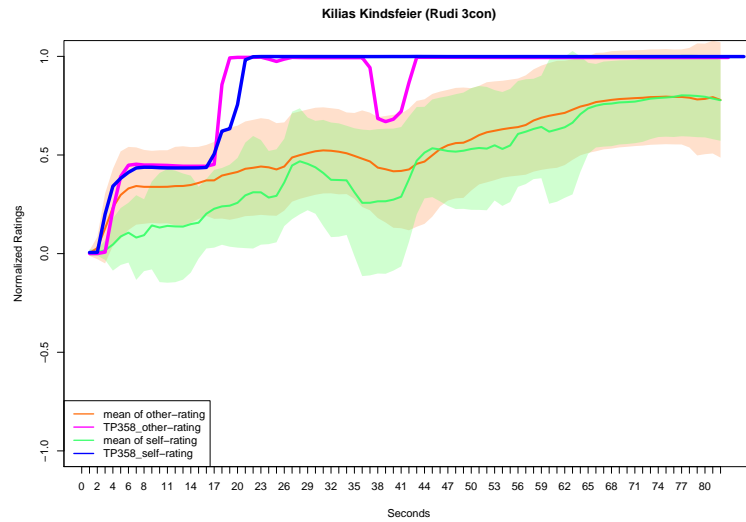
Figure 183: G2 Rudi 3con: Ratings by test-person 358.
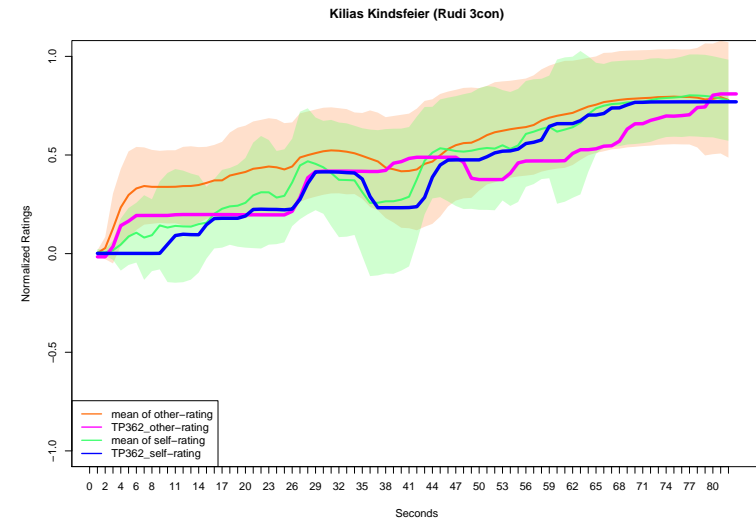
Figure 184: G2 Rudi 3con: Ratings by test-person 362.
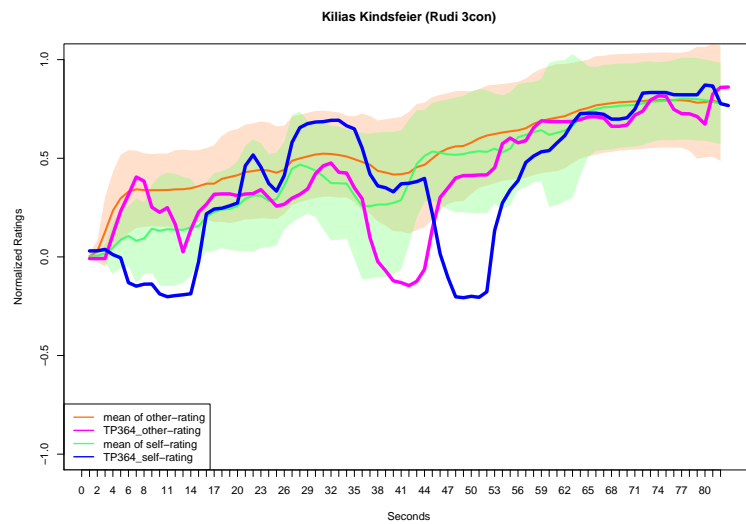
Figure 185: G2 Rudi 3con: Ratings by test-person 364.
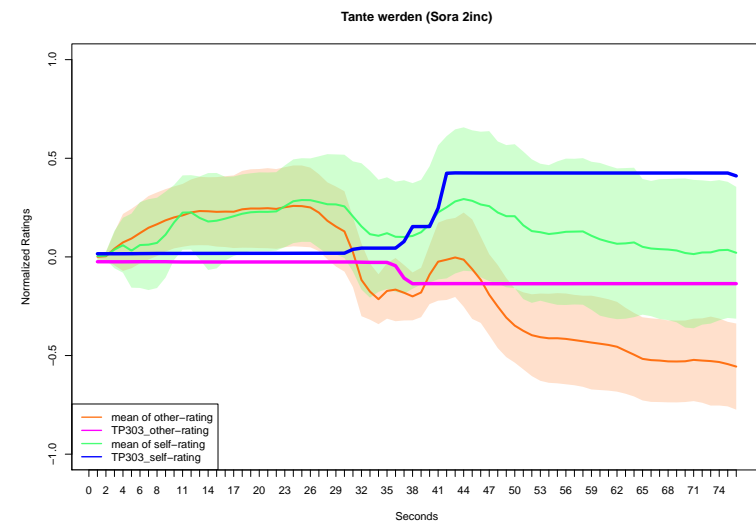
Figure 186: G2 Sora 2inc: Ratings by test-person 303.
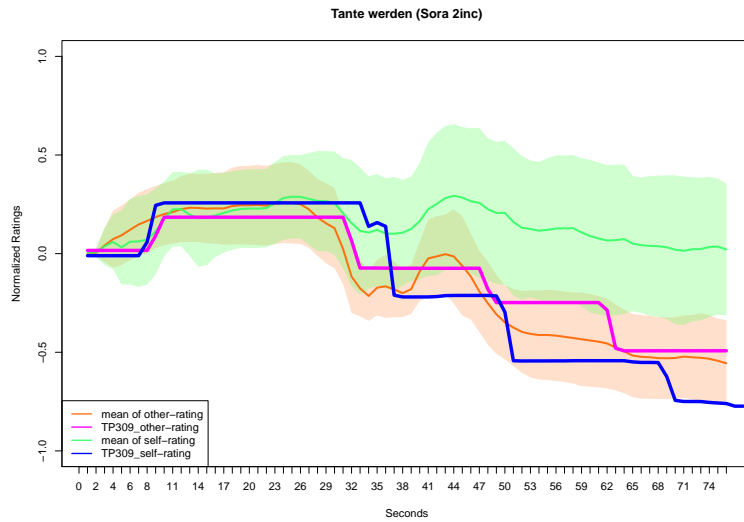
Figure 187: G2 Sora 2inc: Ratings by test-person 309.
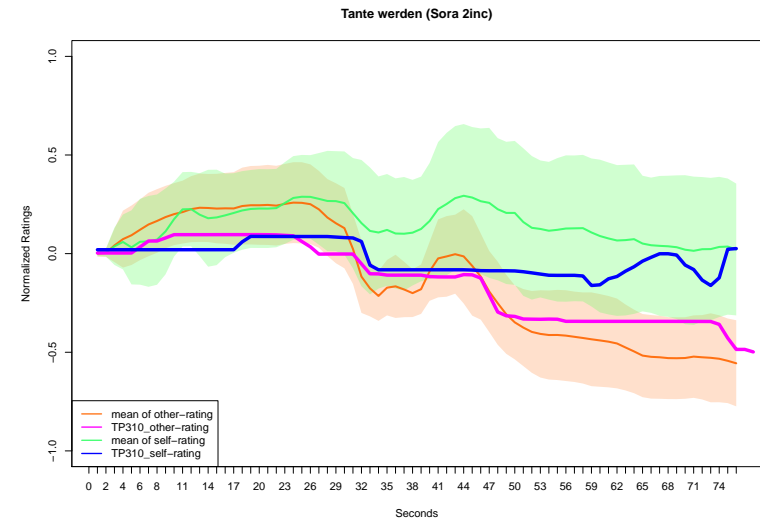
Figure 188: G2 Sora 2inc: Ratings by test-person 310.
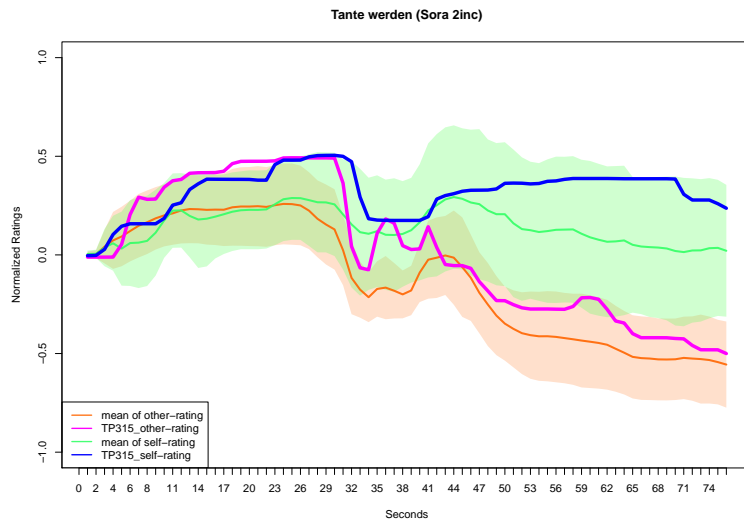
Figure 189: G2 Sora 2inc: Ratings by test-person 315.
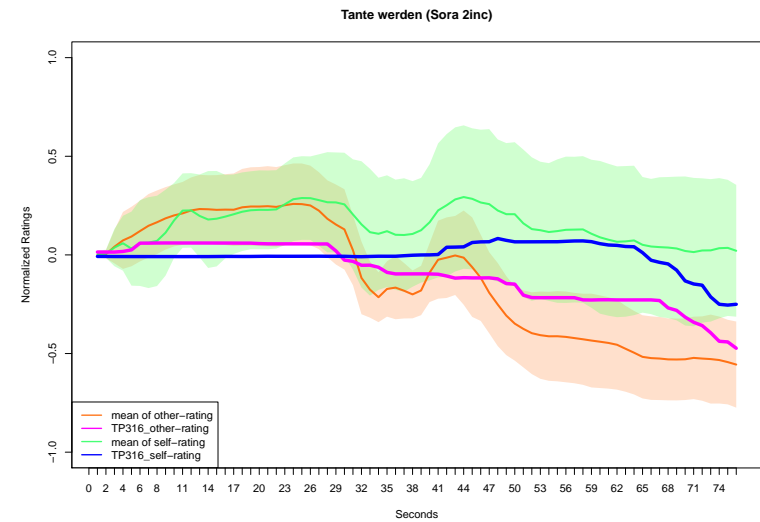
Figure 190: G2 Sora 2inc: Ratings by test-person 316.
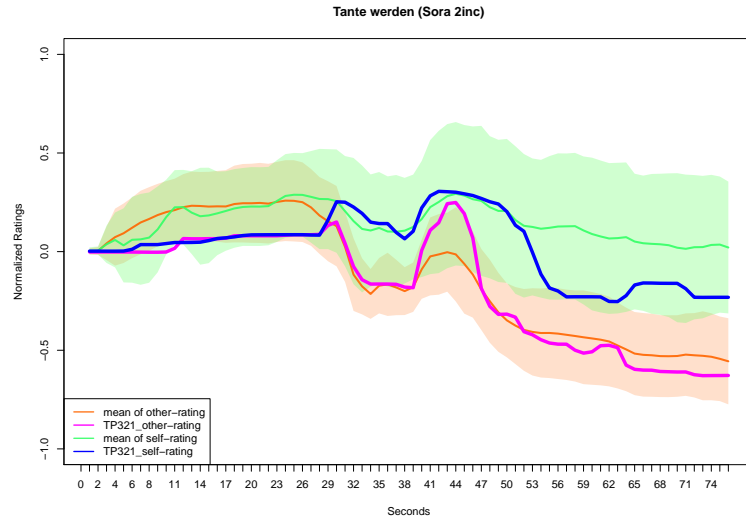
Figure 191: G2 Sora 2inc: Ratings by test-person 321.
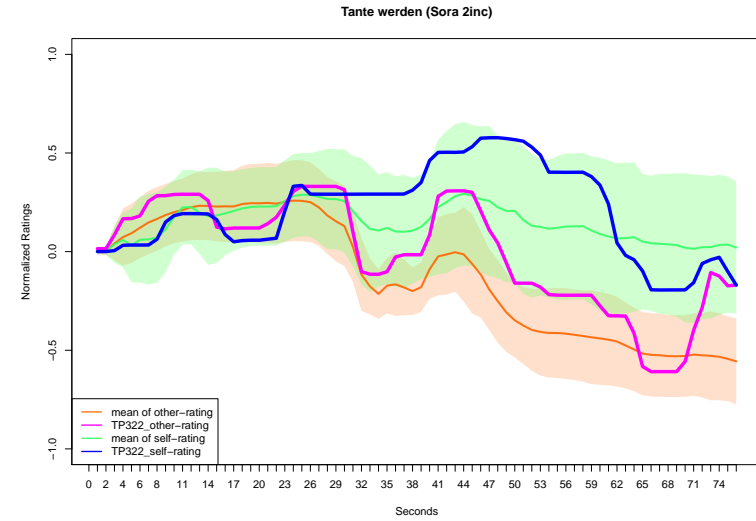
Figure 192: G2 Sora 2inc: Ratings by test-person 322.
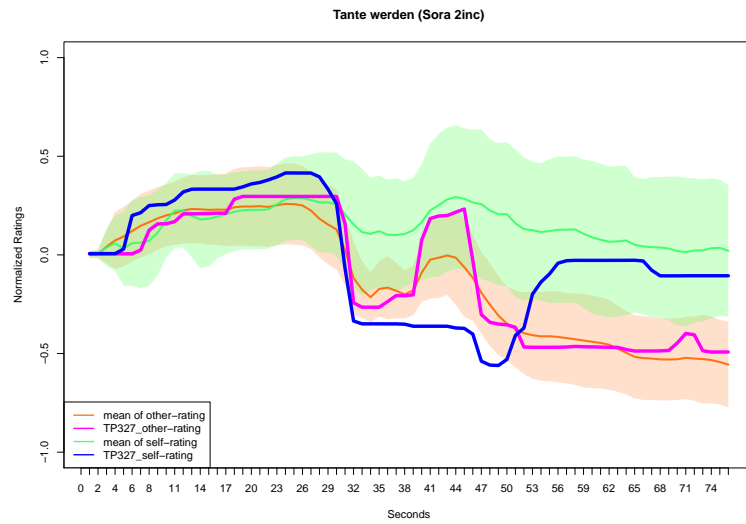
Figure 193: G2 Sora 2inc: Ratings by test-person 327.
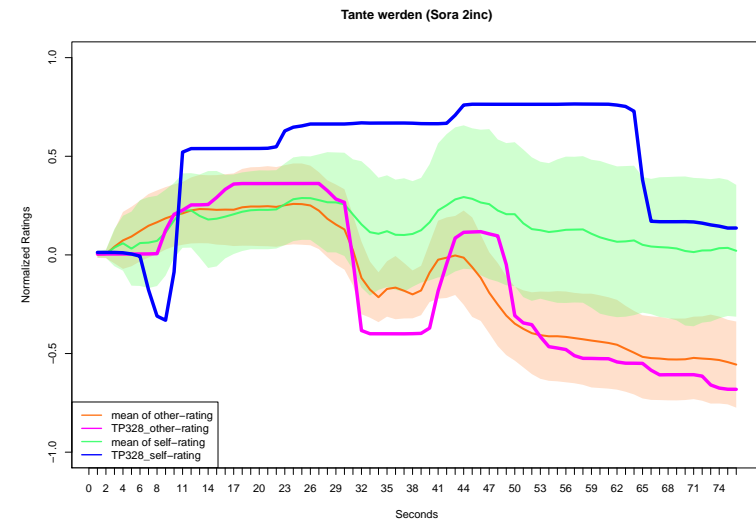
Figure 194: G2 Sora 2inc: Ratings by test-person 328.
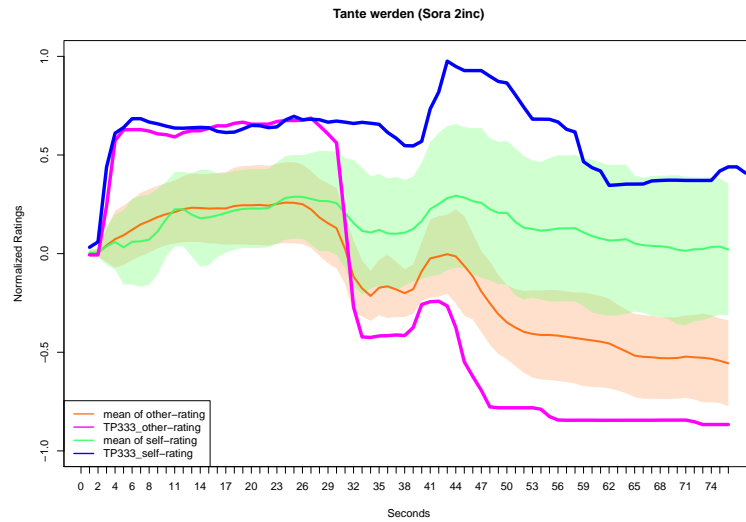
Figure 195: G2 Sora 2inc: Ratings by test-person 333.
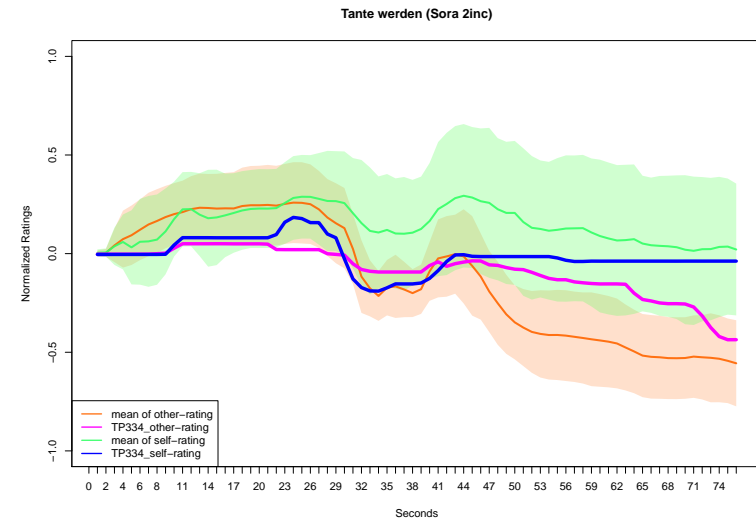
Figure 196: G2 Sora 2inc: Ratings by test-person 334.
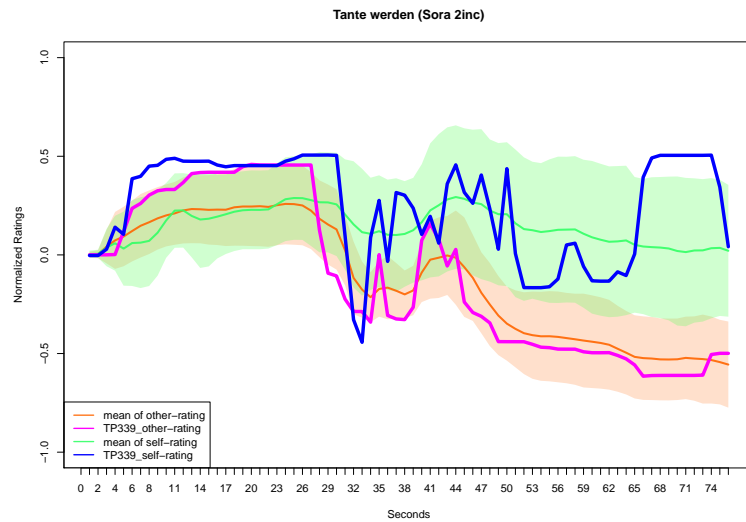
Figure 197: G2 Sora 2inc: Ratings by test-person 339.
