The Processing of Moral Transgressions: Investigating the Role of Affective Evaluations

Dissertation

Mathematisch - Naturwissenschaftliche Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Angelika Ellen Kunkel

aus Altötting

Tübingen

2018

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen

Tag der mündlichen Qualifikation:	07.12.2018
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Hartmut Leuthold
2. Berichterstatter:	Prof. Dr. Barbara Kaup

Contents

Zusammenfassung

Im Alltag passiert es uns oft, dass wir Menschen begegnen, die entgegen unseres Verständnisses darüber, was richtig und was falsch ist, handeln. Dieses Verständnis soll hier unter dem Begriff der Moralität zusammengefasst werden. Seit vielen Jahrzehnten beschäftigen sich Wissenschaftler mit der Frage, wie wir unsere eigenen und die Handlungen anderer hinsichtlich ihres moralischen Gehalts bewerten, und welche Verarbeitungsprozesse dabei im Gehirn ablaufen. In jüngster Zeit wurde vorgeschlagen, dass affektiven im Vergleich zu kognitiven Prozessen eine wichtige Rolle bei moralischen Entscheidungen zukommt. Bisher fehlt es jedoch an einem tieferen Verständnis über die zugrundeliegenden Verarbeitungsmechanismen. Die vorliegende Dissertationsschrift, welche sich dreier Studien bedient, soll einen Beitrag zur Klärung dieser Frage leisten. Hierfür werden bereits bekannte Paradigmen herangezogen und moralisches Sprachverstehen als Spezialfall des emotionalen Sprachverständnisses betrachtet.

Alle drei Studien bauen auf der Vorstellung auf, dass die Verarbeitung moralischer Verletzungen eine wichtige affektive Komponente besitzt. Deshalb wurde in allen Studien neben moralischen Texten zusätzlich die Verarbeitung emotionaler Texte betrachtet. In Studie 1 wurde mittels ereigniskorrelierter Potentiale (EKPs) der Frage nachgegangen, wie während des Lesens kurzer Alltagsszenarien die Verarbeitung moralischer Verletzungen im Gehirn abläuft und ob diese Verarbeitung von der Art der Aufgabe abhängt. Dabei lag der Fokus auf dem zeitlichen Verlauf mentaler Mechanismen bei moralischen und bei affektiven Urteilen. In Studie 2 wurde die Frage untersucht, ob moralische Verletzungen verkörperlicht verarbeitet werden. Mittels der Messung peripherer psychophysiologischer Methoden wurde untersucht, ob beim Lesen konkrete emotionale Zustände simuliert werden. Studie 3 untersuchte sowohl für EKPs, als auch für periphere psychophysiologische Maße die Art der beteiligten Prozesse bei der Verarbeitung moralischer Information beim reinen Lesen.

Zusammenfassend weisen die Ergebnisse der vorliegenden Studien auf eine maßgebliche Beteiligung affektiver Prozesse bei der Verarbeitung moralischer Information hin. Dies spiegelt sich sowohl im EKP, als auch in der verkörperten Verarbeitung moralischer Inhalte wider. Des Weiteren legen die Ergebnisse eine Aufgabenabhängigkeit der beteiligten Prozesse nahe. Beim Lesen und beim emotionalen Urteil wird moralische Information affektiv verarbeitet, wohingegen beim moralischen Urteil kognitive Verarbeitung dominiert.

Abstract

Every day we are confronted with other's actions which clash with our understanding about what is right and what is wrong. This understanding is summarised here by the concept of morality. For many decades, scientists have been investigating how we evaluate the morality of actions and which brain processes underlie such evaluations. For instance, recently it has been suggested that affective as compared to cognitive processes might play a more fundamental role in moral judgement than previously assumed. So far, however, it is still unclear what the underlying mechanisms exactly are. This dissertation, which is based on three studies, intends to contribute to answering this question. By using a language comprehension approach, moral information processing is taken as a special case of emotional information processing.

All three studies are based on the assumption that affective processes play an important role in the processing of moral transgressions. Therefore, in addition to moral texts, the processing of emotional texts was also considered in all studies. In study 1, event-related brain potentials (ERPs) were used to investigate how moral transgressions are processed in the brain during discourse comprehension of everyday scenarios and whether affective processing is task-dependent. The focus was on the time course of mental mechanisms during a morality and an emotional judgement task. Study 2 aimed at examining whether moral transgressions during reading of the same everyday scenarios trigger embodied responses. By using peripheral psychophysiological measures, we investigated whether embodied processing takes place when participants perform an emotional judgement task. Finally, Study 3 analysed both, ERPs and peripheral psychophysiological measures to investigate the type of processes involved in moral information processing when participants merely read for comprehension.

In summary, the results of the studies reported in this thesis point towards the involvement of affective processes in the processing of moral content. This is reflected by ERP correlates indicating affective processing and by peripheral psychophysiological measures suggesting the embodiment of moral content. Moreover, the results suggest a task dependency of involved processes. This means that the goals with which one reads a text influence whether moral information is emotionally salient or not. As a result, cognitive processes dominate over affective processes when moral judgements are demanded and vice versa when affective judgements are required or participants merely read for comprehension.

1 Introduction

As social animals, humans are confronted with incoming information about the actions of others. This information must be rapidly evaluated and the meaning as well as consequences must be processed. For instance, we sometimes find ourselves in situations in which we hear about a person violating a prevailing moral norm (a woman having an abortion) or breaching the law (evading taxes). Often, we experience such actions that violate our sense or our knowledge of rightness and wrongness aversively and spontaneously tend to judge such actions as bad, unsocial, or immoral.

Moral values are standards that an individual or a group holds and on the basis of which we judge what is right or wrong, good or evil. Moral standards regulate the actions of individuals and make social life possible in the first place. In brief, morality is about protecting individuals within a social world (Haidt, 2007). Moreover, when interacting with others, in order to avert threat to the individual, it is important to quickly evaluate others' actions and to identify persons who intend to harm us. This reflects a fundamental aspect of human moral cognition which will be examined in the present thesis. Here, I will adopt a more general concept of morality including moral principles (e.g., "You should not kill.") as well as social conventions (e.g., "As a man you use the men's restroom.").

The primary aim of investigating moral judgement behaviour is to understand how people judge what is right or wrong (Waldmann, Nagel, & Wiegmann, 2012). Over time, various theories have been developed which emphasise different aspects of moral judgement. The starting point for the systematic experimental investigation of moral judgement can be dated back to the 1980s when Kohlberg first presented his approach of moral development (Kohlberg, 1981). He proposed that moral judgements are grounded in highly elaborated reasoning processes, whereas current approaches propose that spontaneous moral judgements are grounded in affective and intuitive processes (Haidt, 2001). In addition, as research of social cognition demonstrated again and again the importance of emotions for a wide range of everyday evaluations and judgements (Forgas, 2012), it is likely that emotions are involved in moral judgement as well. Nevertheless, the exact role of emotions in moral judgement is still debated (cf., Waldmann et al., 2012; Avramova & Inbar, 2013). Moreover, it remains to be investigated whether moral evaluations and judgements are based on domain-specific mechanisms and potentially involve an innate morality module in the human brain, or whether moral evaluations use the same domain-unspecific mechanisms as other social judgements (e.g., reasoning and decision making).

In addressing these questions, current research on moral judgement nowadays combine

behavioural and neuroscientific approaches and methods to reveal the underlying mental processes. For instance, psychophysiological methods offer the opportunity to examine brain processes and other unconscious and automatically triggered bodily responses which are associated with (embodied) affective processing. In the following, I will describe various accounts that discuss the interplay of emotional and cognitive processes in moral cognition. The number of theories and approaches are illustrative for the complexity of the topic. Also, recent studies in other research domains provide convincing evidence for the assumption that emotional and cognitive processes do not work as independent and potentially antagonistic but rather as interacting systems (Lai, Hagoort, & Casasanto, 2012; Pastötter, Gleixner, Neuhauser, & Bäuml, 2013). To reveal the elementary mental mechanisms underlying moral judgements it is in my view crucial not only to examine the occurrence but also the precise time-course of affective and cognitive processes.

The primary aim of the present doctoral thesis is to reveal the covert affective and cognitive mental mechanisms and embodied responses involved in the processing of moral and emotional contents. Three studies will be reported using explicit and implicit tasks and behavioural as well as psychophysiological methods to address this topic. All three studies are inspired by the assumption that emotions play an important role in the processing of moral content during discourse comprehension. In order to elaborate the specificity of processing moral content, it is important to also investigate the processing of emotional content. To this end, a new set of moral and emotional text materials was constructed and pre-tested. These materials include short scenarios describing either moral or immoral actions and neutral or negative situations.

In the following section, first, I will discuss different theories of moral judgement and conceptual and methodological changes this research area underwent. Then I will discuss possible functions of emotions in moral cognition by considering recent studies and will end with highlighting the suitability of psychophysiological methods in addressing this issue.

1.1 Developmental theories of moral judgement

Initial theories of moral judgement originated in developmental and differential psychology. Researchers were interested in how and when children acquire moral beliefs, as well as in the reasons for why individuals' act inconsistently across different situations (Duska & Whelan, 1975). In these theories, moral development is described as the transformation of cognitive structures, which depends on the cognitive development in general and reflects a child's ability to adapt to its social environment. For decades, Jean Piaget was one of the

leading researchers of children's cognitive development, who considered moral cognition as a part of social cognition (Piaget, 1932). In general, he was interested in how the mind comes to respect rules in an attempt to regulate social interaction. According to Piaget, morality is a system of rules that also relies on perspective-taking abilities and that depends on social relationships that are characterised by mutual respect, cooperative activity, and developing autonomy.

Kohlberg's (1981) theory of moral development is grounded in Piaget's social developmental account and emphasises, in contrast to Piaget, that the development of moral cognition is a domain-specific aspect of cognitive development which already starts in infancy. According to his rationalist theory, moral judgements reflect the result of higher cognitive and elaborated reasoning processes of what is right and wrong. In a broad longitudinal study starting in 1963, Kohlberg (1981) interviewed 50 male pupils and young adults ranging in age from 10 to 28. He used moral dilemmas to examine moral judgement behaviour in order to uncover reasons for their decisions. Over 18 years, every participant had been invited every three years to assess their moral reasoning. Kohlberg found that all participants went through the same sequence of developmental stages. He differentiated six hierarchical stages at three distinct levels that include different moral dictums and principles varying with the processes of role- and perspective taking. It is worth noting that the rate of moral development differed between individuals of the same age, therefore, age could not be taken as sine qua non. Although moral development is positively correlated with cognitive development, individuals' moral orientation and the ability of moral reasoning differs. The *pre-conventional level* typically occurs in pre-adolescence when the labels good/bad and right/wrong and the underlying rules are adapted from the authorities who are also the source of punishment when breaking these rules. On this level, children have an egocentric perspective and good behaviour is defined as obeying the rules in order to avoid negative hedonistic and physical consequences for oneself rather than to consider the psychological interests of others. On the *conventional level*, the perspective changes to the self in relation to others of the group to which the individual belongs. This usually happens when children get in contact with their peers, for example in kindergarten or preschool. The individual realises its social affiliation and expects reciprocal commitment to rules. Good behaviour is no longer motivated for personal reasons but by how well it fulfills the order and expectations of the group. Finally, on the *post-conventional level* moral decisions are defined by rights, values, or principles that are agreeable to all individuals and have validity and application apart from an authority or the individual's own identification

with a group. At this level, a conflict between the own moral principles and these of the prevailing legal form and social society could arise when they differ (Kohlberg, 1981). The sixth stage is the highest level of development in moral reasoning and involves universal principles of justice, reciprocity, and the equality of human rights and the respect for the dignity of human beings (Duska & Whelan, 1975).