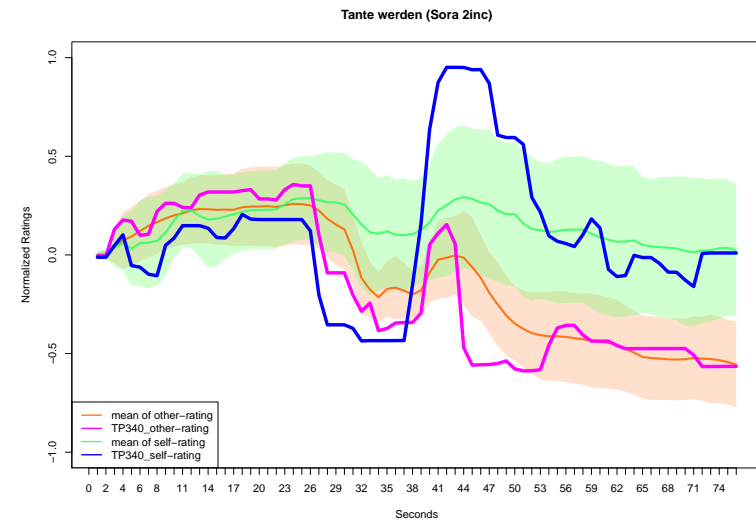
Figure 198: G2 Sora 2inc: Ratings by test-person 340.
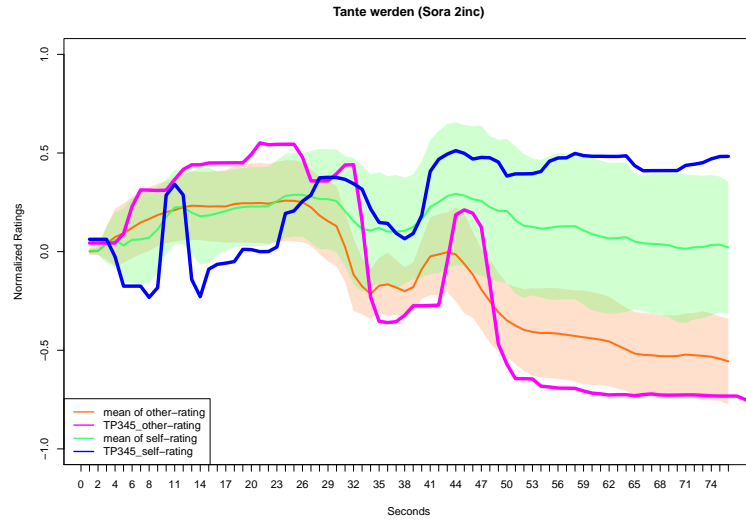
Figure 199: G2 Sora 2inc: Ratings by test-person 345.
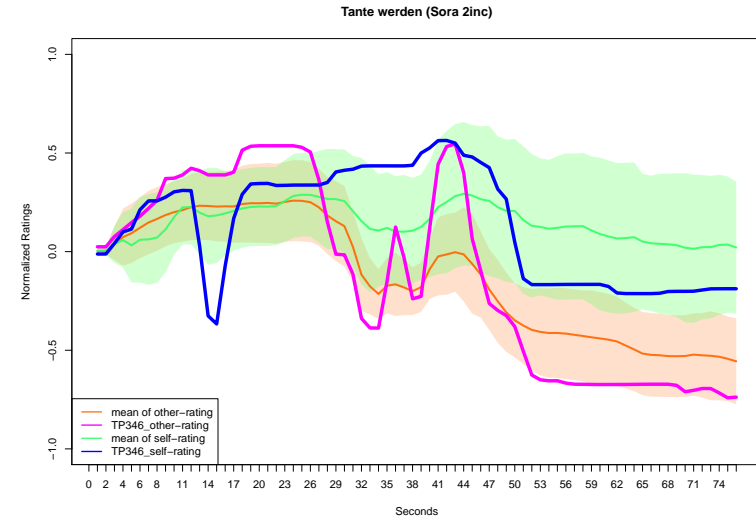
Figure 200: G2 Sora 2inc: Ratings by test-person 346.
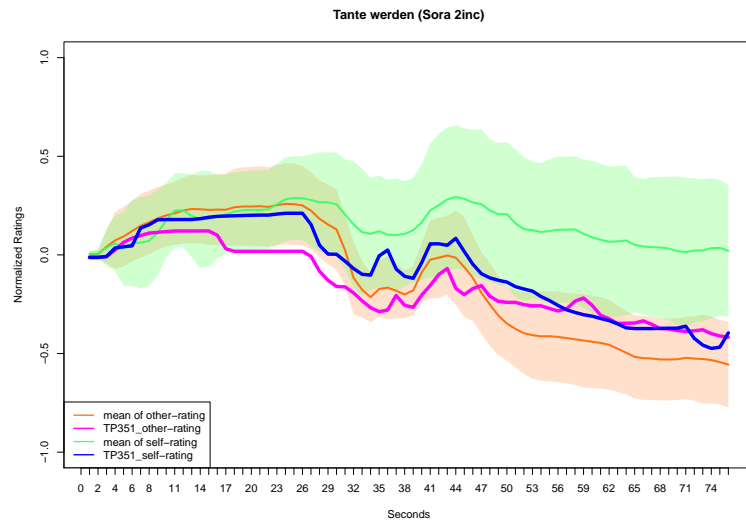
Figure 201: G2 Sora 2inc: Ratings by test-person 351.
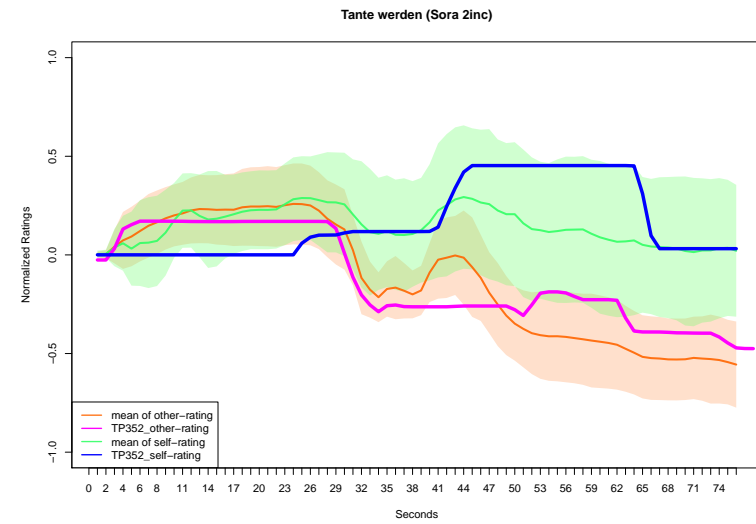
Figure 202: G2 Sora 2inc: Ratings by test-person 352.
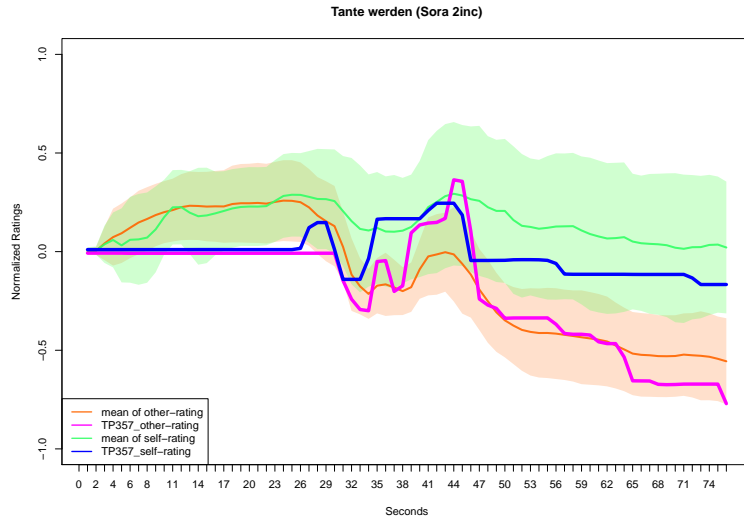
Figure 203: G2 Sora 2inc: Ratings by test-person 357.
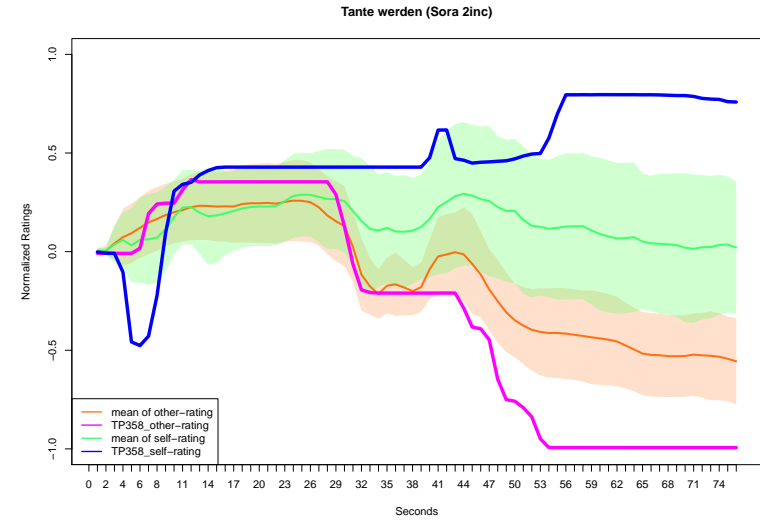
Figure 204: G2 Sora 2inc: Ratings by test-person 358.
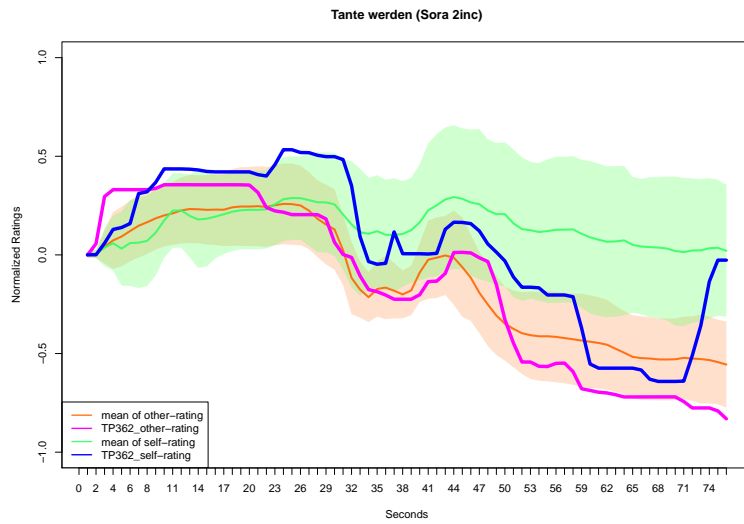
Figure 205: G2 Sora 2inc: Ratings by test-person 362.
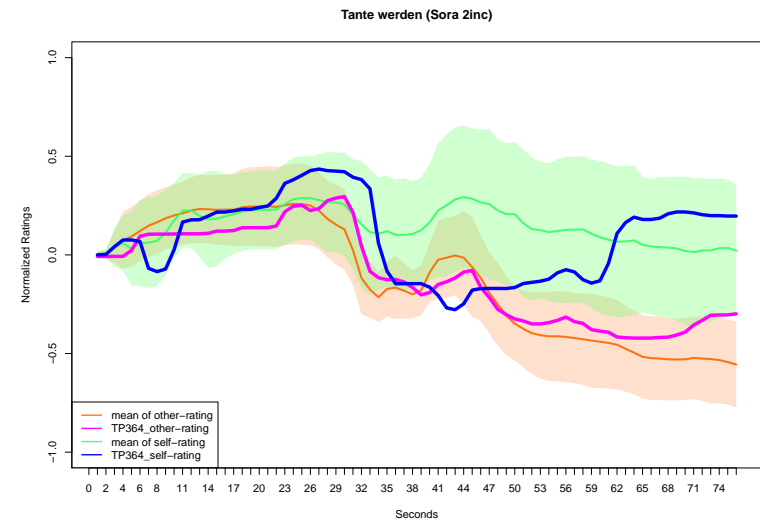
Figure 206: G2 Sora 2inc: Ratings by test-person 364.
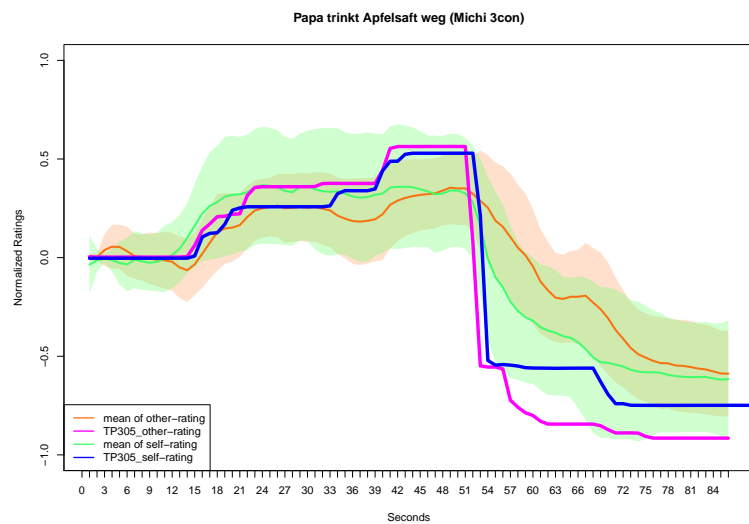
Figure 207: G3 Michi 3con: Ratings by test-person 305.
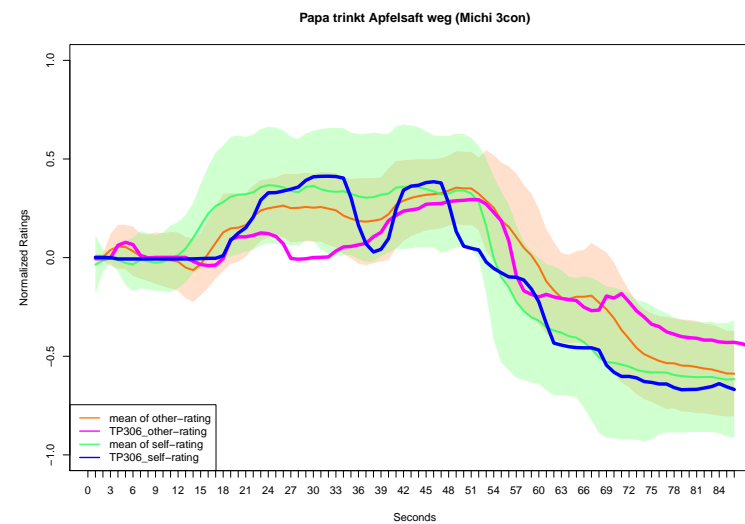
Figure 208: G3 Michi 3con: Ratings by test-person 306.
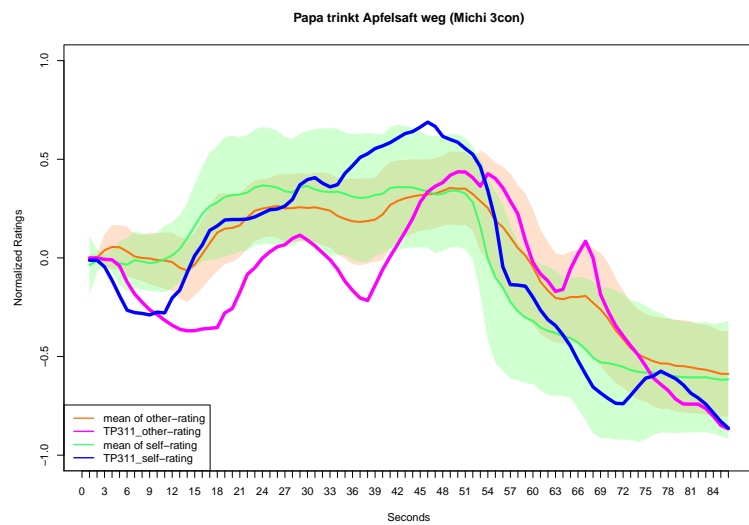
Figure 209: G3 Michi 3con: Ratings by test-person 311.
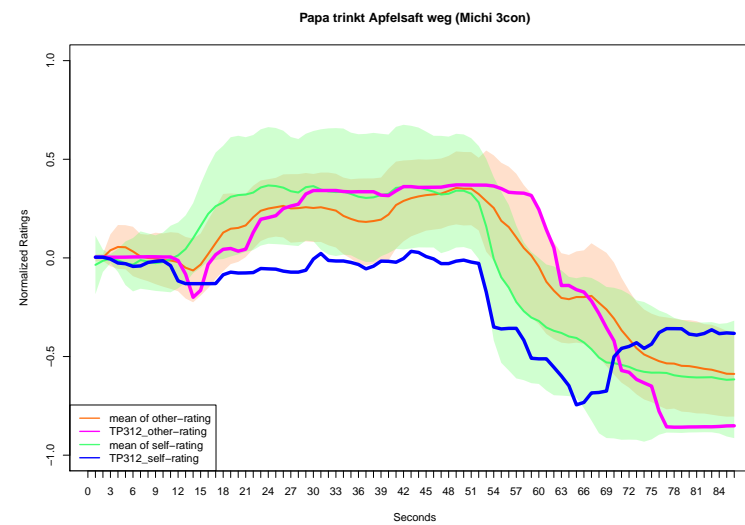
Figure 210: G3 Michi 3con: Ratings by test-person 312.
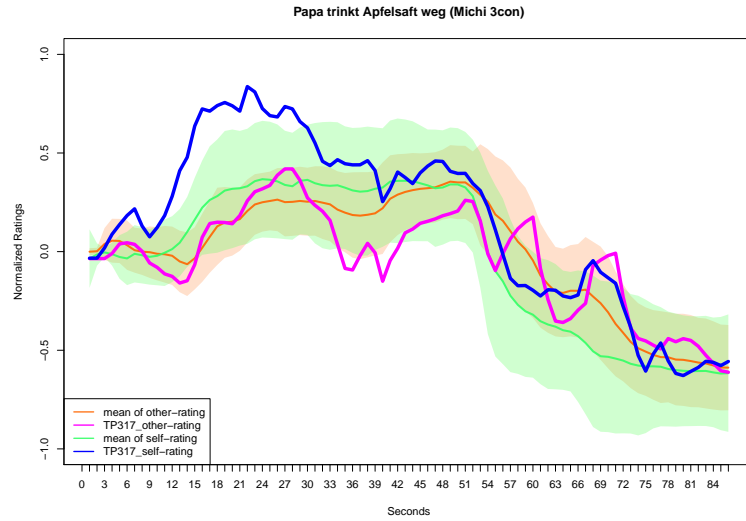
Figure 211: G3 Michi 3con: Ratings by test-person 317.
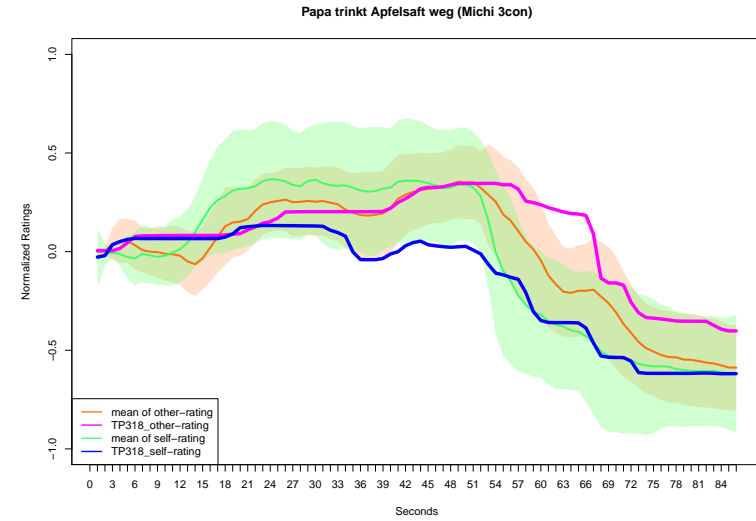
Figure 212: G3 Michi 3con: Ratings by test-person 318.