There are some limitations of Kohlberg's theory. First, the sixth stage is grounded in a hypothetical and abstract assumption based on philosophical and ethical rules such as the golden rule or Kant's categorical imperative rather than on empirical research (Duska & Whelan, 1975). Second, Kohlberg does not provide any explanations for the context-dependancy and variability of moral reasoning within an individual (Waldmann et al., 2012). For instance, whether lying is morally fine appears to depend on the situation. Third, his cognitive developmental account gives conscious and stage-dependent reasoning a central place in moral judgement. He neither believes in an innate moral entity nor in the influence of affective and intuitive processes. Finally, Kohlberg developed his theory and the assumption of six stages of moral development on empirical studies testing only western male students. It is therefore plausible to assume that his theory is subject to cultural and gender biases (Baumrind, 1986). Together, given these limitations, researchers questioned the assumptions of Kohlberg and put forward alternative theories of moral judgement.

1.2 Current moral theories of social psychology and cognitive neuroscience

While rationalist theories dominated research on moral judgement for a long time, the adoption of a social-psychological rather than a developmental approach and the use of neuroscientific methods provided new insights regarding the affective and cognitive processes underlying these judgements. Thus, social psychological accounts assume a central role for emotions and affective processing in moral cognition. According to Haidt (2001), automatic intuitions rather than moral reasoning underlie moral judgements. In his social intuitionist model (SIM), a moral judgement is understood as a social process and not as a private act of cognition (Haidt, 2001, 2007). Haidt conducted a number of experiments where he asked participants to morally judge stories describing taboo violations (e.g., describing sexual abnormalities, incest, or eating pets), for example, "A 25-year-old man likes to masturbate while his dog willingly licks his owner's genitals and seems to enjoy it.", and give reasons for their decisions (cf. Haidt & Hersh, 2001; Haidt, Koller, & Dias, 1993). Crucially, participants were able to quickly judge the behaviour as morally unacceptable,

but they were unable to provide justifications or reasons for their decisions. Haidt and colleagues call this phenomenon "moral dumbfounding", which is defined as the stubborn and puzzled maintenance of a moral judgement without supporting reasons. The speed and automaticity of judgements and the fact that justifications are very time consuming leads to the assumption that spontaneous affective evaluations are of key importance in moral judgements.

Hence, Haidt (2001) concludes that people base their judgements on valence-laden and holistic moral intuitions which are automatic and rapid processes. More specifically, an initial evaluative feeling of good/bad, like/dislike about a person's actions or character takes place and provides the grounds for the moral judgement. This is followed by a slow and controlled reasoning process, in support of the construction of justifications for intuition-based judgements and actions after a judgement has been made (Haidt, 2001). The major claim of the SIM then is that moral intuitions - and not reasoning - directly determine moral judgements. Later, Haidt and Joseph (2007) advanced the theory by addressing the question what the moral intuitions are like and where they originate from (see also Haidt & Joseph, 2004, for a review, see Graham et al., 2013). Haidt and Joseph identified five categories of human's moral intuitions or so-called moral foundations: harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity. More recently, researchers have added a sixth foundation, liberty/oppression, focusing on concerns about domination and coercion (Iyer, Koleva, Graham, Ditto, & Haidt, 2012). Crucially, it is emphasised that every foundation is linked to a specific emotional state, for example purity violations are associated with disgust feelings. In sum, the moral foundation theory is an approach to explain both the universality and the inter- and intra-individual variability of moral judgements across situations and cultures.

Moral psychology received another push around the year 2000 from research in cognitive neuroscience that aimed at identifying brain regions which are involved in moral judgements. For instance, Greene and his colleagues used functional magnetic resonance imaging (fMRI) to examine the brain processes underlying people's moral judgement behaviour with regard to moral dilemma situations that differ in their degree of personal immediacy (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Greene, Nystrom, Engell, Darley, & Cohen, 2004). Think of the classic example: you are to save five people working on a track from being run over by an out-of-control trolley. There are two possible scenarios. One, you could hit a switch to detour the trolley and save the five people, at the same time though kill another person on the other track. Two, you could push a man

off a bridge onto the track to stop the trolley, very well knowing that this action would kill him. These two scenarios offer illustrative examples of the extent to which moral dilemmas can engage us in diverging ways. Due to the physical contact, it is obvious that pushing a man off a footbridge is emotionally more salient than hitting a switch (Greene et al., 2001). Although, you would actually kill a person in both cases, your decision would presumably differ depending on the personal immediacy and the emotional engagement of the dilemma situation. That is why these two types of dilemmas are commonly used to reveal the underlying processes in moral dilemma judgements.

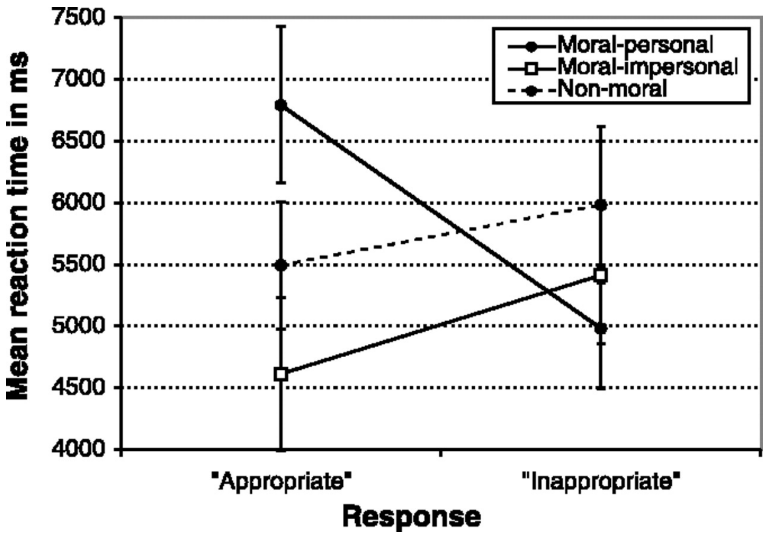


Figure 1. Mean reaction time as a function of condition and response type in Experiment 2 of Greene et al. (2001). Error bars indicate two standard errors of the mean. (Figure taken from Greene et al. (2001) with permission from Springer Link.)

Actually, this is what Greene et al. (2001) found. They asked participants to indicate whether they judged the action to kill a person in order to save others to be *"appropriate"* or *"inappropriate"*. First, in an impersonal dilemma scenario, participants were very likely to act and thereby kill a person (*"appropriate"*-judgement, utilitarian decision). In a personal scenario, however, participants were very likely not to act in order to avoid killing a person (*"inappropriate"*-judgement, deontological decision). Second, in personal dilemmas deontological decisions were faster then utilitarian decisions, whereas, there was no such difference in response time for impersonal dilemmas (see Figure 1). Crucially, fMRI results showed different activation patterns between the two types of dilemmas. For personal dilemmas, brain areas associated with emotion processing (i.e. medial frontal gyrus, posterior cingulate gyrus, angular gyrus) were more active, whereas for impersonal

dilemmas the results showed increased activation in brain areas that are related to working memory operations (i.e. right middle frontal gyrus, bilateral parietal lobe).

Based on the fMRI and behavioural results, the authors proposed a dual-process theory, which assumes both automatic emotional processes and controlled cognitive processes to contribute to moral judgements. According to this theory, there are two dissociable and antagonistic systems, with emotional processes being faster and finishing earlier than cognitive processing under some circumstances (Greene et al., 2001, 2004). The authors did not specify what exactly triggers emotional processing, but their results suggested that material-specific aspects matter. That is, personal dilemmas trigger emotional and impersonal dilemmas trigger cognitive processing. Still, the role of affective processing in moral judgement has not yet been satisfactorily supported (Huebner, Dwyer, & Hauser, 2009). In contrast, there is substantial evidence for controlled cognitive processes in utilitarian moral judgements. For instance, in a study of Greene, Morelli, Lowenberg, Nystrom, and Cohen (2008), in which cognitive load was manipulated, the proportion of deontological compared to utilitarian responses increased only for personal dilemmas under high load conditions. The same pattern was found, when using a task where participants had time pressure for their decisions, suggesting that controlled rational processing has less impact for fast responses (Suter & Hertwig, 2011).

A dual-process theory of moral judgement seems plausible based on Greene and colleagues' data (Greene et al., 2001, 2004, 2008). Nevertheless, this theory and the presumed impact of emotions on moral judgement has been criticised for theoretical and material-specific reasons (Bauman, McGraw, Bartels, & Warren, 2014; Schaich Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006; Conway & Gawronski, 2013; Kahane, Everett, Earp, Farias, & Savulescu, 2015; McGuire, Langdon, Coltheart, & Mackenzie, 2009). First, the theoretical foundation of the utilitarian and deontological principle suffers an important drawback (Conway & Gawronski, 2013). In Greene et al.'s study participants had to select one principle by rejecting the other one. This is only acceptable if they are inversely related. Conway and Gawronski (2013) argue that everyone has both a utilitarian and a deontological inclination which are independent dimensions and differ in their dominance, i.e., the influential strength could be balanced or unbalanced. If dominance differs, participants' responses could be attributed to an increase of one inclination, although it might reflect a decrease of the other one, or both are equally increased or decreased. The authors propose that variations in moral judgements are attributed to variations in the strength of a single moral inclination. Second, from the perspective of

language comprehension research, it is necessary to take a closer look at the moral dilemmas that have been used in Greene's studies. This will be discussed in more detail later (cf. Section 1.2.2).

In summary, Haidt's SIM and Greene's dual-process theory both assume an important role of emotions in moral judgements but the two theories differ with regard to the time point and the circumstances under which emotions become relevant. For Haidt, moral intuitions are grounded in emotions, form our moral judgements, and always precede cognitive processing, whereas for Greene emotions influence moral judgements mainly when the moral content is emotionally salient. At present, the role emotions play in moral judgement is still debated as will be briefly outlined in the following section.

1.2.1 The role of emotions in moral cognition

Evidence such as those provided by Haidt (2001) and Greene et al. (2001) supports the notion that emotions are involved in moral judgements. Their exact function, extent, and specificity is still in the focus of ongoing scientific research.

Regarding specificity, predominately negative emotions are repeatedly reported when participants observe a moral transgression. For instance, it is assumed that anger, contempt, or disgust are felt if someone else did something bad (Chen & Bargh, 1999) and self-referred emotions like shame and guilt are felt if I did something bad myself (Haidt et al., 2003). In general, it is emotionally arousing to violate social and moral rules (Prinz, 2006). Concerning the actual emotions triggered by Greene's moral dilemmas, the off-line rating study of Choe and Min (2011) yielded important results. They examined emotions evoked by Greene et al.'s (2001) personal dilemmas by asking participants to report the major emotion felt during their judgement. Guilt was reported most frequently (about 50%), followed by sadness, disgust, anger, empathy, and anxiety. This was supported by a rating study of Kunkel (unpublished) which found guilt (17.9%), anger (16.3%) and contempt (16.1%) as most frequently reported emotions for both personal and impersonal moral dilemmas. Disgust seems to be an important emotion in moral judgement (Rozin, 2008). Recent behavioural studies illustrated that disgust feelings induced by different methods (olfactory, visual, imaginative) increases the severity of moral judgements (Schnall, Haidt, Clore, & Jordan, 2008; Ugazio, Lamm, & Singer, 2012; Wheatley & Haidt, 2005).

Avramova and Inbar (2013) critically evaluate three claims regarding the function and the extent of emotions within moral information processing. The first function assumes an epiphenomenogical role of emotions that are only the consequence of a moral judgement.

Under this assumption, the controlled and cognitive evaluation of someone's behaviour is followed by an emotional response. Thus, the emotional response is a by-product and does not influence the moral judgement in any way. This is in contrast to the second claim that proposes a causal role of emotions (Greene et al., 2001; Haidt, 2001; Pastötter et al., 2013; Schnall et al., 2008; Ugazio et al., 2012). Accordingly, an affective response is triggered while moral information is processed, hence preceding and determining the moral judgement. The third claim discusses the moderating role of emotions on moral information processing. An individual's moral judgement varies depending on the current emotional state, even if emotions are of the same valence (Janoff-Bulman, Sheikh, & Hepp, 2009; Polman & Ruttan, 2012; Schmader & Lickel, 2006). The idea is that different emotional states evoke opposing motivational tendencies which influnece moral judgements, for example feelings of anger are taken to motivate approach tendencies and to decrease the severity of moral judgements, whereas feelings of disgust are assumed to trigger withdrawal tendencies and to increase the severity (Ugazio et al., 2012).