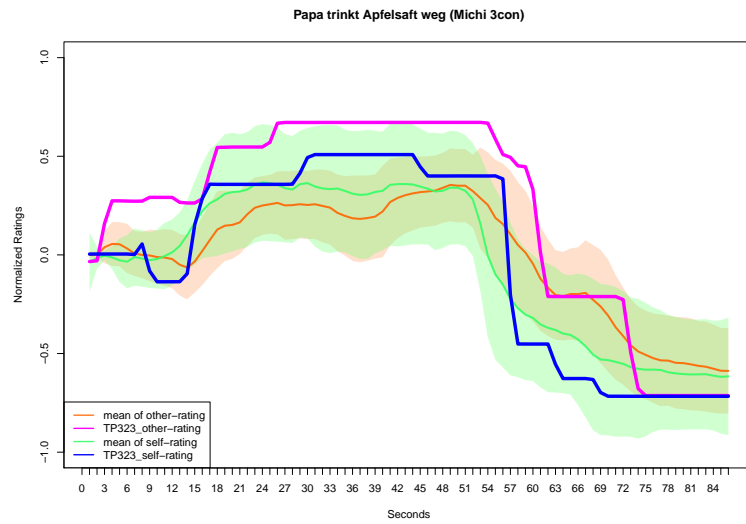
Figure 213: G3 Michi 3con: Ratings by test-person 323.
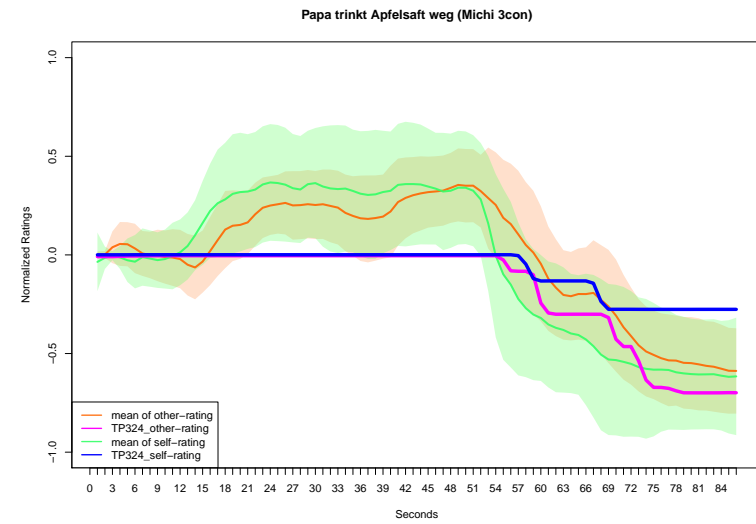
Figure 214: G3 Michi 3con: Ratings by test-person 324.
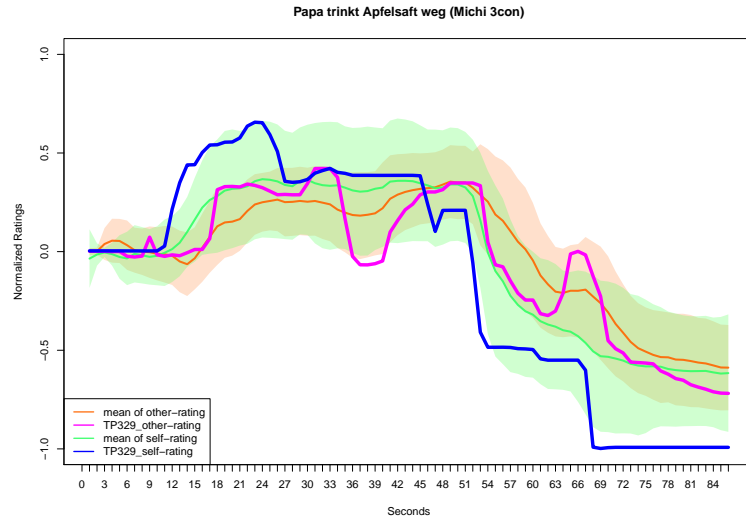
Figure 215: G3 Michi 3con: Ratings by test-person 329.
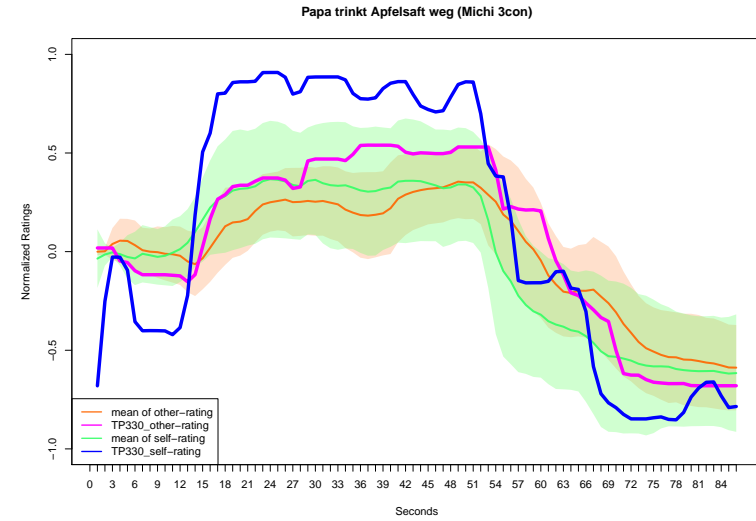
Figure 216: G3 Michi 3con: Ratings by test-person 330.
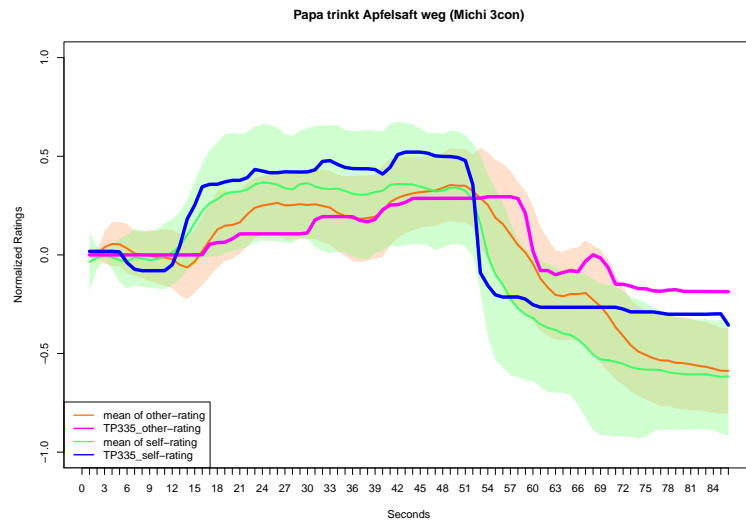
Figure 217: G3 Michi 3con: Ratings by test-person 335.
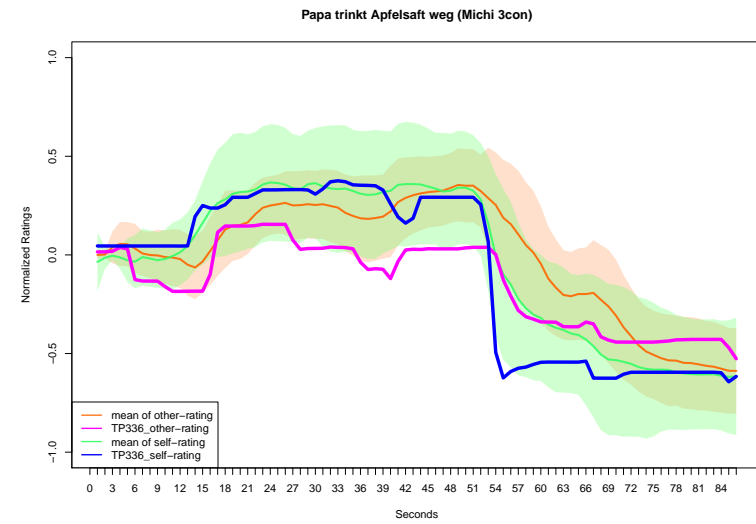
Figure 218: G3 Michi 3con: Ratings by test-person 336.
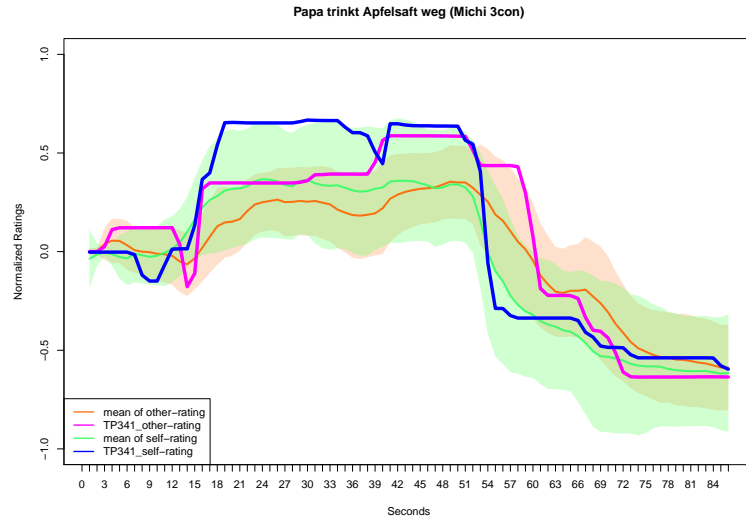
Figure 219: G3 Michi 3con: Ratings by test-person 341.
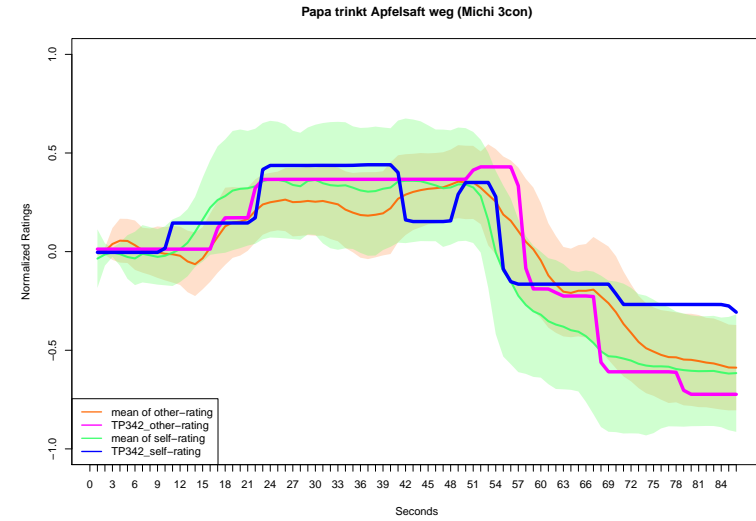
Figure 220: G3 Michi 3con: Ratings by test-person 342.
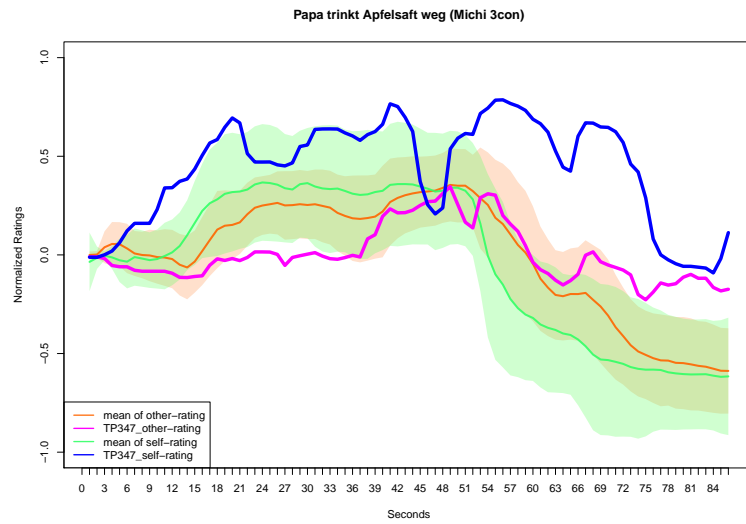
Figure 221: G3 Michi 3con: Ratings by test-person 347.
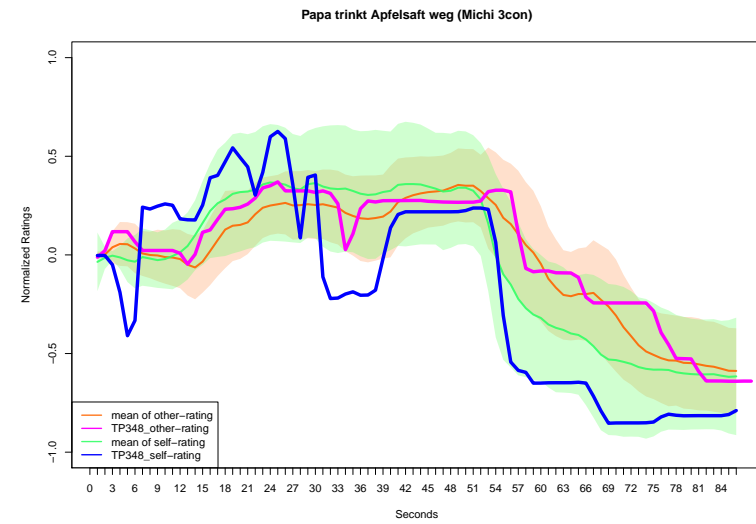
Figure 222: G3 Michi 3con: Ratings by test-person 348.
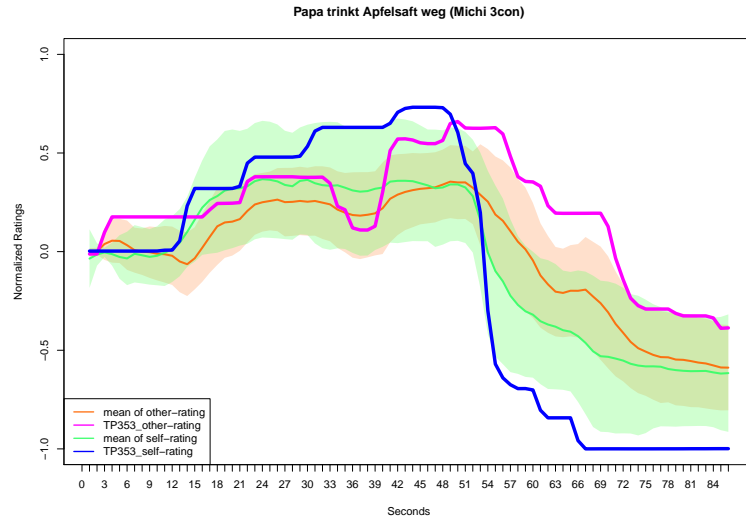
Figure 223: G3 Michi 3con: Ratings by test-person 353.
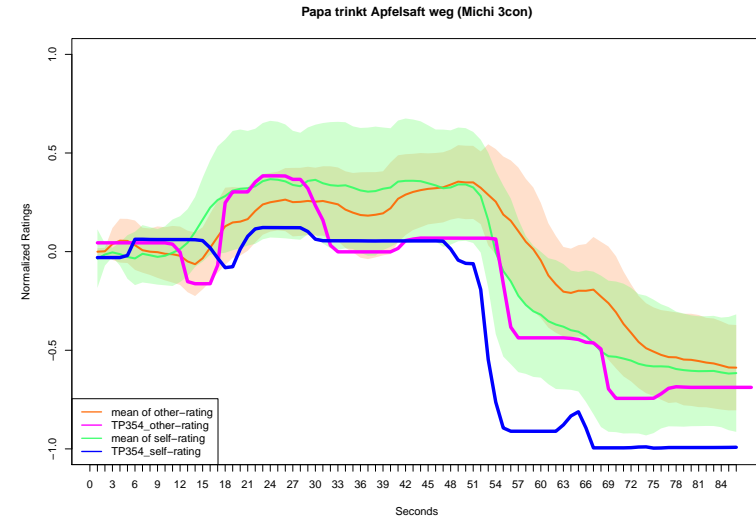
Figure 224: G3 Michi 3con: Ratings by test-person 354.
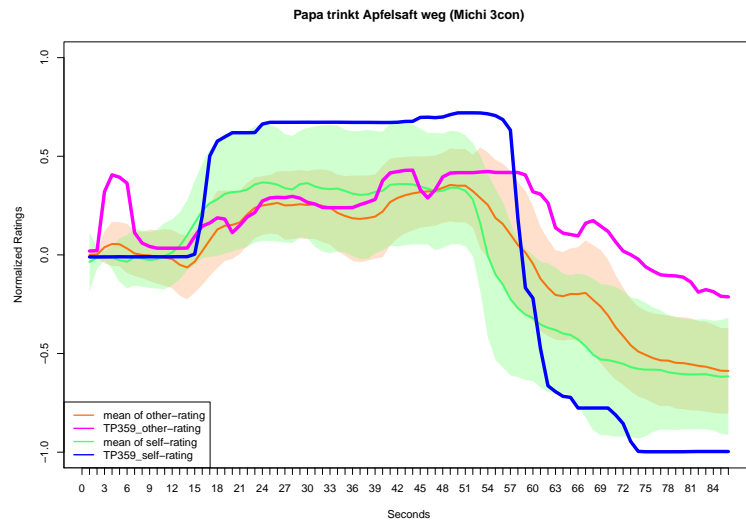
Figure 225: G3 Michi 3con: Ratings by test-person 359.
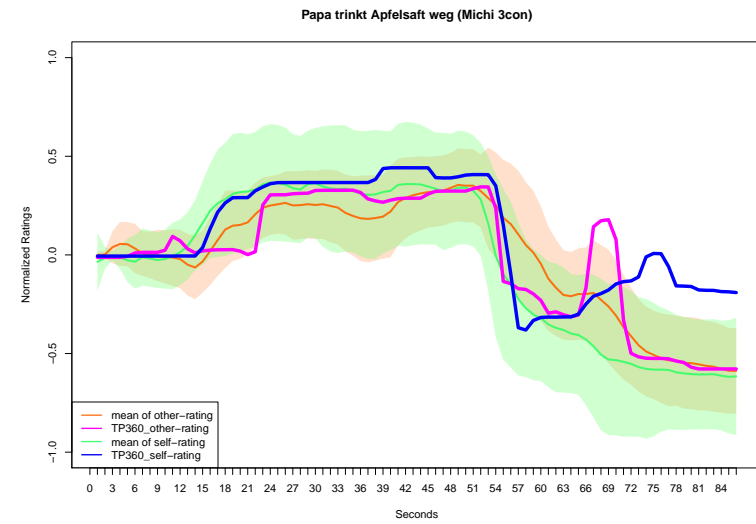
Figure 226: G3 Michi 3con: Ratings by test-person 360.
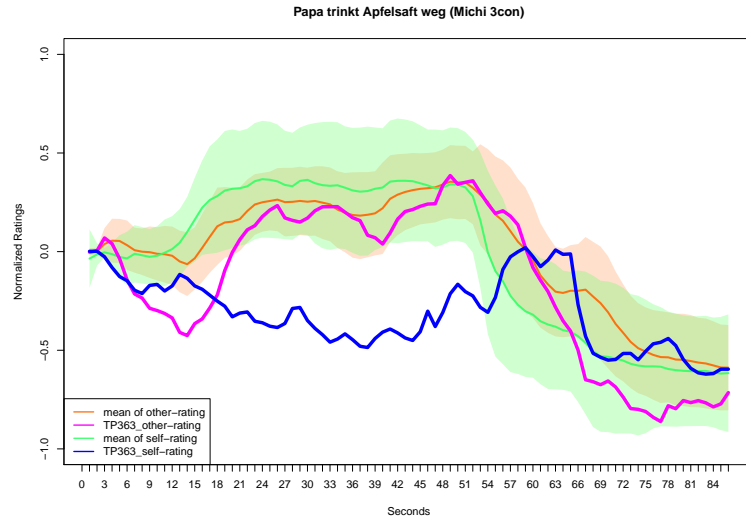
Figure 227: G3 Michi 3con: Ratings by test-person 363.