There are different accounts that consider emotional and cognitive processing in general. The affective primacy assumption proposes a dominance of emotional over cognitive processes (Zajonc, 1984). Emotional stimuli have a facilitated and prioritised access, avoiding extensive and effortful cognitive processing. This is based on the assumption that emotions of different valence differentially relate to the brains' motivational appetitive/approach and aversive/avoidance systems (Bradley & Lang, 2000). In an evolutionary context, the affective motivational system evolved to ensure the survival of organisms by quickly identifying negative or dangerous situations that threaten their lives. Even today, negative stimuli are known to capture our attention (Hajcak, MacNamara, Foti, Ferri, & Keil, 2013; for a review see Hajcak, MacNamara, & Olvet, 2010) and to be processed faster (Ito, Larsen, Smith, & Cacioppo, 1998).

The more recent affect-as-information hypothesis of Clore and colleagues (2001; Clore & Huntsinger, 2007, 2009; see also Clore et al., 2001; Pastötter et al., 2013; Schnall et al., 2008; Schwarz & Clore, 1983) proposes that an emotional response, which is experienced as a reaction to what is being judged, serves as information during the decision process and shapes our judgement. More precisely, the momentary affective state directs one's attention to what is a plausible cause for our current feelings and is used as information to evaluate an object. Moreover, it is assumed that emotions of different valence are linked to particular cognitive styles; a positive affect tends to reinforce relational processing like cognitive, interpretive, category-level and global processing, whereas negative emotions

inhibit this tendency and lead to referential processing like perceptual, item-level, and local processing (Clore & Huntsinger, 2007, 2009). This hypothesis is in line with the causal and moderating role discussed by Avramova and Inbar (2013).

However, other approaches assume that the primacy of affect does not hold universally (e.g., Lai et al., 2012). According to Lai and colleagues, the primacy of either affective or cognitive processing of the same stimulus varies depending on the global context in which they occur and neither on the stimulus per se nor on the task participants perform. The primacy of affect is only given if a stimulus is presented within an affective context.

1.2.2 Criticism of Greene's moral dilemmas

From the perspective of language comprehension, it is necessary not to forget the limitations of the moral dilemmas that have frequently been used in other studies, but have never been validated or tested in regard to their material specific effects. McGuire and colleagues (2009) reanalysed Greene's data by repeating their analysis after removing nine items that have a very low percentage of subjects judging the behaviour to be appropriate. They performed two additional analyses where items instead of subjects were taken as random factor, first, with all and, second, without these nine"bad" items. An analysis with items as random factor reveals whether an effect is based on the experimental manipulation or rather driven by specific items and hence item specific aspects. McGuire et al.'s results show that the important interaction between dilemma type (personal versus impersonal) and response (appropriate versus inappropriate), which was found in the subject analysis with all items, was absent in the same analysis without the"bad" items. Furthermore, all effects of the by-subject analysis were absent in both item analyses. Based on these results, it appears that response time findings of Greene et al. (2001) are driven by just a few particular dilemmas with extremely fast *inappropriate* responses to personal dilemmas (McGuire et al., 2009).

Schaich Borg et al. (2006) criticised the language used in personal versus impersonal dilemmas. Personal dilemmas are described using a more emotive and colourful language and often reference to family members or friends, both as confounding factors explaining greater emotional activation. Also Bauman et al. (2014) doubt the generalisation of results, which emerge from sacrificial moral dilemmas, since they are unrealistic and unrepresentative. The authors further criticise the poor plausibility of the dilemmas which are more amusing than shocking. Hence, they do not elicit the same psychological processes as other moral situations would do. Moreover, Kahane et al. (2015) assumed that ordinary

responses to sacrificial dilemmas tell us little about moral judgement behaviour because of the assumption that dilemmas are too complex, far-fetched, and convoluted.

Greene accepted the criticism regarding his text stimuli (Greene, 2009) which were frequently used by researchers in this field. Current studies use standardised, revised, and newly constructed moral dilemmas with lists of factors that are controlled for including intentionality, involvement of family members, or the number of harmed people among others (cf., Christensen, Flexas, Calabrese, Gut, & Gomila, 2014; Kahane et al., 2012; Lotto, Manfrinati, & Sarlo, 2014).

It is obvious that moral research typically involves quite complex text materials. Nevertheless, from the perspective of language processing research, the limitations of text materials can ideally be overcome if word- or sentence-based effects are minimised. Materials need to be very well constructed and controlled by some important linguistic aspects including an extensive description of target words (i.e., word class, word length, number of syllables). Moreover, rating studies should be used to reveal the characteristics of materials with regard to cloze probability and plausibility, and also the semantic relatedness of target words in relation to a given context or sentence should be analysed. It is also important to have valid morality as well as valence and arousal ratings for each item.

2 Emotional language comprehension

As outlined earlier, moral research typically involves the presentation of complex text stimuli including an agent, a social situation, and a potential moral conflict. To get an idea about the implicit mechanisms during discourse comprehension of moral text stimuli, it appears promising to use a language comprehension approach. A typical research question would be, how and when exactly individuals have processed the meaning of (im-)moral content during the reading process. There are many theories about how language is processed and understood. It is not the intention of the present thesis to go into detail of the various language comprehension models (e.g., Van Dijk, Kintsch, & Van Dijk, 1983; Gernsbacher, 2013). In the following section, I will give a brief overview of underlying theories in order to understand the assumptions of the present studies.

According to the construction-integration model of Kintsch (1988) a text is interpreted at three distinct levels of mental representations. The least abstract representation is a mere phrase structure that captures the exact wording of the original text and is the so-called surface model. For the next level of abstraction, the surface model is converted into a mental representation that describes the text in propositions. A proposition is a predicate

with its arguments, forming the basic unit of meaning that can be assigned a truth value and which is linked to other propositions (Kintsch, 1974). For example, the meaning of the sentence "The cat sits under the table." would be represented by the proposition with the following form: sit [cat, table, under]. Propositional representations form the so-called text-base which also includes information that is not explicitly mentioned in the text but inferred from it. At the last level, propositions will be integrated into a situation model that describes the ideas that the text is about. It is generally assumed that comprehenders encode various text dimensions such as time, space, causation, as well as the intentions and emotions of story characters in these situation models (e.g., Kintsch, 1988; Zwaan & Radvansky, 1998). Situation models provide a coherent mental representation of the people, objects, and events being described in the text. For example, emotional states of the described characters as well as anticipated future events or consequences are inferred from the situation model while reading or listening.

That the meaning of conversations and texts is given by propositional structures that could be combined to form higher-order complexes is a very abstract and disembodied view of language comprehension (Gibbs, 2003). In contrast to the symbolic cognitive model described above, recently, simulation models of language comprehension suggest that linguistic meaning is related to real-world referents (for a review, see Barsalou, 2010). This view emphasises the importance of perception and action in people's understanding of linguistic meaning. Comprehension involves reactivation of experienced-based, mental representations which are grounded in perception and action. In contrast to the idea that text comprehension involves the representation of amodal propositions, simulation models assume a key role of different modal, experience-based representations (e.g., motor, sensory). Therefore, reading a word is not only accompanied by an activation of lexical, grammatical, or phonological representations but also by emotional and perceptual represenations which are based on previous experiences (Zwaan, 2004).

Language provides a set of cues, which the reader integrates and uses to construct a coherent mental representation and simulation of the described situation. It is assumed that in discourse processing, different types of constraints, such as phonology, semantics, and syntax are very quickly taken into consideration during listening and reading (Hagoort, 2003; Hagoort, Brown, & Osterhout, 1999). The activation of previous experiences helps the reader to quickly comprehend linguistic input, for instance, when the processing of incoming emotional information activates a previously experienced emotional feeling.

In the following section, I will describe studies using various psychophysiological mea-

sures that investigated brain and embodied processing of emotional linguistic input. Subsequently, I will describe a new set of text materials which is used for the present doctoral thesis to overcome the limitations with previously used text materials (cf. Section 1.2.2).

2.1 Emotional language comprehension at the level of brain processes

fMRI is ideally suited to identify the brain regions involved in emotional comprehension (e.g., Greene et al., 2001) because this method provides high spatial resolution. Due to its relatively low temporal resolution, however, fMRI is not appropriate for the purpose of identifying the millisecond by millisecond timing of mental processes. For instance, Greene et al. (2001) assumed that automatic emotional processes are involved in moral judgements. Based on their fMRI data, their precise timing, that is whether these automatic processes precede or follow the moral decision, could not be clarified. It turns out that other methods are needed to measure the nature and timing of mental processes triggered by incoming moral information.

In contrast to fMRI, the electroencephalogram (EEG) is a direct and non-invasive measure of neural activity. The EEG has a high temporal resolution, although at the expanse of a limited spatial resolution (Bartholow & Amodio, 2009; Amodio, Bartholow, & Ito, 2013; Luck, 2014). It represents the summation of post-synaptic potentials from populations of synchronously active neurons, that is, mainly pyramidal cells of the cortex. These summated potentials form an electrical field which is strong enough to be measured with electrodes placed on the surface of the scalp. The event-related brain potential (ERP) reflects the neural activity associated with the processing of a specific event like the presentation of a stimulus, a mental event, or a response. The ERP waveform is a voltage by time function and is characterised by a series of peaks and troughs. These individual deflections differ with regard to their polarity, latency, and distribution at the scalp. Crucially, individual deflections are modulated by experimental manipulations which allows to distinguish them with regard to this functional significance of the ERP. That is, individual ERP deflections can be related to theoretical components which are characterised in different ways. There are three major classes of ERP components: exogenous, endogenous, and motor components (Luck, 2014). Motor ERP components accompany the preparation and execution of a motor response. Exogenous components are triggered by the presentation of a stimulus and reflect automatic and obligatory sensory processes which can be modulated by top-down processes (e.g., selective attention). Endogenous components reflect cognitive processes and mechanisms which are task-dependent. Usually, they have a longer latency

than exogenous components, occurring typically 200-300 ms after stimulus onset.

A number of ERP studies investigated brain processes underlying emotional language comprehension (for a reviews see Fischler & Bradley, 2006; Hajcak, Weinberg, MacNamara, & Foti, 2012). In the following, I briefly summarise relevant ERP findings in this research area. The N400 is a language-related component which was firstly described by Kutas and Hillyard (1980). The N400 amplitude has been shown to respond to the predictability of a word within a given constraining context. Kutas and Hillyard (1980) presented sentences that were either correct (e.g., "I take coffee with cream and *sugar*.") or violated a semantic expectancy (e.g., "I take coffee with cream and *socks*."). They used the rapid serial visual presentation (RSVP) of individual words to present the critical sentences. That is, by presenting the sentence word by word, participants were controlled in their reading time. Importantly, the RSVP procedure allowed to time-lock the ERP to the critical word (*sugar/socks*), thereby capturing the brain's on-line responses to the semantic match versus violation. The ERP results showed that after presenting the critical word, there was a negative deflection over centro-parietal regions peaking after 400 ms. The N400 was larger for the word *socks* than *sugar*, which is the semantically inappropriate word in the given context.

This N400 effect reflects the anomaly in semantic processing at the sentence level, for example the eliciting word mismatches the meaning of the sentence context (Kutas & Hillyard, 1980; Van Berkum, Holleman, Nieuwland, Otten, & Murre, 2009). Moreover, later studies found that world knowledge violations also produce a larger N400 than information consistent with world knowledge (e.g., Filik & Leuthold, 2008; Hagoort, Hald, Bastiaansen, & Petersson, 2004; Leuthold, Kunkel, Mackenzie, & Filik, 2015). Based on such findings, it is generally assumed that the N400 reflects the processing cost of meaning construction or the reprocessing of anomalous information (Brown & Hagoort, 1993; Osterhout & Holcomb, 1992; for a review see Kutas & Federmeier, 2011). The worse the semantic fit between a word and its context, the more enhanced is the amplitude of the N400 (Osterhout, 1997).