Figure 228: G3 Rudi 5inc: Ratings by test-person 305.
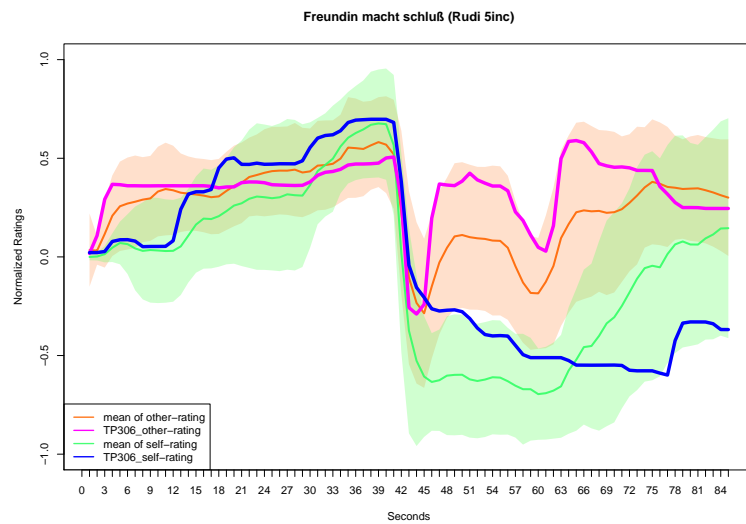
Figure 229: G3 Rudi 5inc: Ratings by test-person 306.
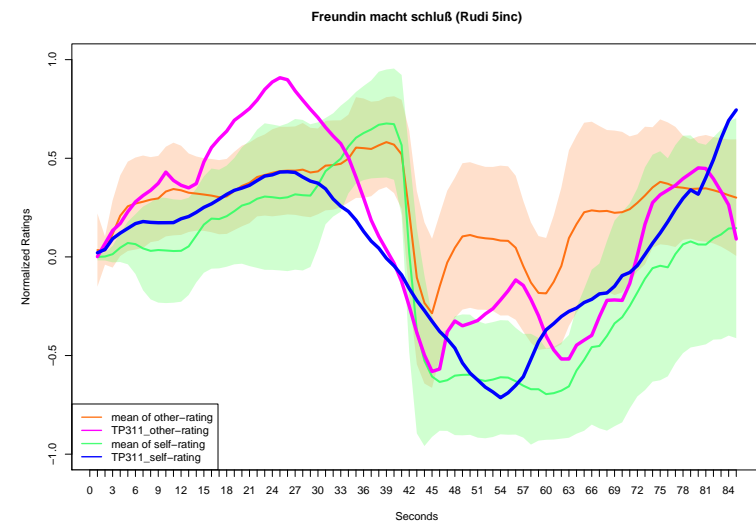
Figure 230: G3 Rudi 5inc: Ratings by test-person 311.
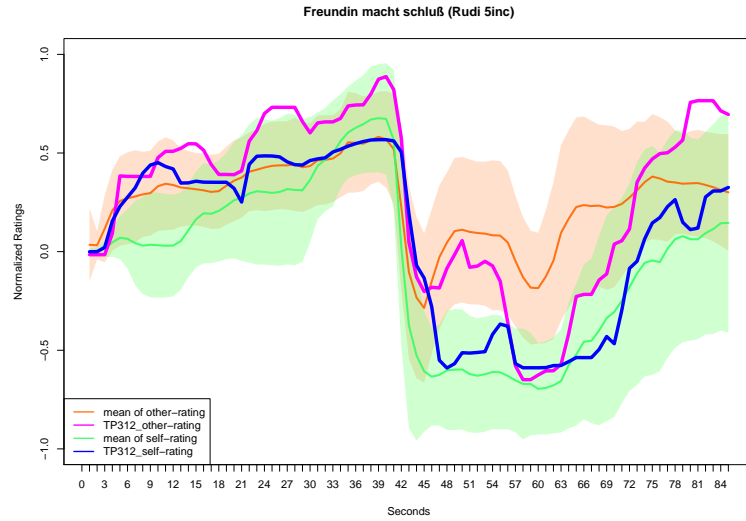
Figure 231: G3 Rudi 5inc: Ratings by test-person 312.
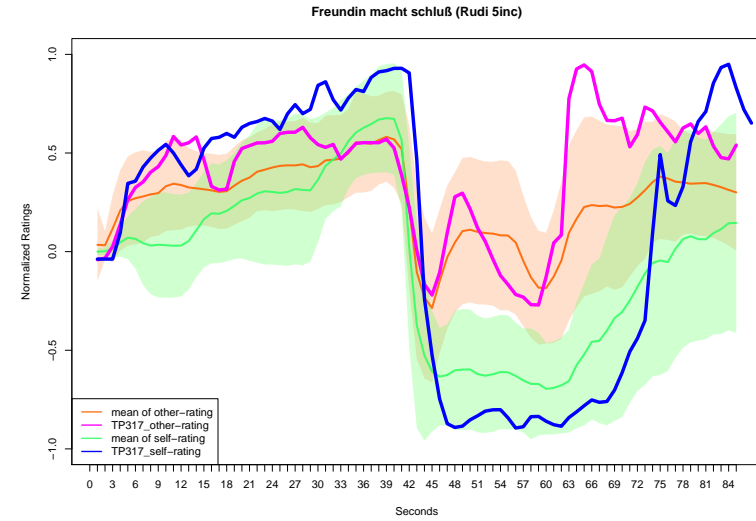
Figure 232: G3 Rudi 5inc: Ratings by test-person 317.
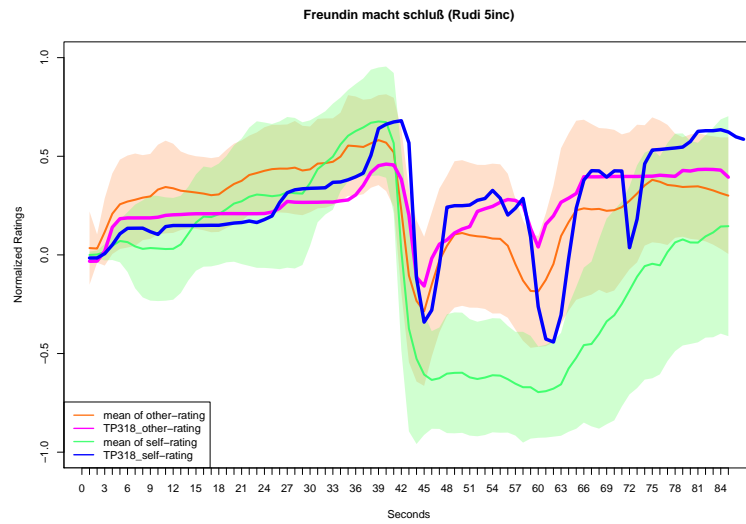
Figure 233: G3 Rudi 5inc: Ratings by test-person 318.
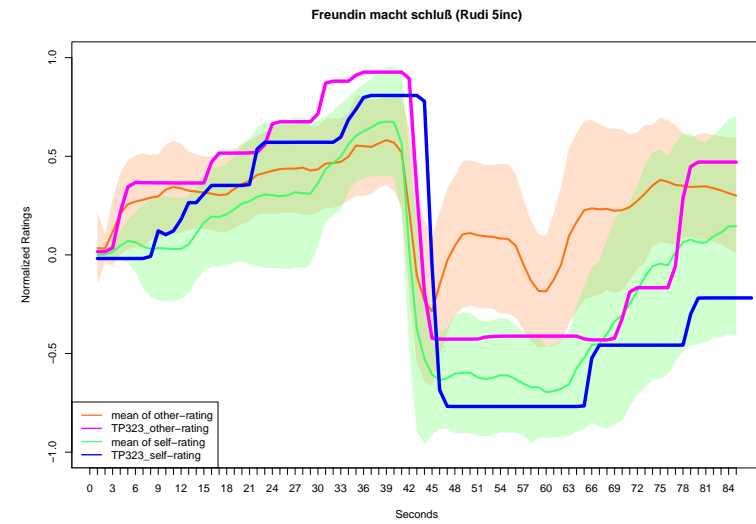
Figure 234: G3 Rudi 5inc: Ratings by test-person 323.
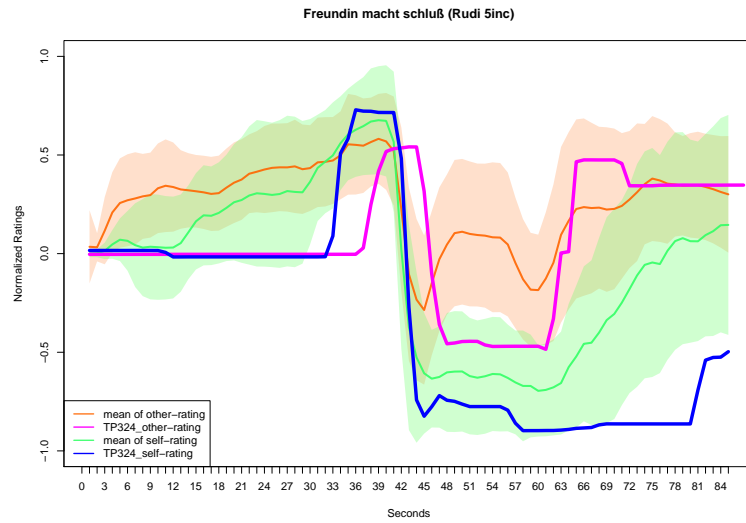
Figure 235: G3 Rudi 5inc: Ratings by test-person 324.
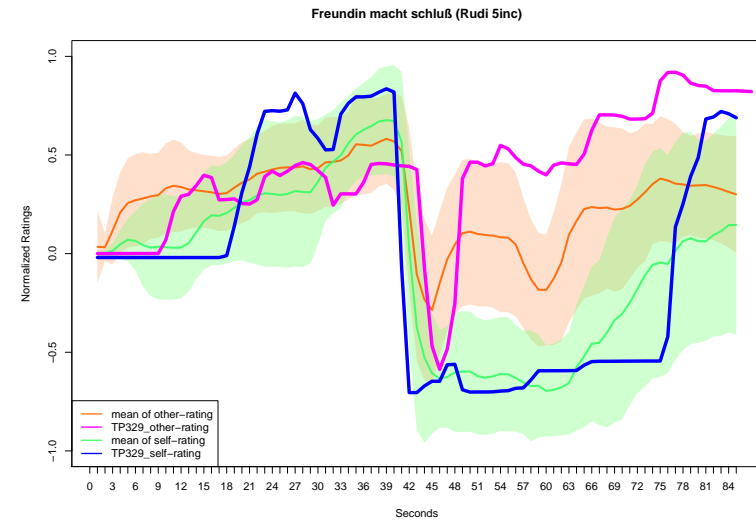
Figure 236: G3 Rudi 5inc: Ratings by test-person 329.
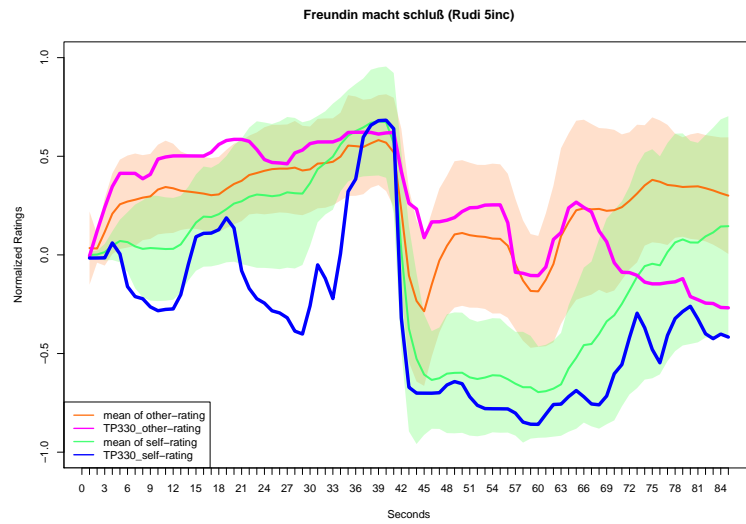
Figure 237: G3 Rudi 5inc: Ratings by test-person 330.
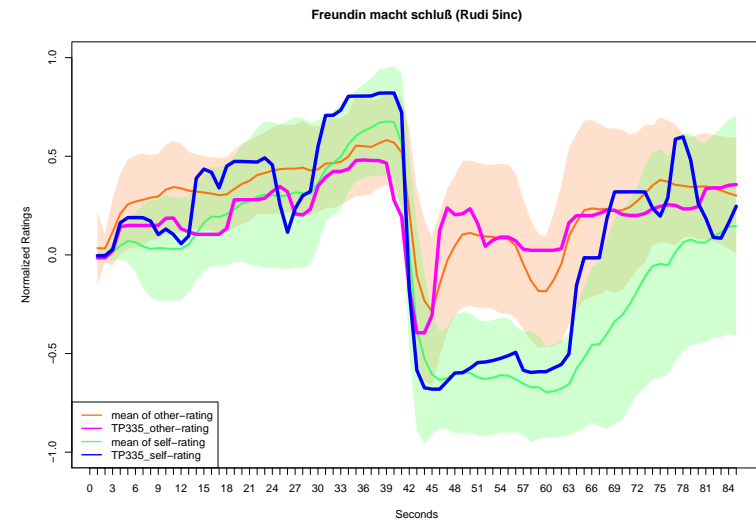
Figure 238: G3 Rudi 5inc: Ratings by test-person 335.
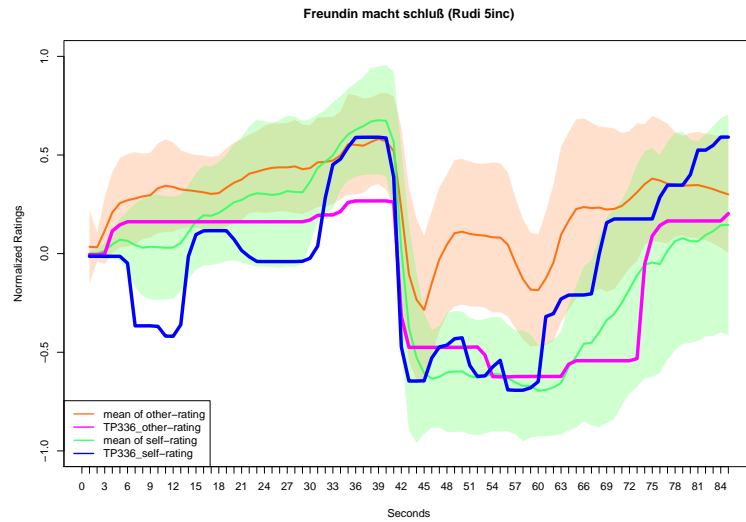
Figure 239: G3 Rudi 5inc: Ratings by test-person 336.
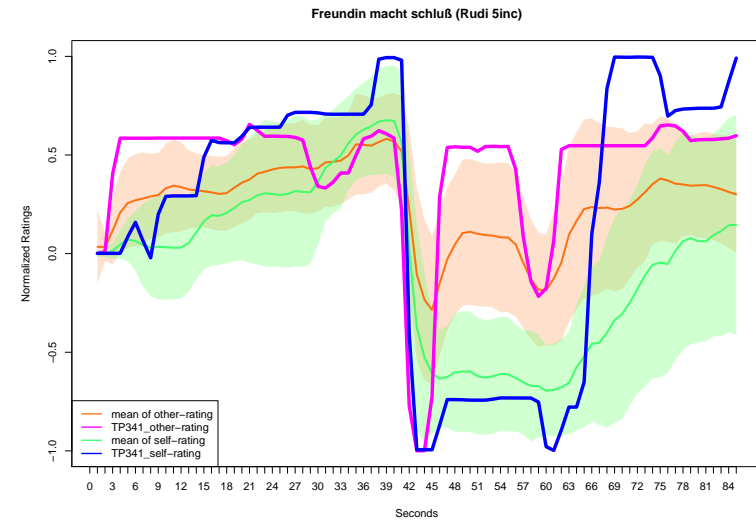
Figure 240: G3 Rudi 5inc: Ratings by test-person 341.
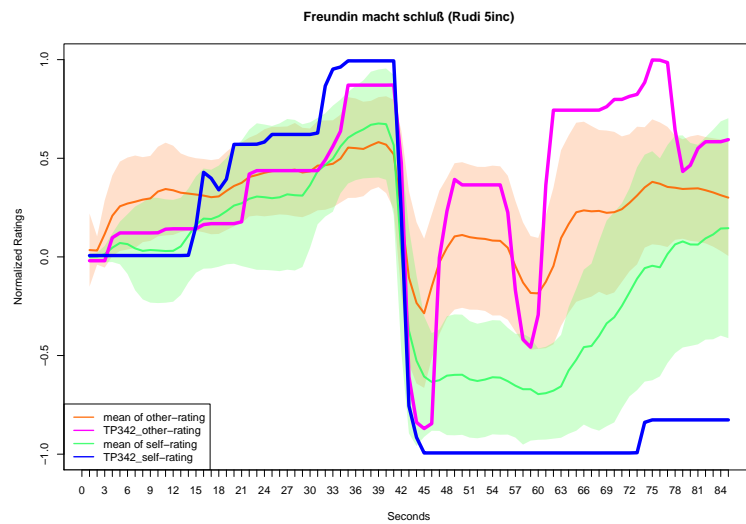
Figure 241: G3 Rudi 5inc: Ratings by test-person 342.
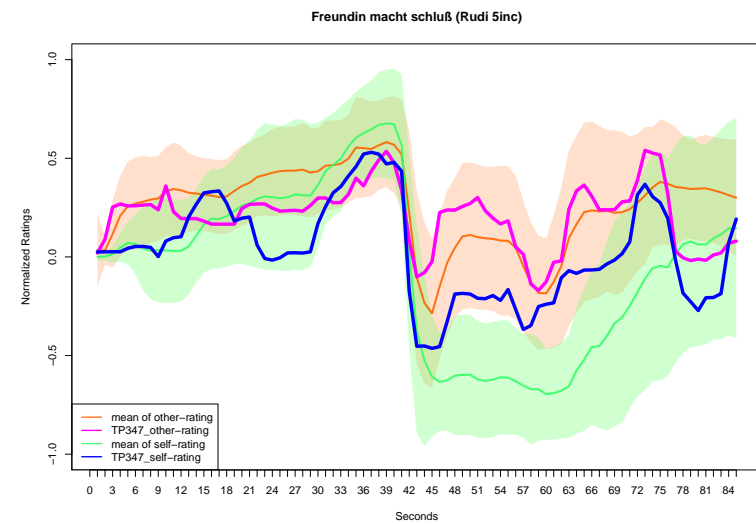
Figure 242: G3 Rudi 5inc: Ratings by test-person 347.
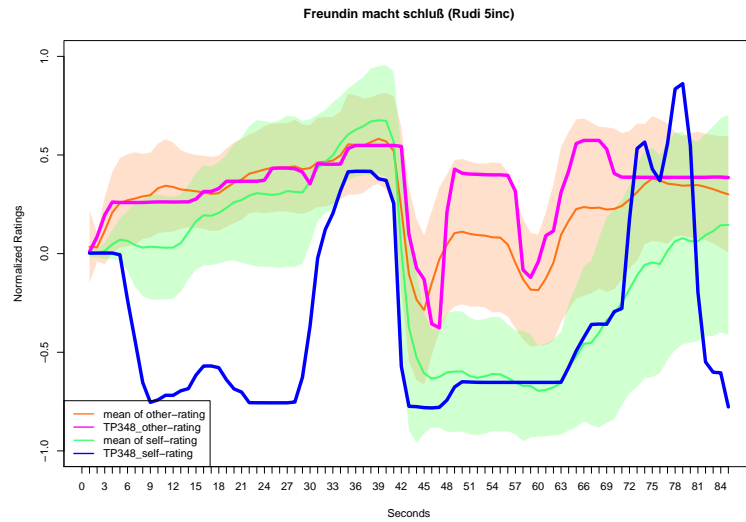
Figure 243: G3 Rudi 5inc: Ratings by test-person 348.
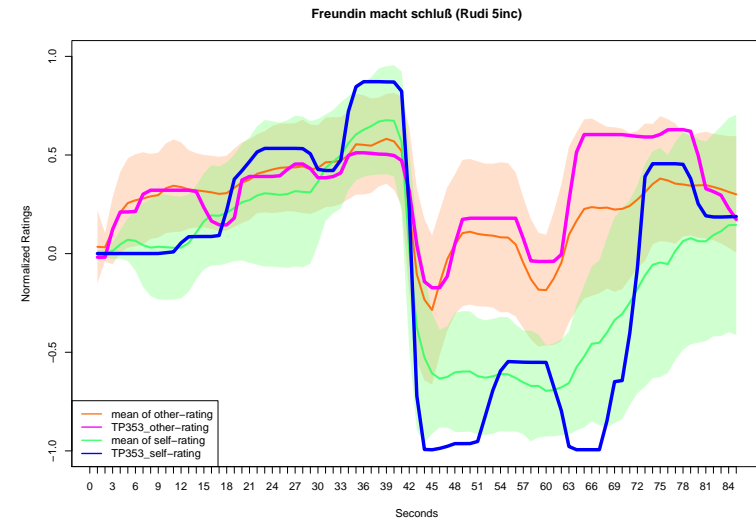
Figure 244: G3 Rudi 5inc: Ratings by test-person 353.
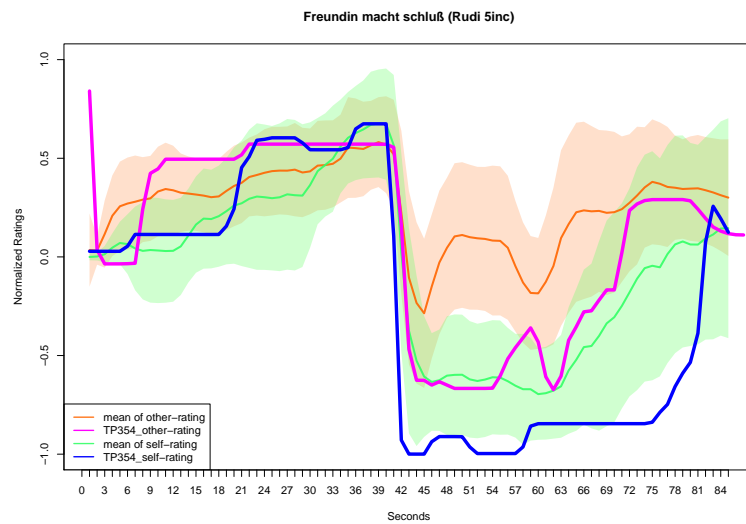
Figure 245: G3 Rudi 5inc: Ratings by test-person 354.
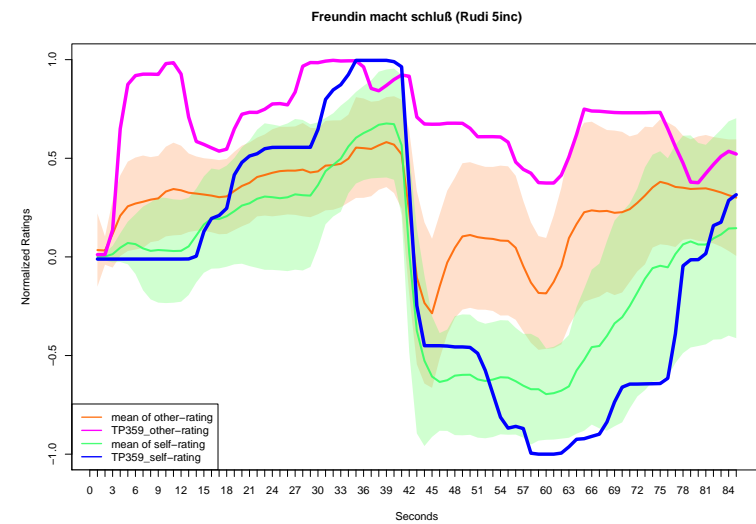
Figure 246: G3 Rudi 5inc: Ratings by test-person 359.