Other studies have shown that the N400 is not only found for semantic anomalies but also when discourse-level expectancies are violated (Van Berkum, Hagoort, & Brown, 1999). As concerns emotional discourse comprehension, León, Díaz, de Vega, and Hernández (2010) presented emotional stories describing a protagonist in an emotional situation (e.g., a young writer, who was going to the premiere of his first theatre play, and the next day, critics write that he might be a new theatre talent), followed by a target sentence

("Today Hector felt totally *fulfilled/failure*."). The emotion implied by the target sentence was either consistent or inconsistent with the preceding context. The results showed, that inconsistent target sentences elicited a larger N400. Leuthold, Filik, Murphy, and Mackenzie (2012) used scenarios where a context describes an agent in a social situation (e.g., doctor-patient interaction) that was followed by a target sentence ending with the agent's socio-emotional response. The described response either matched or mismatched what one might typically expect from the given situation. In contrast to León et al. (2010), the critical word was identical and only the context varied between the conditions. Hence, resulting ERP effects could be attributed to discourse-based processing, ruling out an explanation in terms of word-based emotional effects. The results indicated that when the critical word mismatched the participant's expected feelings, again, a larger N400 was elicited. Therefore, according to Leuthold et al. (2015) the emotional salience of an identical word can modulate the process of meaning construction depending on the discourse context.

Another ERP component is the P300, which is a positive-going, parietally distributed ERP component peaking at about 300 ms or later after stimulus presentation. It is elicited by unpredictable, infrequent changes in stimuli, for example, when presenting tones that differ with regard to pitch and loudness as compared to a series of standard tones (Squires, Squires, & Hillyard, 1975). Importantly, for present purposes, Cacioppo, Crites, Berntson, and Coles (1993) described a P300-like positivity with a longer latency than typically reported in P300 research, which they named late positive potential (LPP). Cacioppo and colleagues took this component to indicate a categorisation process which was associated with an evaluatively inconsistent stimulus in a sequence of stimuli of opposing valence (e.g., a negative stimulus embedded in positive stimuli). This LPP effect was also found for emotional pictorial stimuli with a larger amplitude for negative affect (Cacioppo et al., 1993; Crites, Cacioppo, Gardner, & Berntson, 1995; Crites et al., 1995).

Also, the LPP has been repeatedly reported in response to emotional words, sentences, and texts (for a review see Citron, 2012). The LPP amplitude has been found to be sensitively modulated by various variables, including arousal which reflects the strength of emotional content (extremity, intensity), valence (Bradley & Lang, 2007; Holt, Lynn, & Kuperberg, 2009; Ito et al., 1998), discourse context (Delaney-Busch & Kuperberg, 2013; Fields & Kuperberg, 2016) and task demands (Gable, Adams, & Proudfit, 2015). For instance, it was found in explicit and implicit categorisation tasks (Ito & Cacioppo, 2000) as well as in a pure reading task when no affective judgement was demanded (Leuthold

et al., 2015). The LPP has been also found to be sensitive to evaluative inconsistencies (Herring, Taylor, White, & Crites Jr, 2011). For instance, Holt et al. (2009) presented two-sentence scenarios, where the critical word was either negative, positive, or neutral, within a nonconstraining context (e.g., "Nancy's son ended up just like his father. He was already a *husband/millionaire/criminal* by age 25."). The results showed no N400 effects, indicating the absence of semantic processing difficulties, but instead an LPP amplitude which was positively correlated with items' arousal and valence ratings. Hence, the authors assume that the LPP reflects physiological arousal induced during affective language processing of emotional content.

Extending Holt et al. (2009), Delaney-Busch and Kuperberg (2013) compared the effects of valence congruity of emotional words with the effects of semantic congruity of neutral words to a given constraining context. They presented two-sentence scenarios like "Colin saw a *stunning/horrifying* object on the ground. He realised it was a *snake/diamond* right away.", for the emotion condition, and "Colin saw a *small* object on the ground. He realised it was a *button/giraffe* right away.", for the semantic condition. Like in other studies, emotion-congruent compared to incongruent conditions elicited an N400 if a neutral rather than an emotional context preceded the target sentence. A novel finding in this study was that for emotional words the N400 was followed by an LPP which was not modulated by discourse congruity. According to the authors, the LPP indicated the emotional salience (to arousal and valence) of incoming emotional words while reading the scenarios. Overall, the LPP indicates an evaluation of incoming stimuli, which apparently depends on affective aspects.

In sum, ERPs measure the temporal dynamics of mental activity, therefore, it is a straightforward method to infer the precise time course of hidden mental mechanisms underlying the processing of incoming information. Most of the ERP studies of emotional language comprehension employed materials for which an emotional critical word of a target sentence differed across emotional discourse contexts (e.g., Delaney-Busch & Kuperberg, 2013; Holt et al., 2009; León et al., 2010). Hence, the N400 varies with the degree of discourse-level expectations and is taken to indicate cognitive processing of incoming linguistic information. In contrast, the LPP varies with arousal and valence ratings of either the whole text stimulus or single target words and is therefore taken to indicate affective processing. With regard to the objectives of the present doctoral thesis, the question then is, whether incoming moral information elicits the same general processes in participants' brains as emotional information does or whether there is a domain-specific

activation pattern.

2.2 Emotional language comprehension at the level of peripheral measures

Damasio (1994) stated that repeated experiences in the world can trigger emotional responses of the brain that involve bodily changes and feelings which occur whenever a similar situation arises. Imagine, you are walking along the sidewalk and step into a pile of dog poop. You try to clean your shoe with a tissue. You see the colour, the consistence, you smell it, and you might feel disgusted. The different modality-specific representations of this experience will be stored in the brain. The next time you walk along a sidewalk you probably prioritised your attention to brown spots on the ground. If you see dog poop, you will quickly access the stored multimodal representation which in turn will increase processing speed in the current situation. The interpretation would be "That's disgusting", and your leg muscles will be innervated to immediately step aside. Moreover, only thinking about excrements may reactivate the same representation of disgust. So, the associated knowledge will be activated (brain) and simulated (body), even if the particular entity is not present, for example thinking of disgusting things involves the re-experiencing of neural states that occurred when one felt that emotion before.

Peripheral psychophysiological methods like measuring heart rate, skin conductance responses and muscle activity are straightforward measures to identify these embodied effects. Phasic electrodermal activity (EDA) provides an index of sympathetic arousal of the underlying sudomotor nerve activity (Dawson, Schell, & Filion, 2007). Changes in heart rate are an index of changes in the balance between sympathetic and parasympathetic influences on the emotional state, indicating individual differences in regulated emotional responding (Appelhans & Luecken, 2006). Facial EMG reveals involuntary facial muscle movements, which are related to the expression and hence simultion of specific emotions (Niedenthal, 2007; Niedenthal, Winkielman, Mondillon, & Vermeulen, 2009). More specifically, fEMG activity from the levator labii superioris muscle (wrinkling) and corrugator supercilii muscle (frowning) is taken to indicate negative affect (disgust, anger), whereas zygomaticus major (smiling) activity indicates positive affect (e.g., Larsen, Norris, & Cacioppo, 2003; Tan et al., 2012).

Embodied simulation accounts of language processing also postulate an important role of embodied processing in the understanding of words, phrases, and texts (Barsalou, 1999; Glenberg & Robertson, 1999; Havas, Glenberg, & Rinck, 2007; Havas, Glenberg, Gutowski,

Lucarelli, & Davidson, 2010; Niedenthal et al., 2009). Studies concerned with language comprehension show that emotional information processing involves the automatic and unconscious reactivation and simulation of already experienced, and implied emotions which are relevant to the particular present emotional meaning (Niedenthal, 2007). The knowledge of a specific emotion is assumed to be grounded in ordinary past experiences involving different modalities (sensory, motori, affective) (Niedenthal et al., 2009). Past experiences influence attention, processing speed, and the interpretation of incoming emotional meaning. This assumption fits very well with research in social psychology on the affect-as-information hypothesis of Clore et al. (2001). More precisely, the activated affective state directs one's attention, shapes perceptions, and serves as information in language comprehension.

Nevertheless, the exact role of embodied simulation during language comprehension is still unclear (Barsalou, 1999; Filik, Hunter, & Leuthold, 2015; Glenberg & Robertson, 1999; Gibbs, 2003; Havas et al., 2007; Niedenthal et al., 2009 Glenberg, Webster, Mouilso, Havas, & Lindeman, 2009 for a review). There are a number of studies that measured embodiment in the processing of emotional words and sentences. For instance, embodied effects were showen for isolated emotional words (Niedenthal et al., 2009). Niedenthal and colleagues presented negative (e.g., murder, vomit) and positive (e.g., smile, sun) target words eliciting anger, disgust and joy. Participants either indicated whether the word was written in capital or small letters or whether it was associated with an emotion. Only in the emotion-focused task, embodied responses of facial muscles were observed. Moreover, if the ability of activating corresponding muscles was inhibited, accuracy in judgements suffered. However, Niedenthal and colleagues (2009) suggested that embodied simulation of emotions is involved in the processing of emotional laden words but only if the task demands their affective evaluation.

With regard to discourse comprehension, behavioural studies demonstrated embodied effects for sentences with emotional content (Havas et al., 2007, 2010). Havas et al. (2007) presented sentences describing emotionally laden events (e.g., "The police car rapidly pulls up behind you, siren blaring.") or target words (e.g., embrace, exam) and measured response time in a lexical decision task. They manipulated emotional states with a procedure developed by Strack, Martin, and Stepper (1988) in which participants hold a pen in the mouth to produce either a smile (holding the pen using only the teeth) or a frown or pout (holding the pen using only the lips and not the teeth). It was only for sentences only that emotion simulation influenced the response time of participants. The authors

assumed that emotion simulation during language comprehension appears only at the level of phrases or sentences and not as a general lexical-level phenomenon. By using fEMG, Havas et al. (2010) demonstrated that involuntary facial expressions facilitate the processing of emotional language. They presented sentences with angry, sad, and happy content (e.g., "You hold back your tears as you enter the funeral home." for a sad sentence) in a reading-for-comprehension task and measured participants reading time. They inhibited activation of the corrugator by injecting botulinum toxin-A (a neurotoxin that causes temporary muscular denervation). Results showed that the inhibition of muscle activity necessary for expressing the negative emotion evoked by the sentences slowed reading time. This provides evidence for a causal role of embodied emotions in the processing of emotional language.

Most of the studies examined embodied processing by using affective-laden target words. To my knowledge, only one study so far investigated embodied processing in discourse comprehension. The study of Thompson, Mackenzie, Leuthold, and Filik (2016) investigated the influence of context on emotion simulation when participants read scenarios with ironic content. They presented short scenarios like "Susie texted Linda to say that she hadn't been to the gym at all that week. Linda texted her back to say: You're so motivated." while recording fEMG and EDA. They found evidence for reduced frowning (corrugator activity) and enhanced smiling (zygomaticus activity) for ironic compared to literal criticism, but enhanced frowning and reduced smiling for ironic compared to literal praise, suggesting that irony weakens the emotional impact of language input.

Together, evidence supports the assumption that emotion simulation facilitates emotional language comprehension. Psychophysiological methods are also suited to test assumption whether emotional information processing involves embodied representations of language comprehension. There is less evidence whether emotion simulation takes place in discourse comprehension of emotional scenarios and whether this is dependent on the task like for words (Niedenthal et al., 2009).

2.3 Moral language comprehension

In the last decade, morally loaded language processing has been taken as a special case of emotional language processing (Van Berkum et al., 2009). In the following section, I will illustrate how ERPs and peripheral psychophysiological methods were used to investigate moral language processing and which questions have not yet been answered.

2.3.1 Event-related brain potentials

So far, only a few ERP studies investigated the neural mechanisms of moral cognition in an attempt to establish the time course of moral information processing. Chen, Qiu, Li, and Zhang (2009), for instance, were to my knowledge the first to investigate neural correlates associated with processing of moral dilemmas. Approximately one month after the Sichuan earthquake in 2008 they invited volunteers who have experienced it. They presented two kinds of word pairs consisting either of relatives (e.g., *father-mother*, *sister-brother*), or of strangers (*stranger A-stranger B*). Participants were asked to choose as quickly as possible between one of the two persons which to rescue from the earthquake. After half of the blocks participants were informed by the experimenter that there would be an aftershock after a few hours. When choosing between relatives rather than between strangers, the results showed a larger P2 amplitude for relatives compared to strangers. According to the authors, this indicates conflict detection during early stimulus evaluation (Chen et al., 2009). Furthermore, only after hearing the aftershock warning, a larger positivity between 350 and 450 ms and a longer reaction time was found in connection to the choice between relatives as opposed to strangers. The authors related this late positivity to the P300 component and suggested that when making decisions, for example, between mother and father a stronger dilemma conflict affects the resolution processes indicated by the larger P300. In my view, *stronger emotional arousal* would be an alternative interpretation of Chen et al.'s P300 results because the participant's relation to the victim described in the dilemma influences the emotional impact of the situation (Schaich Borg et al., 2006; Christensen et al., 2014).

Similarly, Sarlo et al. (2012) used moral dilemmas adapted from Greene et al. (2001) and Cushman, Young, and Hauser (2006) and measured the temporal dynamics of emotional and cognitive processing in participants' decisions. The moral dilemmas were categorised into two types of dilemmas depending on the agent's intention to harm another person. The first category included dilemmas for which the death of one or more persons was instrumentally used to save the lives of many others. For the second category, the death of one or more persons was a foreseen but unintended, incidental side effect of the agent's action. After presenting the dilemma scenarios, two action options were presented successively on different screens: one in which the main character let the people die (non-utilitarian); and another one in which the main character kills one individual to save these people (utilitarian). On the following screen, participants were asked to choose a utilitarian or non-utilitarian option while decision time was limited. ERPs were time-locked

to the decision screen. ERP results showed, first, that instrumental dilemmas elicited a larger positivity over anterior electrodes (P260) than incidental dilemmas, whereas a later posterior positivity slow wave between 600 and 750 ms was larger for incidental rather than for instrumental dilemmas. Sarlo et al. (2012) suggested that the late positive slow wave reflects the allocation of attentional resources and working memory load. Second, self reported emotional experience indicated that incidental dilemmas were perceived as less unpleasant than instrumental dilemmas. According to the authors, this suggests that the decision for incidental dilemmas seems to require more cognitive effort for a cost-benefit computation and appears to be unrelated to emotional activation, because there is no correlation between slow wave amplitudes and affective ratings. In contrast, the correlation analysis showed that cognitive processing of instrumental dilemmas is modulated by emotional intensity, indicated by a larger P260 amplitude. It is important to note, that arousal ratings did not differ between the two types of dilemmas, but correlated positively with the LPP amplitude across all dilemmas.

The studies of Chen et al. (2009) and Sarlo et al. (2012) do not give satisfiable evidence for affective processes involved in moral judgement. In my view, this resulted from an inappropriate interpretation of the late positivity observed in both studies. There is an alternative interpretation for the late positivity in terms of arousal differences of materials that are reflected in ERPs. Unfortunately, the authors do not report the emotional characteristics of materials (cf. arousal and valence rating)

A key ERP study for the present thesis is that of Van Berkum and colleagues (2009) because they used a language comprehension approach to examine moral information processing. Especially, they examined whether and when individual values influence the linguistic analysis of meaning. Participants were selected from members of two political parties with opposing value systems (Christians vs. non-Christians). They were asked to rate their agreement to critical statements like "I think euthanasia is an *unacceptable/acceptable* course of action." with the critical word *acceptable* being consistent to non-Christians and inconsistent to Christians, and vice versa for the critical word *unacceptable*. Van Berkum et al. (2009) used the RSVP paradigm described earlier. They found that value-inconsistent words initially elicited a larger, broadly distributed positivity between 200 and 250 ms (P200), followed by a standard centroparietal negativity peaking at 400 ms (N400), and finally a broadly centroparietally distributed late positivity (LPP) between 500 and 650 ms, peaking approximately 600 ms after stimulus onset.

Van Berkum and colleagues interpreted their P200 findings to indicate the affective

salience of the value-based critical word which leads to its enhanced semantic analysis. Moreover, they took the N400 findings to indicate that readers immediately and automatically evaluate incoming information with respect to their personally held values, giving rise to rapid, valence-based influences on meaning construction. They further speculated that this N400 effect cancels a single sustained LPP effect with an earlier onset than the N400, therefore emerging as a larger P200 and LPP for value-inconsistent than value-consistent statements, an ERP signature they take to reflect the activation of the affect system.

Van Berkum et al.'s study has two major limitations. First, the statements were only usable in order to test members of two political parties of opposing value systems rather than general held moral values. Second, they used explicit moral topics (e.g., euthanasia, abortion, divorce), which are highly dependent on personally held moral values. This constrains participants' (implicit) expectations regarding the likely sentence endings. The critical word in value-inconsistent conditions is less expected and, hence elicits a larger N400 than the value-consistent word. Therefore, Van Berkum et al.'s interpretation of the N400 effect is questionable, because it may reflect a consistency effect regarding highly expected critical words rather than the immediate and automatic evaluation of incoming information with respect to the personally held values.

In conclusion, despite the potential limitations of previous ERP studies, the recording and analysis of ERPs is perfectly suited to produce insights in the time course and mechanisms underlying moral judgements. That is, ERPs give access to covert processes at different levels, which manifest themselves in distinct ERP components reflecting, for example, the attentive processing of incoming information (P200), the construction of meaning (N400), and the implicit affective evaluation of input (LPP).

2.3.2 Embodied cognition

With regard to the potential impact of affective simulations during the processing of moral information, less is known about embodied responses in moral cognition. Recent studies mostly induced specific emotions rather than measuring embodied responses triggered by moral information and they used only behavioural methods. Hence, these studies provide only indirect evidence and do not allow making assumptions about the precise time course of embodied responses (e.g., Schnall et al., 2008; Ugazio et al., 2012; Wheatley & Haidt, 2005).

Cannon, Schnall, and White (2011) were the first to investigate embodiment by using

peripheral psychophysiological measures in response to short audible statements describing either moral or immoral actions. The moral content of statements varied as defined by the moral foundation theory of Haidt and Joseph (2007). Participants were asked to listen to the statements and make a moral judgement ("How negative or positive was this behaviour?"), while the fEMG of three muscles was recorded. fEMG results varied depending on the moral foundation, which let the authors conclude that the muscle activation depends on the specific emotional response elicited by a given statement. Stronger levator activation (nose wrinkling) was found with statements describing purity and fairness transgressions indicating disgust, whereas harm statements elicited stronger corrugator activity (frowning) indicating anger. For authority and ingroup transgressions, the fEMG activity was not modulated by valence. Together, the authors related both increased levator and corrugator muscle activity to the processing of moral transgressions. Their data suggest that although spontaneous facial affect predicted the extremity of moral judgements, there was no unique pattern of fEMG activity in relation to specific moral foundations (Cannon et al., 2011).

Recently, Krumhuber, Tsankova, and Kappas (2018) presented socio-cultural norm vignettes describing, for example, hygiene, gender equality, or personal space violations. The vignettes described the action of an agent in situations where he acted against or outside of a social/cultural norm. Krumhuber and colleagues measured fEMG at two time points: when participants read the vignettes and when participants imagine that they were the recipients of the agent's action. Only the levator muscle was sensitive to violations compared to neutral vignettes during both reading and imaging. The authors assumed that their results are in line with the assumption that the feeling of disgust is highly correlated with moral violations. Therefore, they suggested that socio-cultural violations involve a moral component.

Despite this evidence for embodied effects elicited by moral or socio-cultural content, it remains to be investigated whether written moral information is immediately affectively evaluated during reading. Moreover, it is still unclear whether moral information processing requires domain-specific mechanisms or whether it is processed in the same way as emotional content.

2.4 Interim summary

Theories of moral cognition suggest an important role of emotions in moral judgement. Researchers aim to disclose the occurrence and precise time course of cognitive and affective

processes. So far, however, there is no convincing evidence that emotions matter during moral information processing. Reasons for this are that previous research used unrealistic dilemma situations, differing emotional target words, or constraining contexts that describe explicit emotional states. Crucially, it is still unclear whether everyday moral text materials trigger emotional processes and whether moral judgements are simply a special case of domain-general cognitive processes.

A language comprehension approach appears useful to shed light on some of these open research questions, especially when combined with an on-line measurement of the processing of moral and emotional content. To this end, psychophysiological and measures with high temporal resolution, such as ERPs and fEMG promise to allow inferences about the precise time course of unconscious mechanisms underlying moral cognition. With regard to ERPs, there are two components that have been related to different mental processes during comprehension: the N400 indicating cognitive-semantic processing and the LPP indicating affective processing. With regard to peripheral psychophysiological measures, there are only a few studies on affective simulation during on-line moral and emotional language processing.

3 The processing of moral transgressions: Investigating the role of affective evaluations using everyday scenarios

It is the goal of the present doctoral thesis to investigate the time course of affective processing in discourse comprehension using newly constructed morality and emotion materials and different psychophysiological methods. In the following section, I will describe first the text materials, the methods, and finally the exact research questions.

3.1 New moral stimuli

The text materials of previous studies concerned with moral processing (e.g., moral dilemmas of Greene et al., 2001) are not suitable when adopting a language comprehension approach, as used in the present work. This is because dilemma materials were not constructed in such a way as to control for important linguistic aspects in order to minimise material-specific effects. Often, the materials lacked for an extensive description of their linguistic characteristics (i.e., word class, word length, cloze probability/semantic relatedness of target words), and other descriptive ratings (plausibility, valence, arousal). To overcome these limitations, first, materials need to be designed well and control for some

Table 1

Example socio-normative and general world knowledge materials (critical word in bold). (adapted from Leuthold et al., 2015)

Socio-Normative	
Morally-acceptable Context	Tinas Opa leidet an Krebs und wird bald sterben. Zu seinem 85. Geburtstag hat er ein großes Fest geplant und wünscht sich vor allem, dass alle seine Kinder und Enkel kommen. [Tina's grandfather suffers from cancer and will die soon. For his 85th Birthday he has planned a big party and wishes nothing more than that all his children and grandchildren attend.]
Morally-unacceptable Context	Tinas Chef macht schon seit einer Weile eindeutige Anspielungen. Nun hat er sie zu einem Essen in ein teures Restaurant eingeladen. Sie weiß, dass er seit 20 Jahren verheiratet ist und drei Kinder hat. [Tina's boss has been making explicit innuendos for some time. Now he has invited her out for dinner to an expensive restaurant. She knows that he has been married for 20 years and is the father of three children.]
Target sentence[a]	Sie hat die Einladung *angenommen*. [She has the invitation *accepted*.]
World Knowledge	
Knowledge-consistent Context	Bei einem Frankreichaustausch isst Frau Lehmann eine bekannte französische Delikatesse. [During a France exchange Mrs. Lehmann eats a famous French speciality.]
Knowledge-inconsistent Context	Frau Lehmann geht in ein schwäbisches Restaurant und bestellt eine lokale Spezialität. [Mrs. Lehmann goes to a Schwabian Restaurant and orders a local speciality.]
Target sentence[a]	Sie erhält als Gericht einen Teller voller *Schnecken* und Weißbrot. [She receives as dish a plate full of *snails* and white bread.]