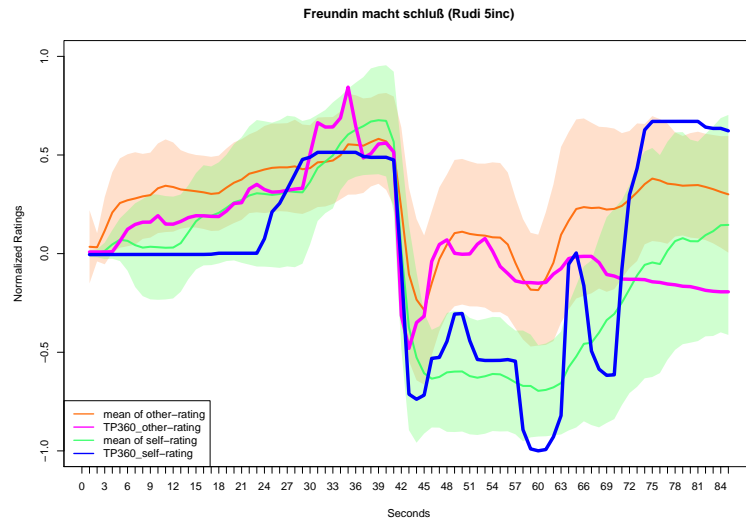
Figure 247: G3 Rudi 5inc: Ratings by test-person 360.
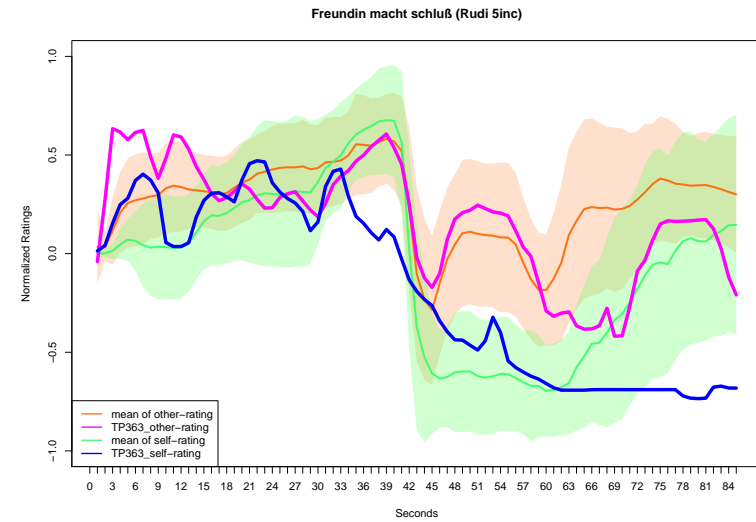
Figure 248: G3 Rudi 5inc: Ratings by test-person 363.
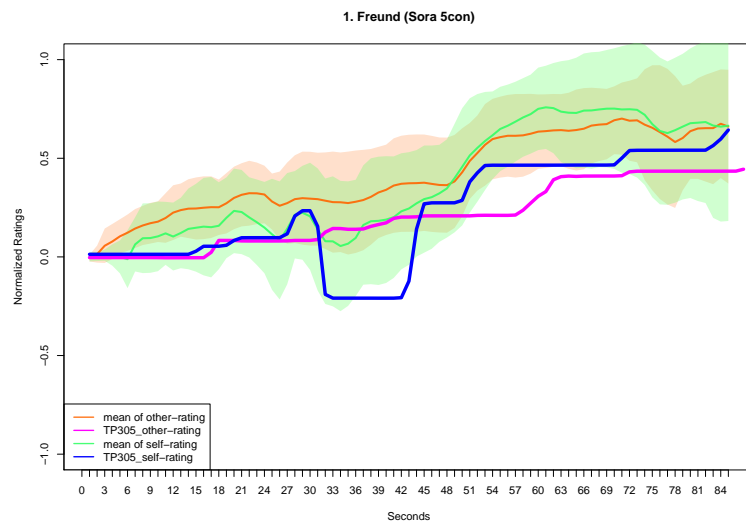
Figure 249: G3 Sora 5con: Ratings by test-person 305.
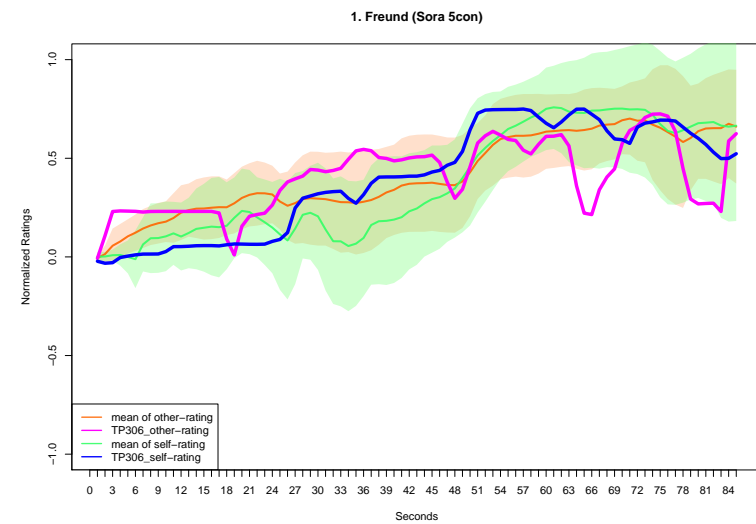
Figure 250: G3 Sora 5con: Ratings by test-person 306.
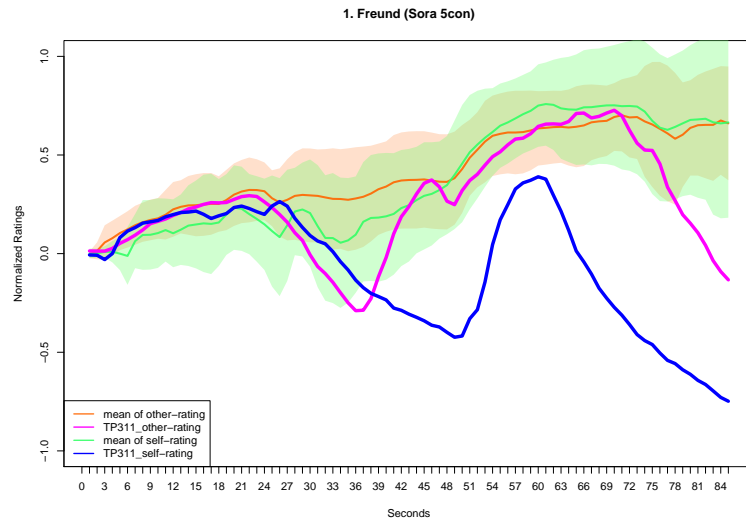
Figure 251: G3 Sora 5con: Ratings by test-person 311.
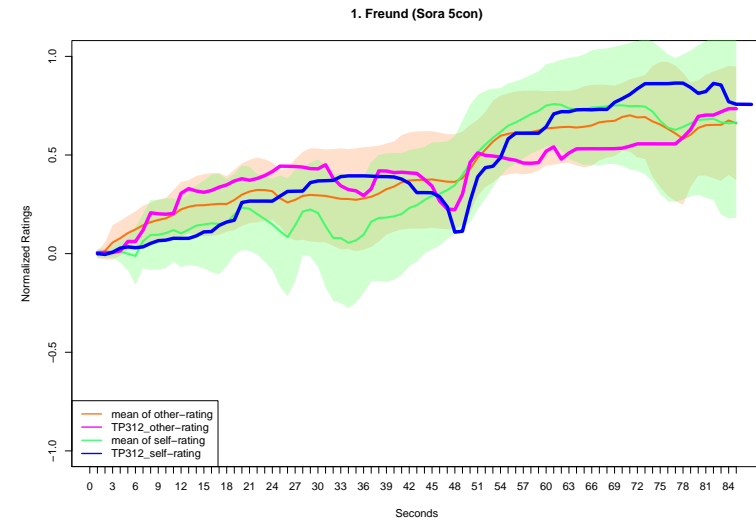
Figure 252: G3 Sora 5con: Ratings by test-person 312.
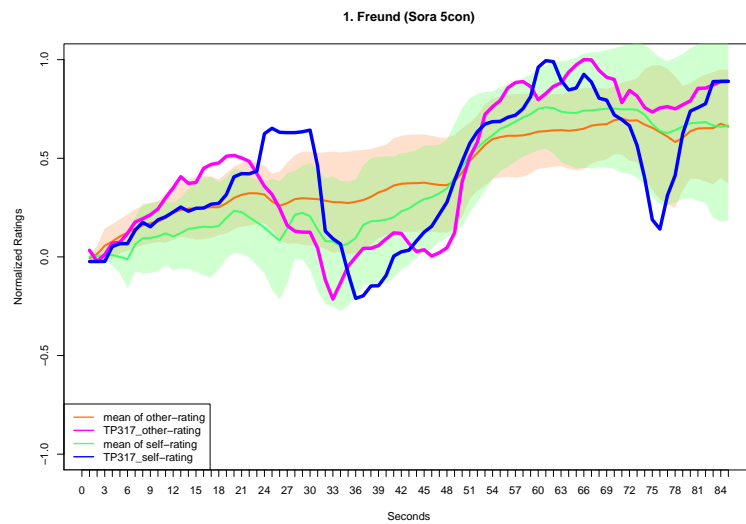
Figure 253: G3 Sora 5con: Ratings by test-person 317.
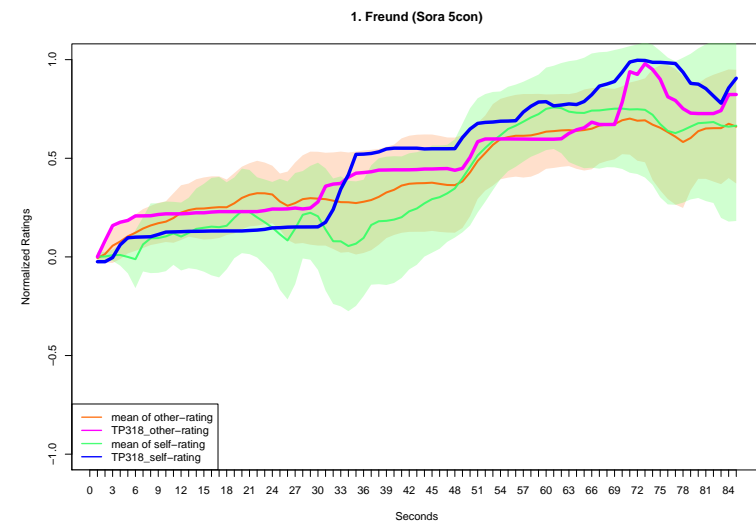
Figure 254: G3 Sora 5con: Ratings by test-person 318.
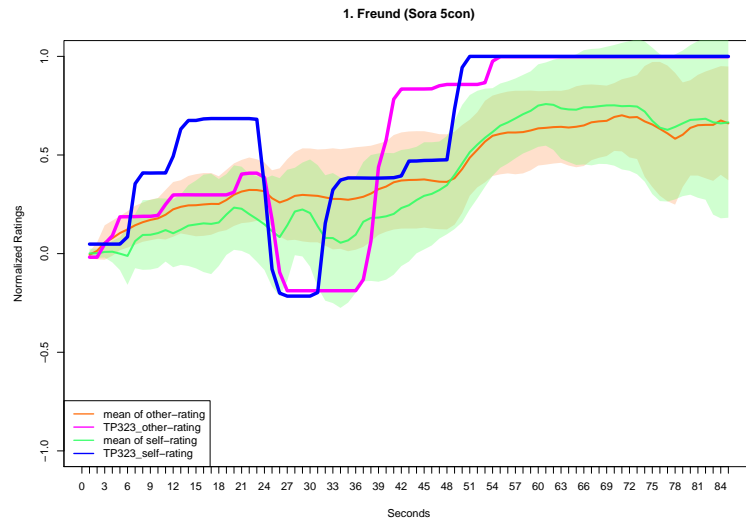
Figure 255: G3 Sora 5con: Ratings by test-person 323.
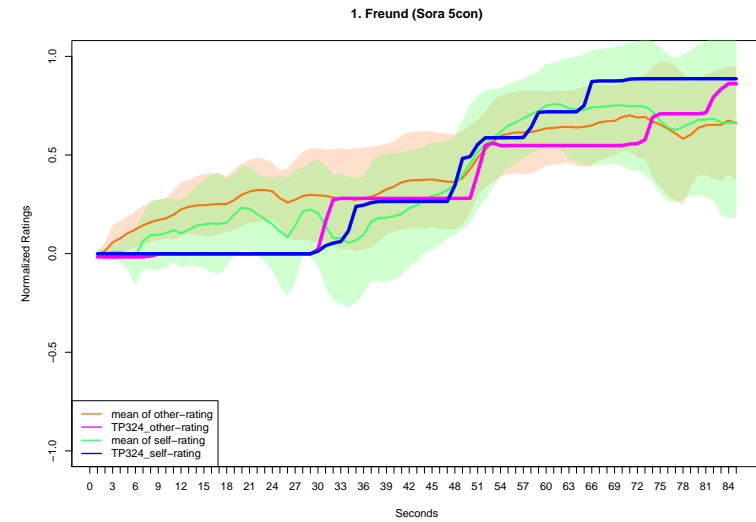
Figure 256: G3 Sora 5con: Ratings by test-person 324.
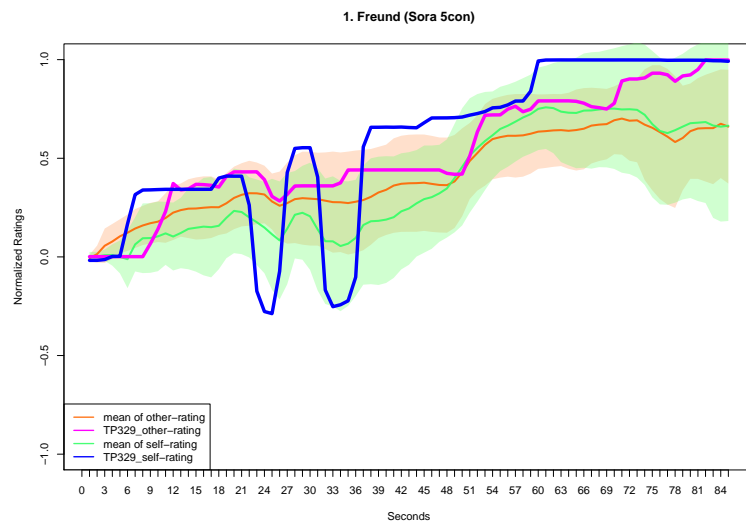
Figure 257: G3 Sora 5con: Ratings by test-person 329.
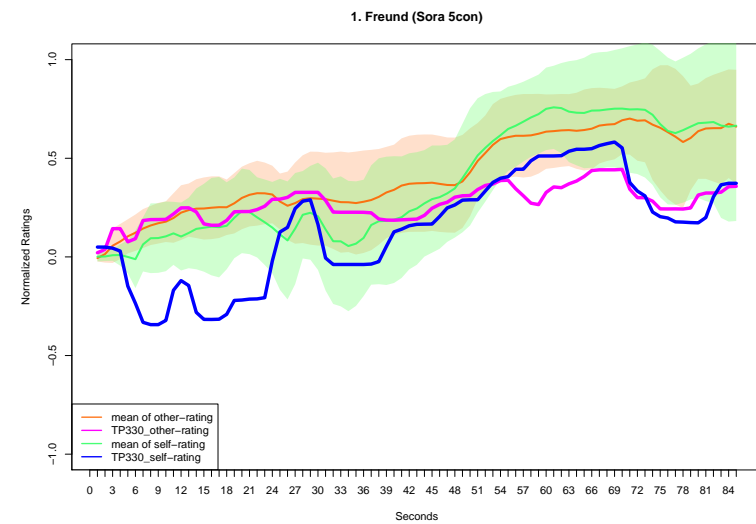
Figure 258: G3 Sora 5con: Ratings by test-person 330.
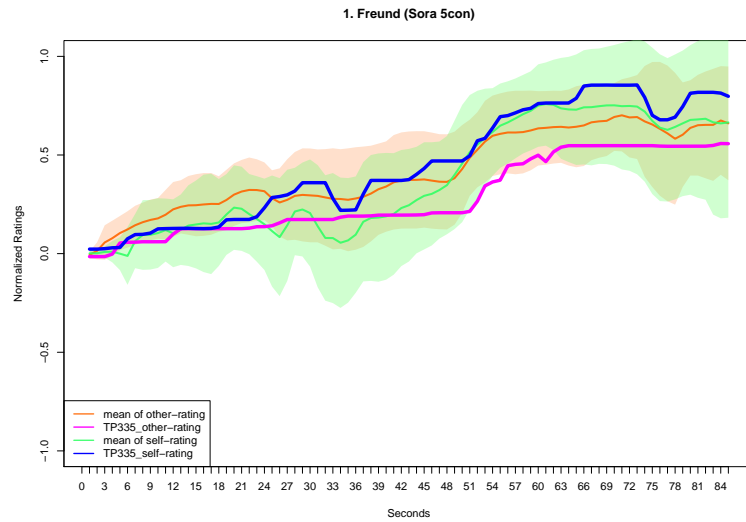
Figure 259: G3 Sora 5con: Ratings by test-person 335.
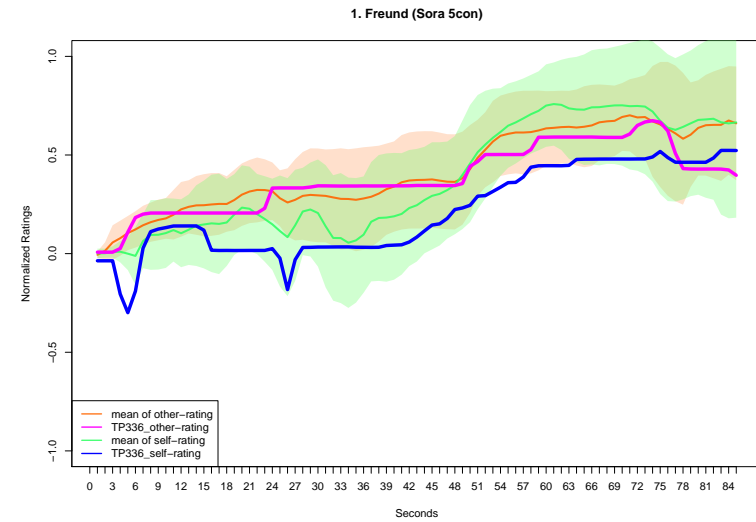
Figure 260: G3 Sora 5con: Ratings by test-person 336.
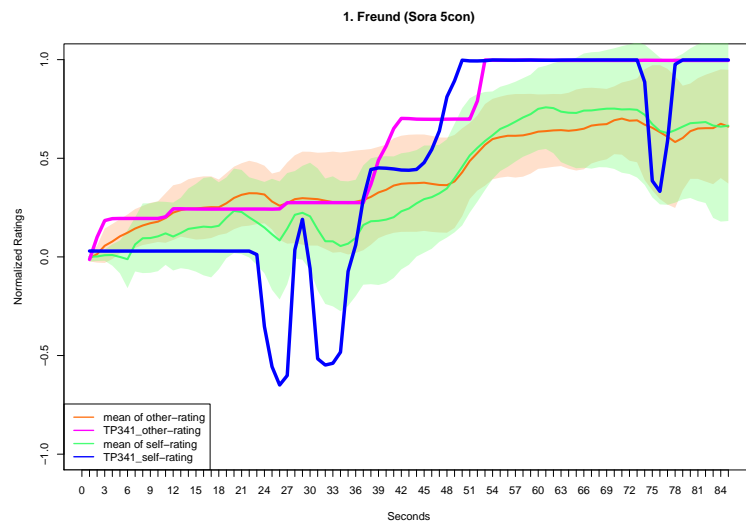
Figure 261: G3 Sora 5con: Ratings by test-person 341.
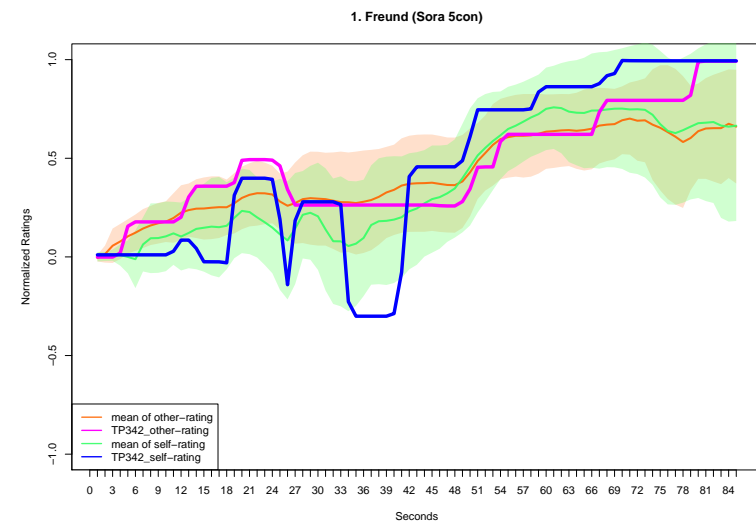
Figure 262: G3 Sora 5con: Ratings by test-person 342.
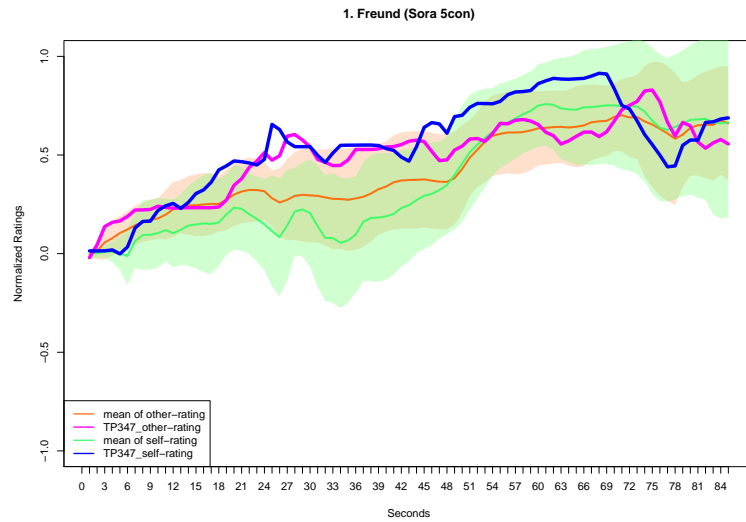
Figure 263: G3 Sora 5con: Ratings by test-person 347.
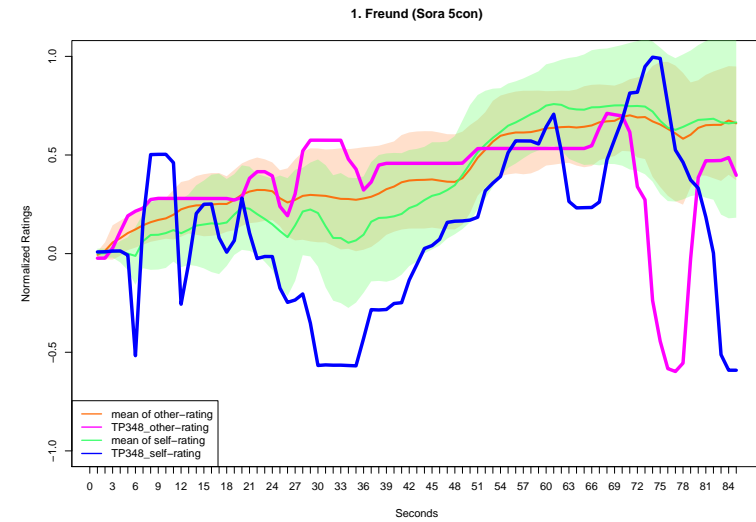
Figure 264: G3 Sora 5con: Ratings by test-person 348.
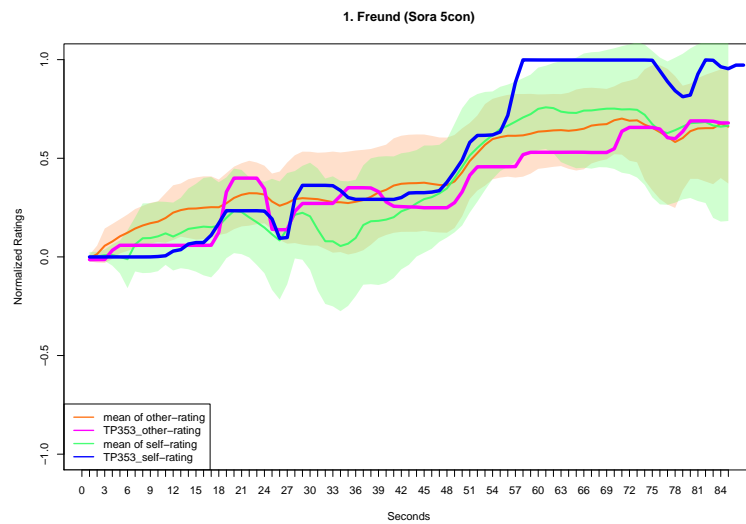
Figure 265: G3 Sora 5con: Ratings by test-person 353.
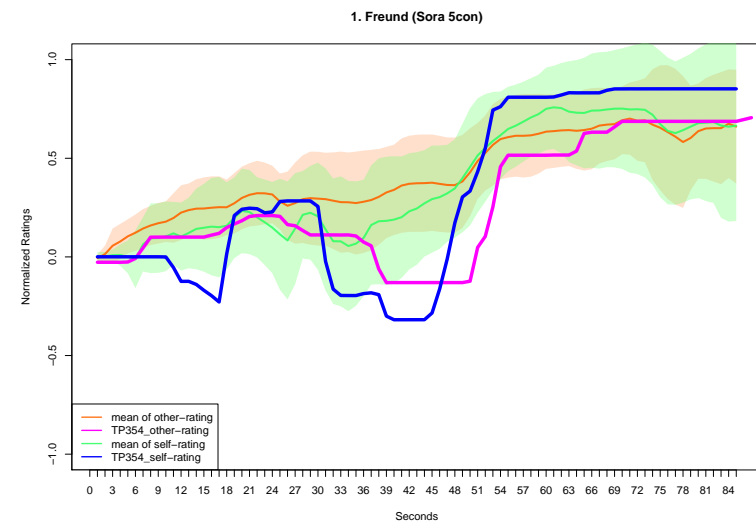
Figure 266: G3 Sora 5con: Ratings by test-person 354.
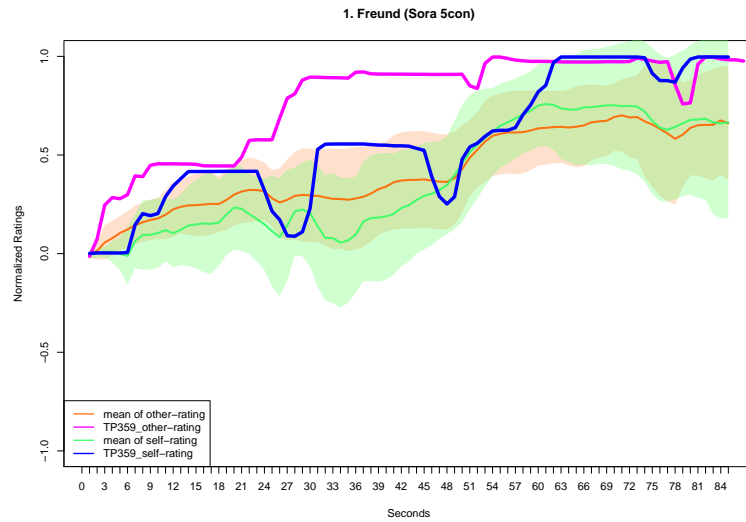
Figure 267: G3 Sora 5con: Ratings by test-person 359.
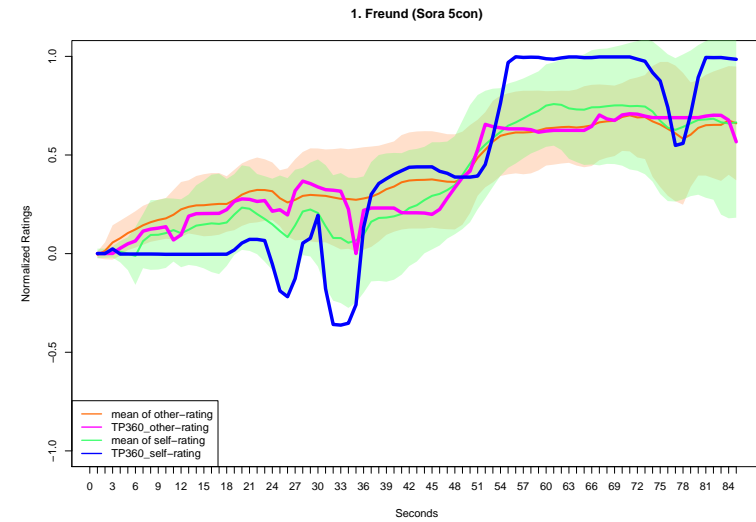
Figure 268: G3 Sora 5con: Ratings by test-person 360.
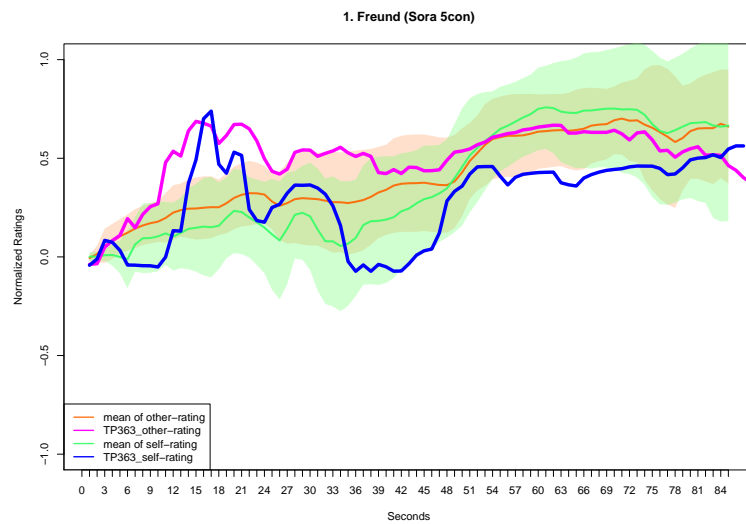
Figure 269: G3 Sora 5con: Ratings by test-person 363.
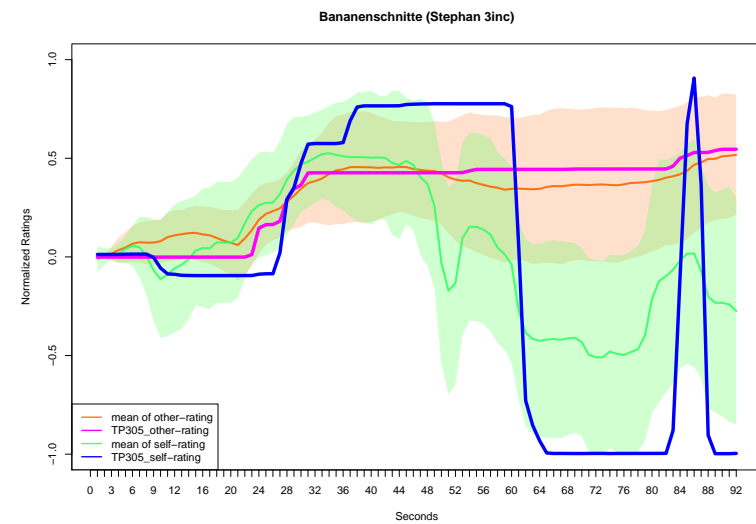
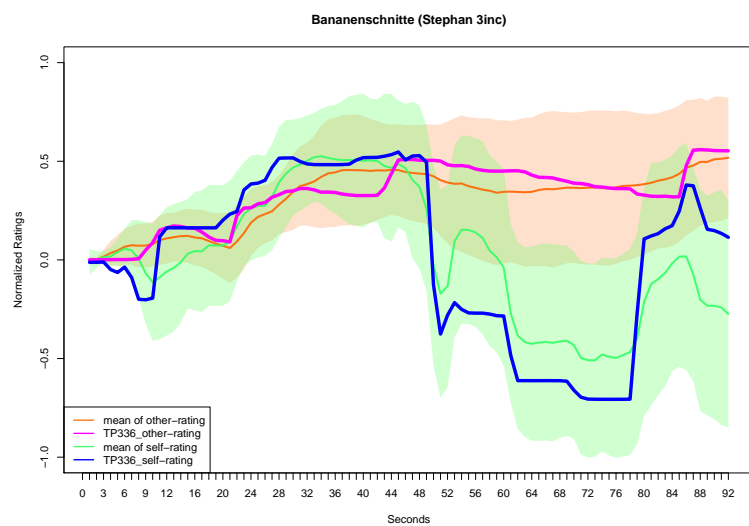**Figure 270:** G3 Stephan 3inc: Ratings by test-person 305.
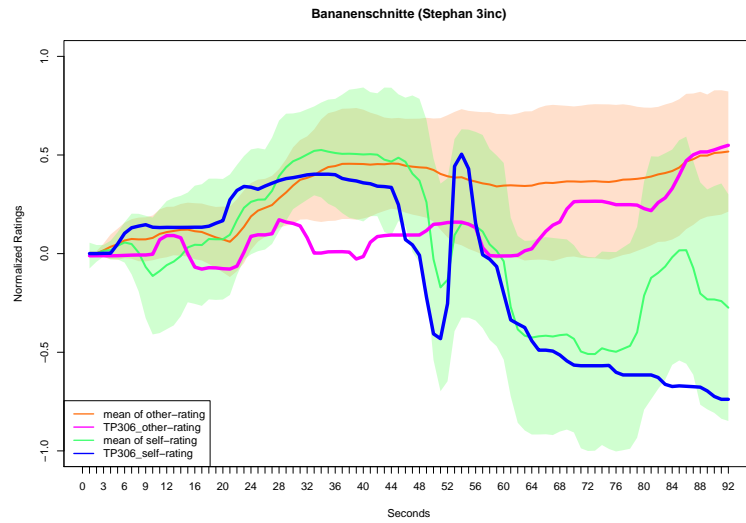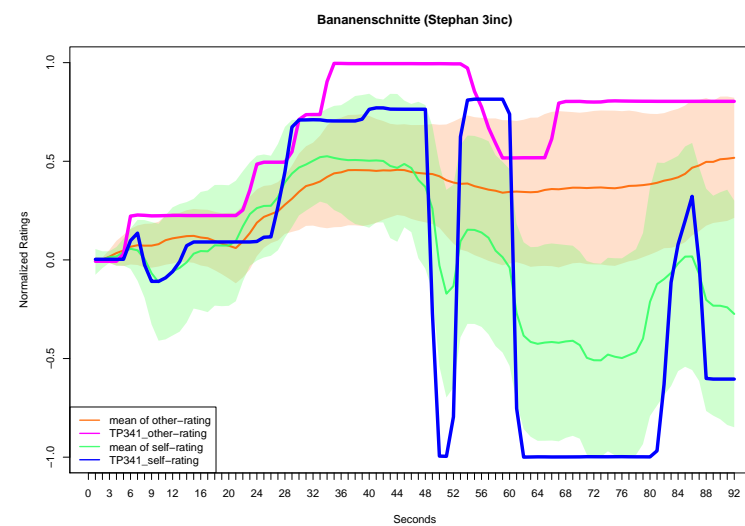
Figure 271: G3 Stephan 3inc: Ratings by test-person 306.
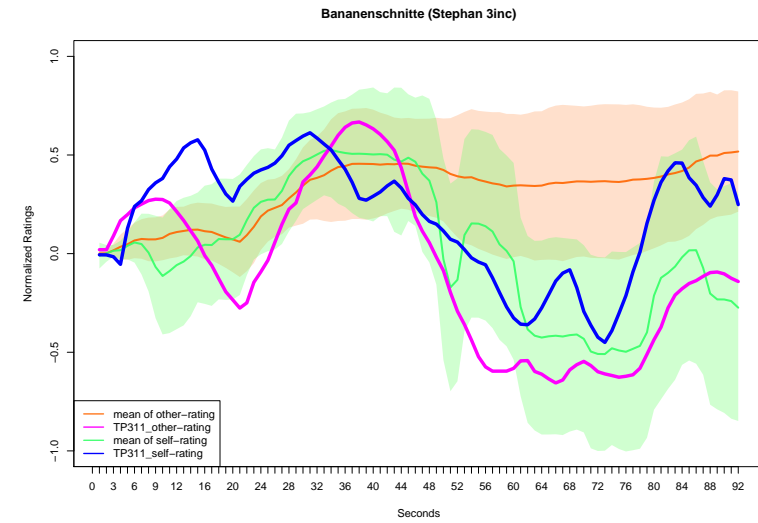
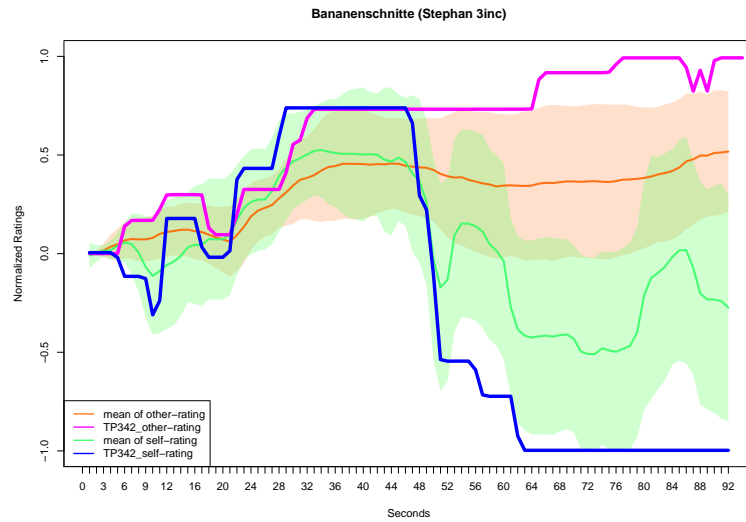**Figure 272:** G3 Stephan 3inc: Ratings by test-person 311.
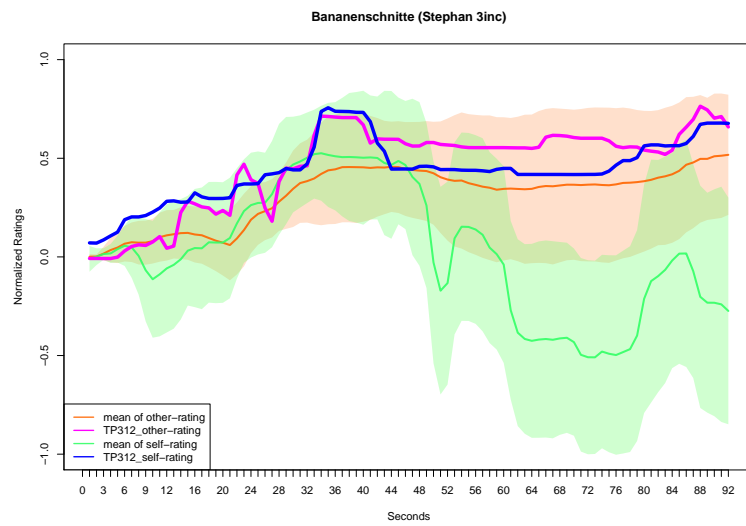
Figure 273: G3 Stephan 3inc: Ratings by test-person 312.
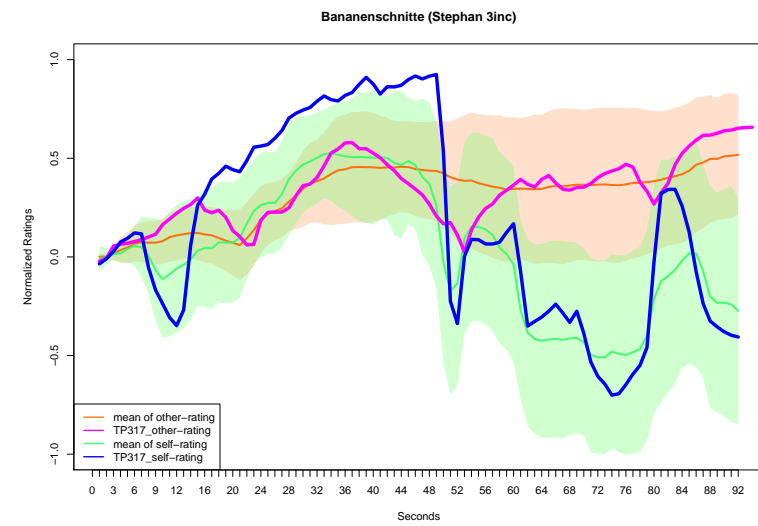
Figure 274: G3 Stephan 3inc: Ratings by test-person 317.
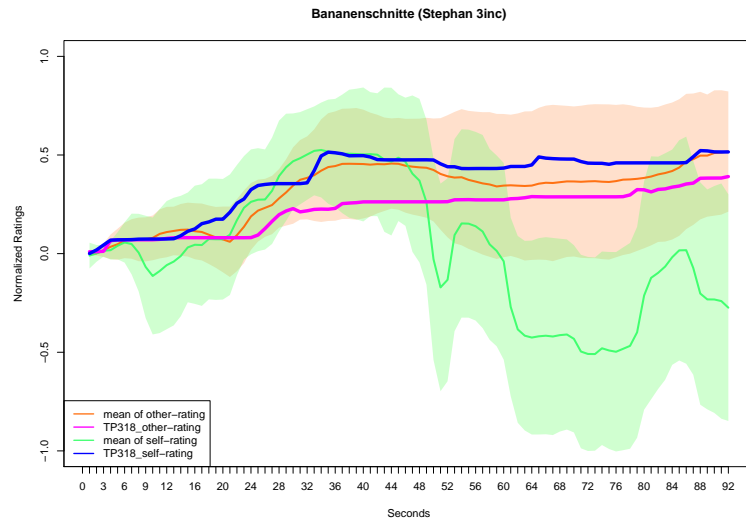
Figure 275: G3 Stephan 3inc: Ratings by test-person 318.
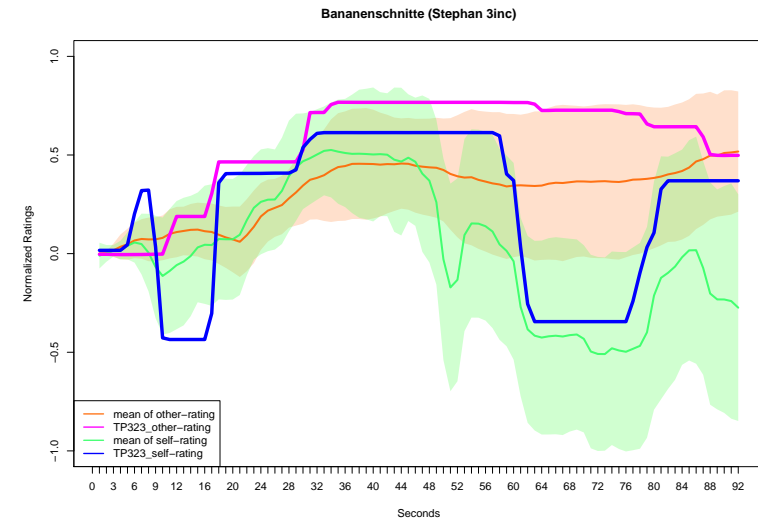
Figure 276: G3 Stephan 3inc: Ratings by test-person 323.
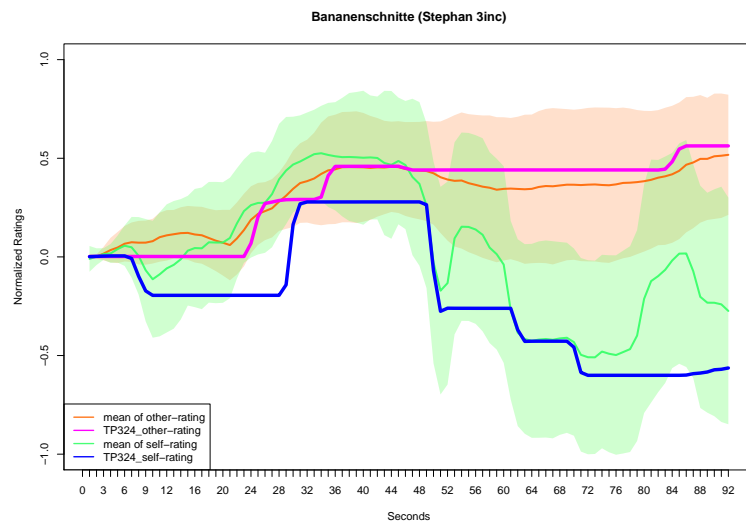
Figure 277: G3 Stephan 3inc: Ratings by test-person 324.
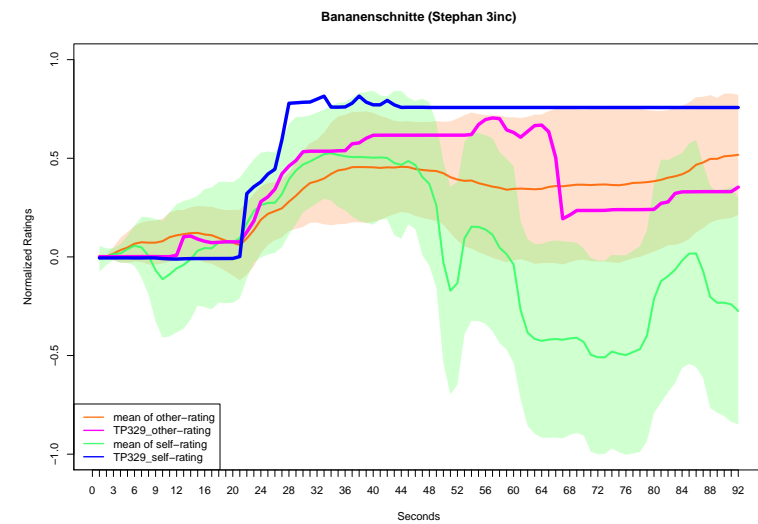
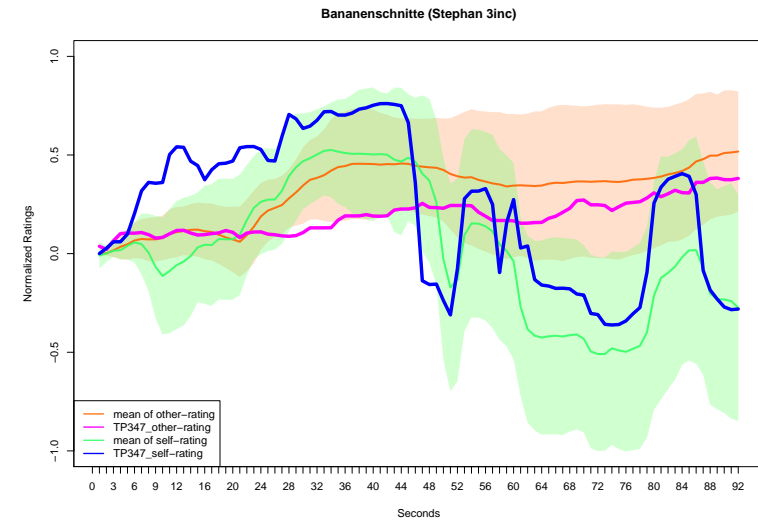**Figure 278:** G3 Stephan 3inc: Ratings by test-person 329.
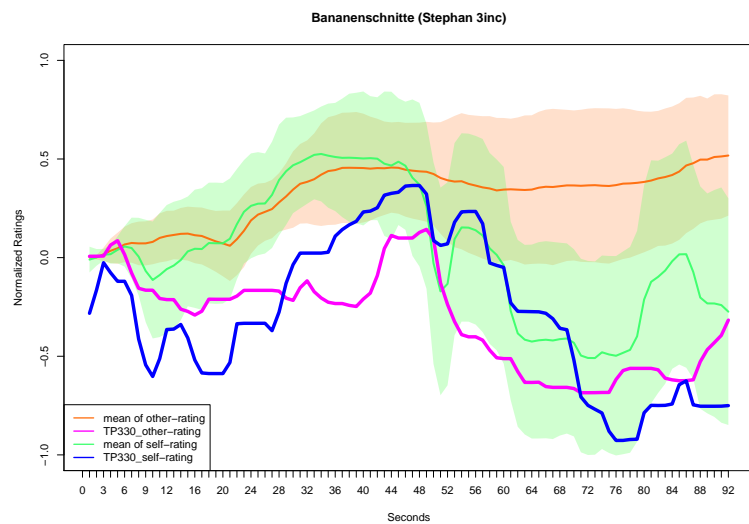
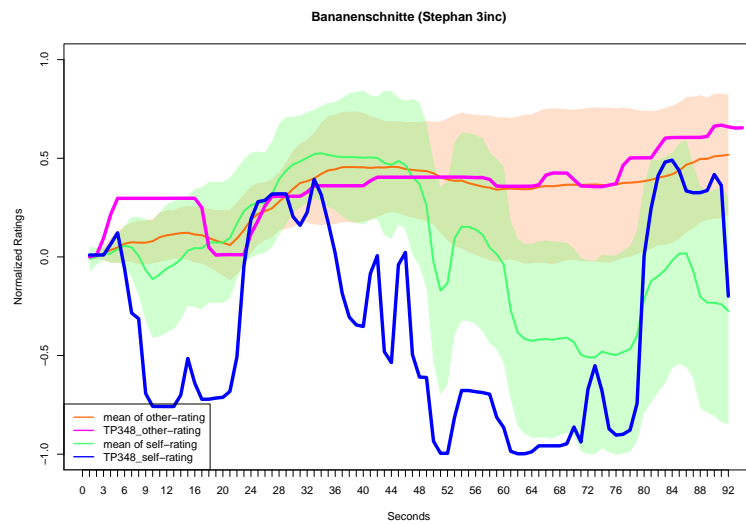**Figure 279:** G3 Stephan 3inc: Ratings by test-person 330.
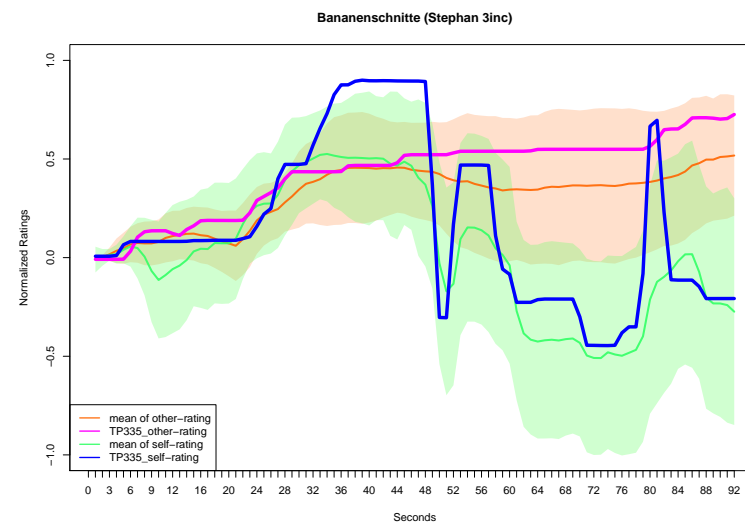
Figure 280: G3 Stephan 3inc: Ratings by test-person 335.