[a] Note that target sentences are translated word by word to indicate the position of the critical word, hence disregarding the appropriate word order in English.

important linguistic aspects. Furthermore, previous studies used procedures for which it was unclear when on-line processes were triggered. For instance, ERPs were measured during the decision phase (e.g., Sarlo et al., 2012) after reading, when automatic processes could already be gone. Appropriate methods and task procedures are hence crucial to give access to processes as they unfold in real time after their elicitation.

In a first study, Leuthold et al. (2015) addressed these aspects by using a language comprehension paradigm and word-by-word presentation of the critical target sentence. They constructed a new set of moral materials describing single everyday moral transgressions with an agent acting either morally or immorally (cf. Table 1). Assuming that the processing of moral transgressions corresponds to that of world knowledge violations, they expected an N400 for immoral than for moral behaviour. As a control and in order to test the N400's sensitivity regarding moral violations, they further constructed short scenarios that describe either violations or matches of world knowledge (cf. Table 1). Both text materials were constructed in a way that the meaning needed to be inferred while reading. The moral meaning of identical target sentences, and hence identical critical words (cf. Table 1), should be processed differently depending on the moderately constraining contexts. That is, the ambiguous situation set up by the context was resolved by the final critical sentence, which was presented using RSVP. Participants merely read the scenarios for comprehension while ERPs were measured.

The ERP pattern for world knowledge violations compared to matches was similar to that of Van Berkum et al. (2009), showing a larger P200, N400, and LPP amplitudes. The larger P200 and N400 amplitudes for world knowledge violations were taken to reflect enhanced (perceptual) processing and increased demands for meaning construction, respectively. In contrast, immoral as opposed to moral scenarios did not trigger an N400 effect but a larger positivity starting at already 320 ms after the critical word. Leuthold et al. (2015) related this early ERP positivity to the LPP component reflecting an implicit, affective evaluative categorisation process (Ito & Cacioppo, 2000; Van der Cruyssen, Van Duynslaeger, Cortoos, & Van Overwalle, 2009). Crucially, this evaluative categorisation process is specifically engaged during the on-line processing of moral transgressions.

The ERP results of Leuthold et al. (2015) indicated that readers process incoming information of either world knowledge or moral transgressions differently. They provided first evidence for discourse based effects during the processing of immoral as opposed to moral scenarios. Still, the assumption that the LPP reflects an implicit, affective evaluative process during discourse comprehension would need to be further substantiated.

One possibility to achieve this would be to investigate if a discourse-based emotional effect indicated by the LPP is also apparent for emotional scenarios that are constructed in the same way as moral scenarios.

3.2 Comparing moral and emotional information processing

For the present three studies, we were interested in how everyday moral transgressions are processed. Leuthold et al.'s morality scenarios are suitable for our purpose and meet the requirements regarding language comprehension research. These materials provide examples of everyday moral situations rather than relatively unrealistic moral dilemmas. I slightly modified them and added newly constructed short emotion scenarios which were not related to any moral content. Looking at the rating data, valence for moral scenarios was rated as neutral or even slightly positive. Therefore, neutral scenarios suited better to compare them to moral scenarios than positive ones. Emotion scenarios were taken to reliably indicate emotional processing like in previous text comprehension studies. That is, using emotional-negative and neutral contexts, emotion-related processing differences as indicated by the LPP were expected to critical words of the identical target sentences.

For both material sets, items are constructed in a way that moral and emotional content must be inferred from the context. For a given item, the context varies between the respective material-specific conditions followed by the identical target sentence. The critical word of the target sentence is embedded in a meaningful context and disambiguates the situation in a moral or immoral and neutral or negative way. Like previous studies, the RSVP procedure for the critical sentence was used in the present studies (cf., Van Berkum et al., 2009; Leuthold et al., 2015). Crucially, one major advantage of the present approach is that resulting ERP effects reflect discourse-based influences on information processing rather than influences related to the target word.

To study emotional processing and its response diversity, we followed the recommendations of Bradley and Lang (2000) to use an exhaustive measurement and not a single method. Consequently, we collected behavioural data (response time, explicit judgement behaviour), used ERPs to produce insights in the mental mechanisms underlying moral language comprehension as well as peripheral psychophysiological measures to indicate autonomic arousal (ECG, EDA) and involuntary facial expressions (fEMG) to indicate valence.

The main research questions are:

1. Does moral information undergo affective processing?

2. Does moral content of incoming linguistic information trigger embodied processing?

3. Does moral information processing differ from the processing of emotional information?

Study 1

The first study serves to investigate brain processes and the time-course of underlying elementary mechanisms and (affective) states when participants perform different judgement tasks after reading the morality and emotion scenarios. In two experiments, participants were asked to explicitly judge the moral acceptability of morality scenarios and the emotionality of emotion scenarios (Experiment 1), or only the emotionality of both materials (Experiment 2). Brain processes are measured by using ERPs. Based on the results of Leuthold et al. (2015), in Study 1 it was assumed that moral acceptability will be inferred from the context and involve the affective evaluation of linguistic input as indicated by the LPP. It is now important to assess whether similarly constructed emotion materials, which are not related to moral content also elicit the same LPP effect indicating the affective evaluation of incoming emotional linguistic information. In the light of evidence suggesting that the processing of incoming information depends on the task, it was expected that when explicit moral judgements are required, semantic-cognitive analysis of incoming information should dominate, whereas affective processing should dominate for emotional judgements (Lai et al., 2012; Sevinc & Spreng, 2014). Accordingly, if incoming moral and emotional information is affectively processed alike, then affective processing should also dominate for morality judgements, that is independent of the task. By contrast, if the processing of moral information depends on the task, there should be different processing patterns for moral and emotional content.

Study 2

If moral transgressions elicit affective evaluations as Study 1 suggests, there emerges another issue worth investigating, namely, whether the same materials trigger embodied responses, indicating affective responses outside conscious awareness of (alleged) affect-evoking test stimuli. During discourse comprehension of morality and emotion scenarios, the emotional salience of incoming linguistic information is measured by using three peripheral psychophysiological methods: (1) facial EMG, (2) EDA, and (3) ECG. Participants are asked to perform the same emotional judgement task as in Study 1. An additional control experiment measured the sensitivity of all three methods with stimuli that usu-

ally evoke psychological and physiological emotional reactions. To this end, we presented emotional-negative compared to neutral pictures and words.

Study 3

Study 3 was conducted to examine whether the same brain processes and embodied responses are still present if participants merely read for comprehension rather than perform an explicit judgement task. Study 1 and 2 are modified in a way that morality and emotion materials are presented intermixedly in a passive reading-for-comprehension task. In two separate experiments we measure ERPs (Experiment 1), as well as fEMG and EDA (Experiment 2).

3.3 Studies

The following chapters were written as separate manuscripts. This is why they differ from the rest of the monography in layout and style. Overlapping contents between the introduction and empirical chapters are therefore the case.

3.3.1 Study 1: Kunkel, A., Filik, R., Mackenzie, I.G., & Leuthold, H. (2018). *Task-dependent evaluative processing of moral and emotional content during comprehension: An ERP study*

Copyright Notice

Kunkel, A., Filik, R., Mackenzie, I. G., & Leuthold, H. (2018). Task-dependent evaluative processing of moral and emotional content during comprehension: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience*, 18, 389-409. doi: 10.3758/s13415-018-0577-5

Authors involved:

- Prof. Dr. Hartmut Leuthold (HL)

- Dr. Ian G. Mackenzie (IGM)
- Prof. Dr. Ruth Filik (RF)
- Angelika Kunkel (AK)

List of author contributions

- Study conception and design: HL and AK
- Material preparation: AK and HL
- Programming: IGM
- Acquisition of data: AK and bachelor students
- Analysis: IGM, AK, and HL
- Interpretation of data: HL, AK, IGM, and RF
- Writing paper: AK, HL, IGM, and RF

Task-dependent evaluative processing of moral and emotional content during comprehension: An ERP study

Angelika Kunkel

Department of Psychology, University of Tübingen, Germany

Ruth Filik

School of Psychology, University of Nottingham, United Kingdom

Ian Grant Mackenzie & Hartmut Leuthold

Department of Psychology, University of Tübingen, Germany

Address correspondence to:

Angelika Kunkel
Department of Psychology
Eberhard Karls University of Tübingen
Schleichstrasse 4
72076 Tübingen
Germany

Email: angelika.kunkel@uni-tuebingen.de
Phone: +49 (0)7071 29-75590

Abstract

Recently, we showed that when participants passively read about moral transgressions (e.g., adultery) they implicitly engage in the evaluative (good–bad) categorization of incoming information, as indicated by a larger event-related brain potential (ERP) positivity to immoral than moral scenarios (Leuthold, Kunkel, Mackenzie, & Filik, 2015). Behavioral and neuroimaging studies indicated that explicit moral tasks prioritize the semantic-cognitive analysis of incoming information but that implicit tasks, as used in Leuthold et al. (2015), favor their affective processing. Therefore, it is unclear whether an affective categorization process is also involved when participants perform explicit moral judgments. Thus, in two experiments, we used similarly constructed morality and emotion materials for which their moral and emotional content had to be inferred from the context. Target sentences from negative vs. neutral emotional scenarios and from moral vs. immoral scenarios were presented using rapid serial visual presentation. In Experiment 1, participants made moral judgments for moral materials and emotional judgments for emotion materials. Negative compared to neutral emotional scenarios elicited a larger posterior ERP positivity (LPP) about 200 ms after critical word onset, whereas immoral compared to moral scenarios elicited a larger anterior negativity (500-700 ms). In Experiment 2, where the same emotional judgment to both types of materials was required, a larger LPP was triggered for both types of materials. These results accord with the view that morality scenarios trigger a semantic-cognitive analysis when participants explicitly judge the moral content of incoming linguistic information but an affective evaluation when judging their emotional content.

Keywords: Moral judgment, emotion judgment, affective evaluation, LPP, anterior negativity

We often find ourselves in situations in which a person is violating a prevailing social norm or moral value. For instance, if we find out that someone cheats in an exam or is telling a lie, we tend to spontaneously judge such behavior as bad or immoral. This reflects a fundamental aspect of human moral cognition and it has been proposed that such judgments are based on affective or intuitive processes (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Haidt, 2001). In line with this view, a recent event-related brain potential (ERP) study demonstrated that when participants passively read about everyday moral transgressions, they implicitly categorize the described behaviors as good or bad as early as about 320 ms after the presentation of the critical word (cf. Leuthold, Kunkel, Mackenzie, & Filik, 2015), which was argued to be reflected by a late posterior positivity (LPP). However, there is also evidence in the literature suggesting that the semantic-cognitive analysis of incoming information dominates when explicit morality judgments are required, whereas the affective analysis is prioritized when emotion judgments are demanded (cf. Lai, Hagoort, & Casasanto, 2012; Sevinc & Spreng, 2014). Therefore, it remains unclear whether such a rapid evaluation process, as indicated by the LPP during a passive reading task, indeed would be found for explicit moral judgments as well. In addition, we aim to provide further support for the proposal that the LPP reflects affective processing of incoming information by also investigating whether an LPP is elicited by similarly constructed everyday emotional scenarios without a moral component. To this end, we will record ERPs that are elicited by scenarios describing moral transgressions and emotional events, to see whether, and to what extent, cognitive and affective processes are involved during discourse comprehension when explicit moral or emotional judgments are required.