Figure 281: G3 Stephan 3inc: Ratings by test-person 336.

Figure 282: G3 Stephan 3inc: Ratings by test-person 341.

Figure 283: G3 Stephan 3inc: Ratings by test-person 342.

Figure 284: G3 Stephan 3inc: Ratings by test-person 347.

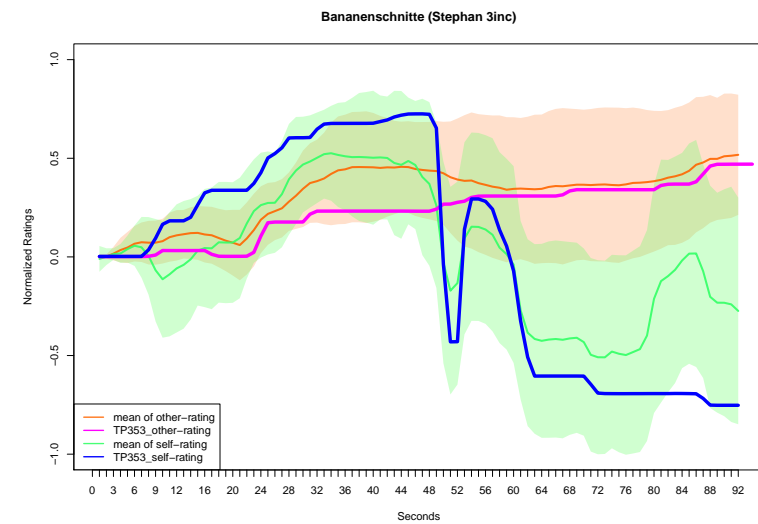
Figure 285: G3 Stephan 3inc: Ratings by test-person 348.