Van Berkum, Holleman, Nieuwland, Otten, and Murre (2009) are to our knowledge the first to use a text comprehension approach to reveal the ERP correlates

of moral cognition. Specifically, they investigated whether and how rapidly an individual's values influence the online linguistic meaning analysis of moral statements when explicit judgments were required. Male participants with two opposing value systems (members of a Dutch strict-Christian party vs. voters of parties with opposite moral-ethical programs, referred to here as non-Christians) were asked to rate their agreement with critical statements such as, "If my child were homosexual, I'd find this hard/easy to accept". The individual words forming these statements were presented using rapid serial visual presentation (RSVP), affording the measurement of immediate ERP responses to the critical word. They found that value-inconsistent compared to value-consistent critical words (e.g., *easy* vs. *hard* for strict-Christians and *hard* vs. *easy* for non-Christians, respectively) initially elicited a larger broadly distributed positivity between 200 and 250 ms (P200), followed by a larger centroparietal negativity between 375 and 425 ms (N400), and finally a larger LPP between 500 and 650 ms.

N400 amplitude has been shown to respond to the predictability of a word within a given context (e.g., Kutas & Hillyard, 1984), to semantic anomalies at the discourse level (Van Berkum, Hagoort, & Brown, 1999), as well as to violations of world knowledge (e.g., Filik & Leuthold, 2008; Hagoort, Hald, Bastiaansen, & Petersson, 2004), reflecting the demands of meaning construction (for a review, see Kutas & Federmeier, 2011). Thus, Van Berkum et al. (2009) interpreted, their N400 findings as indicating that readers immediately and automatically evaluate incoming information with respect to their personally held values, giving rise to a rapid value-based influence on meaning construction. They further speculated that this N400 effect overlaps with that of a single sustained ERP positivity that has an earlier onset than the (overlapping) N400, therefore emerging as a larger P200 and LPP for value-inconsistent than value-consistent statements. Van Berkum and colleagues ruled out a cognitive,

decision-related account of this LPP effect for two reasons. Firstly, it has been demonstrated that self-referential (true vs. false) statements that are unrelated to a person's value system do not elicit such an effect (Fischler, Bloom, Childers, Arroyo, & Perry, 1984). Secondly, negative compared to positive and neutral stimuli tend to elicit a larger LPP - reflecting a negativity bias in affective processing (e.g., Ito, Larsen, Smith, & Cacioppo, 1998) - even in language studies where participants either merely read for comprehension or made explicit decisions to critical emotion words (e.g., Holt, Lynn, & Kuperberg, 2009). Therefore, Van Berkum et al. took their LPP effect to reflect the automatic activation of the affect system, in accord with the view that the LPP relates to the implicit evaluative processing of motivationally salient stimuli (cf. Hajcak, MacNamara, & Olvet, 2010).

However, in Van Berkum et al.'s (2009) study, the values held by the participants may have constrained their (implicit) expectations regarding the likely sentence endings. For instance, when persons holding strict Christian values read a statement (taken from their Table 1) beginning with "In a bad marriage, divorce is an ...", based on their personal beliefs, they would not expect it to be continued with the word "acceptable". Hence, similar to N400 effects driven by discourse- or world-knowledge-based expectations (e.g., Filik & Leuthold, 2008; Hagoort et al., 2004; Van Berkum et al., 1999), it is conceivable that the larger N400 elicited by value-inconsistent than value-consistent statements reflects an (implicit) emotional congruity effect that depends on the relation between the emotional features of the preceding context and the critical word.[1] Crucially, a larger N400 to emotion words that were incongruent rather than congruent with the preceding context has been shown not only in studies using

[1] We would like to thank an anonymous reviewer for highlighting the issue of emotion congruity as compared to emotion effects on ERP amplitudes.

sequential prime-target tasks (e.g., Eder, Leuthold, Rothermund, & Schweinberger, 2012; Morris, Squires, Taber, & Lodge, 2003; Zhang, Lawson, Guo, & Jiang, 2006; but see Herring, Taylor, White, & Crites, 2011) but also in discourse comprehension studies using strongly constraining emotional contexts, for instance, when someone is described as being happy in a context which outlines either a positive or a negative event (e.g., León, Díaz, de Vega, & Hernández, 2010; Leuthold, Filik, Mackenzie, & Murphy, 2012). Accordingly, the N400 effect might reflect the more intense lexical or semantic processing for incongruent than congruent moral statements, that is, a morality-unspecific language-related effect. If this conjecture would hold true, Van Berkum and colleagues' interpretation of the P200 and LPP effect in terms of an affective evaluation of statements could be challenged as well. That is, the P200 effect reported might be attributed to the enhanced visual processing of incongruent or very unexpected linguistic inputs (e.g., Bohan, Leuthold, Hijikata, & Sanford, 2012; Ferretti, Singer, & Patterson, 2008; Leuthold et al., 2015), and the larger LPP following incongruent statements might reflect a P600-like semantic effect that is found in response to various types of semantic anomalies (for a review, see Kuperberg, 2007) and has been related to a continued reanalysis of linguistic input following a semantic processing conflict (cf. Kuperberg, 2007; Van de Meerendonk, Kolk, Chwilla, & Vissers, 2009).

A recent text comprehension study by Leuthold et al. (2015) used a different approach to examine the implicit rather than explicit evaluative processing of everyday (fictional) scenarios that involved descriptions of moral transgressions (e.g., cheating on one's partner). Specifically, participants read the scenario context followed by the RSVP of the target sentence containing the critical word (cf. Table 1). The context determined whether the target sentence described a moral or an immoral event. As a control, participants read materials in which the target sentence was either consistent or

inconsistent with their knowledge of the world, to assess the ERP correlates elicited by the linguistic processing of moral-neutral world knowledge violations (e.g., a target sentence of "She receives as a dish a plate full of *snails* and white bread.", following a context that would make this statement either consistent with the participants' knowledge of the world, e.g., "During a France exchange Mrs. Lehmann eats a famous French specialty." or inconsistent, e.g., "Mrs. Lehmann goes to a Schwabian restaurant and orders a local specialty."). Morality and world knowledge materials were randomly interleaved and no explicit judgments were required.

Crucially, a larger P200 amplitude was found both for moral transgressions and for world knowledge violations, indicating domain-unspecific, enhanced attentive processing of materials conflicting with the discourse context. Subsequently, a large posterior N400 was found for general world knowledge violations only. In accord with previous studies from our lab (e.g., Filik & Leuthold, 2008; 2013) and with the N400 literature in general (cf. Kutas & Federmeier, 2011), this was taken to reflect the increased semantic memory demands involved in retrieving and integrating conceptual information during meaning construction when knowledge-based expectations are violated (e.g., Filik & Leuthold, 2008; Hagoort et al., 2004). By contrast, moral transgressions did not trigger a larger N400 but only a larger central-maximal ERP positivity after about 320 ms. Leuthold and colleagues took this finding to reflect an LPP effect, proposing that incoming socio-normative information is, during a first step, implicitly evaluated and categorized as good or bad (see also Cunningham & Zelazo, 2007). This is in line with theoretical views that assume a central role of emotional-intuitive processes for moral judgment (Greene et al., 2001; Haidt, 2001).

More generally, the ERP study of Leuthold et al. (2015) demonstrates the practicality of approaching the (implicit) mechanisms contributing to moral cognition

by having participants read fictional scenarios with moral content. In contrast to Van Berkum et al. (2009), a passive reading task was used in which the moral versus immoral nature of the (identical) target sentences had to be inferred depending on the discourse context. That is, the materials did not involve incongruent moral statements but instead described scenarios that participants in a pre-test had judged as either clearly morally good versus bad, which would explain the absence of an N400 (congruity) effect. Also, since no explicit moral judgments were required, we consider it more likely that the LPP effect reported by Leuthold et al. reflects the implicit (affective) evaluation of morality-related materials. Emotion effects on the ERP waveform are known to depend on the emotional features of the critical item (e.g., valence, arousal), with emotional stimuli such as positive or negative words, pleasant and unpleasant pictures, and arousing stimuli reliably eliciting larger LPP amplitudes than neutral or less arousing stimuli, and this effect is more pronounced when participants judge the emotional content (e.g., in an affective judgment task) rather than an emotion-irrelevant stimulus dimension (e.g., in a semantic classification or passive reading task) (for reviews, see Citron, 2012; Fischler & Bradley, 2006; Hajcak, Weinberg, MacNamara, & Foti, 2012). Hence, it seems reasonable to assume that the LPP elicited in the Leuthold et al. study reflects an emotion effect.

As stated above, the study conducted by Leuthold et al. (2015) did not involve any explicit judgment task. However, there is behavioral evidence suggesting that task demands influence whether an affective versus semantic-cognitive analysis is prioritized (e.g., Lai et al., 2012). Importantly, for the present purposes, evidence from functional magnetic resonance imaging (fMRI) studies corroborates this conjecture for the processing of moral content. That is, fMRI studies consistently indicate that brain areas concerned with both cognitive and emotional processing are activated during moral

judgment tasks using dilemma scenarios (e.g., Greene et al., 2001) and socio-normative scenarios (e.g., Moll, de Oliveira-Souza, Bramati, & Grafman, 2002). Crucially, in a meta-analysis of a total of 40 fMRI studies (Sevinc & Spreng, 2014), brain areas concerned with cognitive processing were more strongly activated than areas linked to emotional processing in studies using explicit moral judgment tasks, whereas the reverse pattern of brain activation was found in studies using implicit (e.g., reading) tasks. In line with these findings, evidence from social cognition research suggests that the impact of automatic evaluations is reduced when participants deliberately rather than implicitly process incoming information (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996). Of course, given the limited temporal resolution of fMRI, the precise time course of task-dependent emotional versus cognitive influences on moral judgments is not yet completely understood (Avramova & Inbar, 2013). Thus, it is essential to investigate this issue using a combined behavioral and ERP approach in order to test whether an explicit moral judgment task would enforce a semantic-cognitive analysis of morality materials as indicated by the N400. We will address this issue by conducting an experiment in which participants read the materials used by Leuthold et al., but in the context of an explicit morality judgment task.

If one assumes that moral acceptability is inferred from the context and involves the affective evaluation of linguistic input, it is important to assess whether similarly constructed emotional materials without moral content also elicit an LPP effect. At present, we are not aware of any published ERP studies investigating the processing of materials where target sentences are identical across conditions, and the emotional meaning of the target needs to be inferred from the context in which it appears. Specifically, previous ERP studies examining discourse-based emotion effects used contexts that were strongly constraining (e.g., León et al., 2010; Leuthold et al.,

2012) or employed materials for which the critical words differed across emotion conditions (e.g., Delaney-Busch & Kuperberg, 2013; Holt et al., 2009; León et al., 2010). For example, Delaney-Busch and Kuperberg found a larger N400 (300-500 ms) for incongruent than congruent neutral words following a neutral context, and this congruity effect was larger over anterior than posterior midline electrodes. Crucially, when an emotional discourse context preceded valence congruent or incongruent emotion words, no congruity effect was observed in N400 amplitude. Rather, a larger LPP (500-700 ms) to pleasant and unpleasant emotion words was elicited irrespective of the valence of the preceding emotional discourse context. In accordance with the affective primacy hypothesis (Storbeck & Clore, 2007), these findings led Delaney-Busch and Kuperberg to suggest that for emotional contexts, the processing of incoming information is dominated by the analysis of their motivational (e.g., approach vs. avoidance) rather than semantic significance. It is therefore unclear whether emotion materials for which the target sentences and the critical (emotion) words are identical across conditions, and hence the emotional meaning has to be inferred from the discourse context, elicit an LPP effect as well. Thus, it is also a major aim of the current work is to close this research gap concerning our understanding of emotional language processing, and this is suited to strengthen the interpretation of the LPP as reflecting the affective evaluation of linguistic input during discourse comprehension.