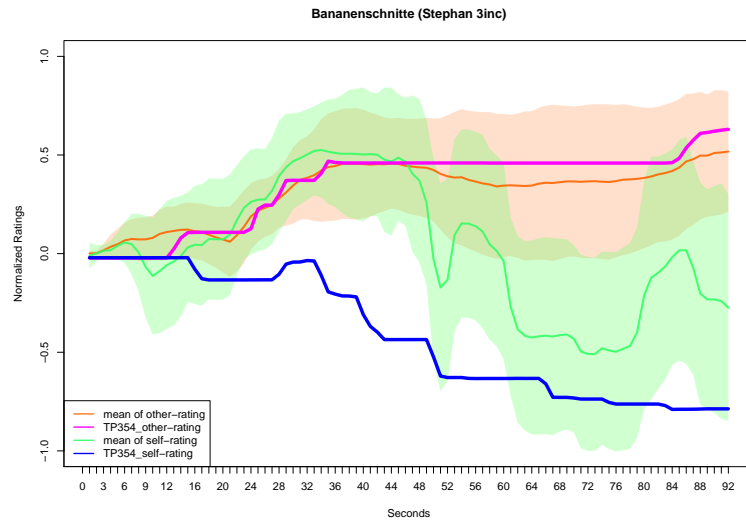**Figure 286:** G3 Stephan 3inc: Ratings by test-person 353.

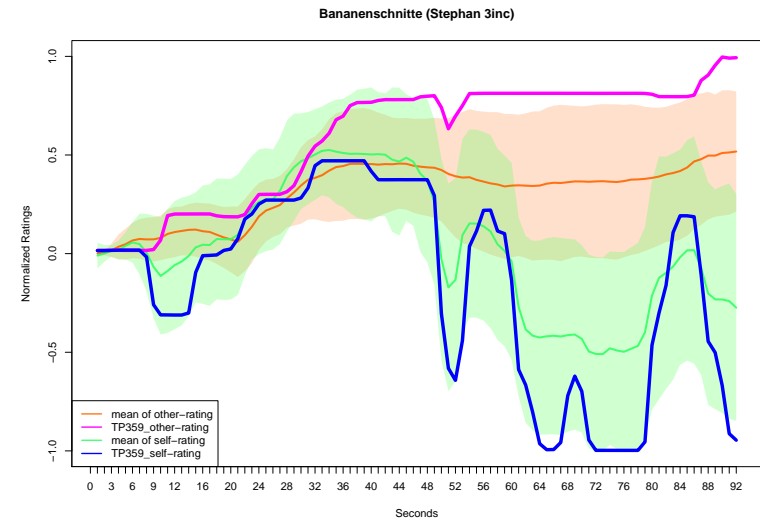**Figure 287:** G3 Stephan 3inc: Ratings by test-person 354.

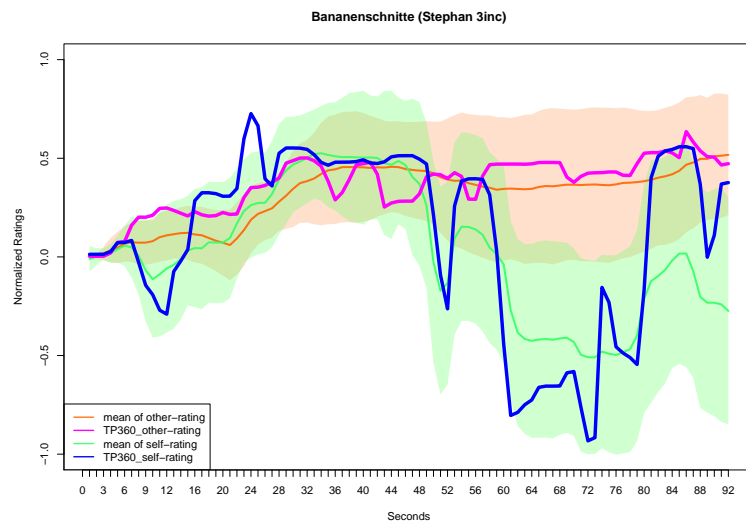**Figure 288:** G3 Stephan 3inc: Ratings by test-person 359.
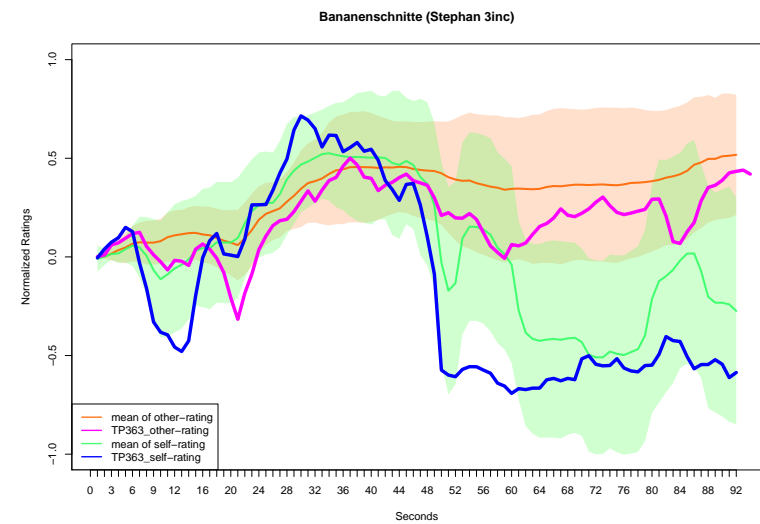

Figure 289: G3 Stephan 3inc: Ratings by test-person 360.

Figure 290: G3 Stephan 3inc: Ratings by test-person 363.

Curriculum Vitae

Angaben zur Person

Name	Angelika Längle
Geburtsdatum	02.07.1984
Geburtsort	Wien, Österreich
Staatsbürgerschaft	Österreich

Ausbildung

Seit Oktober 2012	**Hochschullehrgang Psychologisches Propädeutikum (HoPP)** (am Postgraduate Center der Universität Wien)
Seit Oktober 2005	**Psychologiestudium an der Universität Wien** Schwerpunkt: klinische Psychologie
1994 – 2002	**Rudolf Steiner-Schule Pötzleinsdorf**, Wien Abschluß mit Matura
1990 – 1994	**Volksschule Selzergasse**, Wien

Arbeitserfahrung

Juli/August 2011	6-Wochen-Praktikum in der Neuropsychologischen Sektion der Otto-von-Guericke-Universität (Universitätsklinik für Neurologie und Stereotaktische Neurochirurgie), Magdeburg/Deutschland. Arbeitsschwerpunkt war Testung und Befundschreiben von an Demenz erkrankten Personen.
2002	Sozialpraktikum in Ringwood, Großbritanien Kennenlernen sozial bedürftiger Menschen sowie Institutionen, die sich die Betreuung dieser Menschen zur Aufgabe gestellt haben, 3 Wochen
Mai 2015	Presentation of the validation of the "Test of Existential Motivations" (TEM) at the 1st World Congress of Existential Therapy, London, Great Britain.

Sprachen und besondere Kenntnisse

• Deutsch	Muttersprache
• Englisch	flüssig in Wort und Schrift
• Spanisch	flüssig in Wort und Schrift
• Russisch	Grundkenntnisse
• Französisch	Grundkenntnisse
• Österreichische Gebärdensprache (ÖGS)	Level A.2

- SPSS, LaTeX, E-Prime, R, Matlab
- Adobe PageMaker bzw. Adobe InDesign und Adobe Photoshop