Objectives of the present study

In summary, it remains to be investigated first whether the rapid affective evaluation of descriptions of moral transgressions during text comprehension (i.e., when there is no explicit judgment task), as inferred from the LPP effect by Leuthold et al. (2015), is also observed when participants perform explicit moral judgments. If such an

LPP effect but no N400 effect would be present, this outcome would lend support to the view that incoming linguistic information undergoes an implicit (or task-independent) affective evaluation. By contrast, if these task conditions enforce a semantic-cognitive analysis of morality materials, a larger N400 to immoral than moral items should be triggered. Second, it is crucial to investigate the electrophysiological correlates of discourse-based emotion comprehension, specifically, whether an LPP effect is also elicited when the emotional meaning is inferred from the discourse context. The assumption that the LPP indicates the discourse-dependent affective processing of linguistic input, as assumed in the moral ERP study of Leuthold et al. (2015), would be corroborated by showing that for the same participants, discourse-dependent negative compared to neutral (or positive) items elicit a similar LPP effect to discourse-dependent immoral compared to moral items. Therefore, we created novel emotion materials that were similar to the morality materials with regard to critical dimensions such as the cloze probability of the critical words, their semantic relatedness to the discourse context, critical word frequency, as well as their emotionality in terms of valence and arousal (cf. Methods). Of course, since morality and emotion materials differ with regard to the wording of the critical sentences and hence are not matched regarding all potentially relevant word-level or discourse-level dimensions, this allows only an indirect comparison of the ERP effects triggered by these materials.

We recorded ERPs in two experiments to investigate task-related influences on the online processing of scenarios describing everyday moral compared to emotional situations and the nature of the underlying, potentially affective, processes. Thus, our setup was identical to that of Leuthold et al. (2015) except that (a) instead of world knowledge violations, we used emotional scenarios without moral content as a control condition, and (b) that participants performed explicit judgments of the materials. More

specifically, we used prototypical scenarios for which the protagonists and situations were introduced by the context sentences (for an example, see Table 1).

For morality materials, the target sentence either described a morally acceptable or unacceptable (that is, moral vs. immoral) action, and for emotion materials, the target sentence described either a relatively neutral versus a negative event, which was determined by the context for both materials.

We used RSVP for the final critical sentence, with participants performing their judgment response (yes/no) after the presentation of the final word. We chose a binary judgment task in line with recent moral dilemma and moral judgment studies (e.g., Greene et al., 2001). In Experiment 1 participants made moral judgments for morality materials (i.e., "Is the behavior morally acceptable?") and emotional judgments for emotion materials (i.e., "Are you emotionally moved by the text?"). Experiment 2 required emotional judgments for both of material.

Generally, we hypothesized that linguistic input is affectively evaluated, that is, independently of the specific content of the materials (cf. Bargh et al., 1996; Cunningham & Zelazo, 2007). It is then reasonable to assume that moral actions and neutral (or mildly positive) events are evaluated as potentially "good" and immoral and negative events as potentially "bad". Since such differential affective evaluations are taken to be reflected by the LPP component (cf. Fischler & Bradley, 2006; Hajcak et al., 2012), LPP amplitude should be larger for both immoral and negative compared to moral and neutral (or mildly positive) scenarios. We further reasoned that if these evaluations are automatic in the sense that they are produced by a fast-operating process that is independent from task goals, qualitatively the same LPP effects should be observed in the two experiments for immoral versus moral scenarios. However, if the requirement to judge the moral content prioritizes the semantic-cognitive processing of linguistic

input in Experiment 1, a larger N400 rather than a larger LPP might be elicited for immoral than moral items.

Experiment 1

Participants were presented with morality and emotional scenarios in separate blocks of trials. In the case of a morality scenario, they judged whether someone's behavior was acceptable or not. Here, we predicted that immoral compared to moral scenarios would be judged as less acceptable, hence producing fewer yes-responses. For emotion materials, participants judged whether they were emotionally moved by the text or not. We predicted that negative compared to neutral scenarios are more moving and therefore would produce more yes-responses.

It is important to note that we performed rating studies (see method section for details) to assess the moral acceptability of morality items as well their plausibility, valence, and arousal value and that the same dimensions were assessed for emotion materials (except their moral acceptability). Based on these results, the morality items were classified as either moral or immoral while the emotion items were classified as either neutral or negative. This procedure guarantees that the respective materials are neatly matched across conditions. To examine whether item-specific arousal and valence characteristics, as obtained from the rating studies, contribute to present binary emotion and morality judgments in addition to condition-specific effects, a logistic regression approach was used.[2]

Moreover, since binary morality and emotion judgments (yes- versus no-response) are required in the present experiments, it is conceivable that participants may apply decision criteria that lead, at least sometimes, to judgments that are inconsistent

[2] We would like to thank an anonymous reviewer for suggesting the logistic regression analysis and also for his most helpful comments regarding the judgment-dependent analysis of ERP amplitudes.

Table 1. Example for Moral Materials with Context for Moral and for Immoral Items, for Emotional Materials with Context for Neutral and for Negative Items, as well as the respective Target Sentences Containing the Critical Word (in italics).

Morality		Emotion	
Moral	Immoral	Neutral	Negative
Herr Zimmermann arbeitet in einem Chemiekonzern. Seit einigen Wochen gibt es Sicherheitsprobleme aufgrund fahrlässigen Verhaltens von Kollegen, sodass Menschenleben gefährdet sind. (Mr. Zimmermann works at a chemical company. For several weeks, there have been safety issues caused by colleagues' careless behavior, putting human lives at risk.)	Herr Zimmermann arbeitet in einem Chemiekonzern als Sachbearbeiter in der Personalabteilung. Er ist ehrgeizig und will rascher aufsteigen als seine Kollegen. (Mr. Zimmermann works in human resources of a chemical company. He is ambitious and aims at being promoted faster than his colleagues.)	Sarah trainierte den ganzen Nachmittag im Fitness-Studio und geht zum Abschluss in die Sauna, wonach sie sich wohlig gut fühlt. (Sarah had been at the gym all afternoon before completing her work out with a visit to the sauna, after which she feels pleasantly well.)	Sarah hat einen sehr anstrengenden Job. Sie muss dafür seit mehreren Monaten jeden Tag um 5 Uhr aufstehen und den ganzen Tag hart arbeiten. (Sarah has a very demanding job. For several months she has been getting up at 5 am and working very hard all day.)
Im monatlichen Gespräch berichtet er seinem Chef deren *Fehler*. (In a monthly commitment talk he points out to his boss their *mistakes*.)		Sie ist sehr *erschöpft*. (She is very *exhausted*.)	
Ben ist wegen sexueller Belästigung angezeigt worden. Er war an dem fraglichen Nachmittag bei einem Freund. Dieser Freund wird von der Polizei befragt. (Ben was sued for sexual harassment. The day in question he was at a friend's. This friend is questioned by the police.)	Ben ist wegen sexueller Belästigung angezeigt worden. Er behauptet an diesem Nachmittag bei einem Freund gewesen zu sein, was jedoch gelogen ist. Sein Freund wird nun befragt. (Ben was sued for sexual harassment. He claims that he was at a friend's, which is a lie. This friend is questioned by the police.)	Dominik fliegt mit seinen Freunden im Hochsommer nach Südspanien. Sie gehen lieber an den Strand als die Stadt zu besichtigen.	Dominik beeilt sich das Essen für seine wartenden Gäste zu servieren. Ohne zu überlegen greift er in den Kochtopf mit den Eiern, die noch im Wasser schwimmen.

Er verschafft ihm das *Alibi*.
(He is his *alibi*.)

Erik studiert Mathematik. Für die Übungsaufgaben hat er sich die Lösungen einer besonders kniffligen Aufgabe beim Professor besorgt. Dieser hat ihn darum gebeten, aus Fairness die Lösungen an die anderen Studierenden weiterzureichen.
(Erik majors in math. He asked the professor for the answers to a particularly tricky homework problem. As a matter of fairness, the professor asked him to pass on the answers to his fellow students as well.)

Er hat den anderen die Lösungen *gegeben*.
(He *passed on* the answers to them.)

Frau Bauer möchte zum Abteilungsessen gehen und sucht einen Babysitter. Sie erhält die Anfrage einer Frau, von der man weiß, dass sie gegenüber Kindern bereits handgreiflich wurde. (Mrs. Bauer is looking for a babysitter because she wants to attend the company dinner. A lady contacts her, of whom it is known that she beat children in the past.)

Sie hat der Babysitterin *abgesagt*.
(She has *not invited* the babysitter to work for her.)

Erik studiert Mathematik. Für die bevorstehende Klausur hat er sich bei einem Freund des höheren Semesters die Lösung besorgt. Dieser hat ihn allerdings darum gebeten die Lösungen nicht herumzureichen, weil er sonst exmatrikuliert werden könnte.
(Erik majors in math. He got the answers for an upcoming exam from a friend, who is already a senior. The friend asked him not to pass on the answers to anyone else, otherwise he would get in trouble and have to leave the school.)

Frau Bauer möchte zum Abteilungsessen gehen und sucht einen Babysitter. Sie erhält die Anfrage einer jungen Studentin, die am Telefon sehr sympathisch wirkt, aber mit ausländischem Akzent spricht. (Mrs. Bauer is looking for a babysitter because she wants to attend the company dinner. A young student girl contacts her who sounds very friendly but speaks with a foreign accent.)

Es ist zu heiß.
(It is too *hot*.)

Herr Meier hat Streit mit seiner Frau und rennt wütend aus dem Haus. Noch voller Zorn, achtet Herr Meier nicht auf die Straße und sieht den heranfahrenden PKW nicht. (Mr. Meier has an argument with his wife and storms out of the house in anger. Still upset, he does not mind the traffic and overlooks the approaching car.)

Herr Meier wird von dem Auto *angefahren*.
(Mr. Meier is *hit* by the car.)

Magda überrascht ihren Freund, da sie soeben eine Last-Minute-Reise gewonnen hat und diese gerne mit ihm verbringen möchte. (Magda has won a last-minute trip and surprises her boyfriend by inviting him to join her.)
Er stürmt ins Zimmer und packt den *Koffer*. (He goes and packs his *bags*.)

Herr Meier recht den Rasen, während sein Sohn dort mit dem ferngesteuerten Spielzeugauto spielt. (Mr. Meier rakes the lawn, while his son plays there with a remote-controlled toy car.)

Magda hat wiederholt einen heftigen Streit mit ihrem Freund, der bereits beim letzten Mal drohte sie zu verlassen und aus der Wohnung auszuziehen. Der Disput eskaliert. (Magda and her boyfriend have a serious fight again. The last time he already threatened to leave her and move out. The fight went out of control.)

with the rating-based morality or emotion classification of the materials. That is, some items pre-classified as moral might be judged as morally unacceptable, or some items pre-classified as negative might be judged as neutral and vice versa. Therefore, we performed additional ERP amplitude analyses in waveforms averaged for moral items that were judged as appropriate (yes-response) and for immoral items judged as inappropriate (no-response). Likewise, such judgment-dependent ERP analyses were also conducted for neutral items that were judged as not moving (no-response) and negative items judged as moving (yes-response).

Method

Participants. Thirty-two native German speakers from the University of Tübingen received course credits or payment for participating. Data from four participants were excluded due to excessive alpha activity. For all analyses, we used the dataset from the remaining 28 participants (M = 24.5 years, 19 females).

Materials and design. Morality materials were taken from Leuthold et al. (2015) and modified.[3] These materials consisted of a total of 160 items, resulting from the combination of 80 identical target sentences each with two different discourse contexts, thereby creating 80 moral and 80 immoral items. The 160 emotion materials were newly generated and analogously constructed (see Table 1 for examples; the full set is available from the first author). Both morality and emotion materials were pre-tested (see below).

All scenarios consisted of two parts. The first part consisted of two or three sentences describing the context, and the second part was the target sentence containing

[3] The original morality materials were changed in order to disambiguate the meaning of some items and to consistently present critical words towards the end of the target sentence. As a result, two items were replaced. 52 discourse contexts were slightly shortened or the protagonist's name changed. Of the 80 target sentences, 15 were shortened and for eight target sentences the critical word was replaced.

the critical word. In order to eliminate possible sentence-level and word-based effects, the same target sentence was used for moral and immoral conditions and the same held true for neutral and negative emotional conditions (with the context varying across conditions, see Table 1). The critical word was always presented towards the end of the target sentence, most frequently as the sentence-final word (84.4%). Critical words were predominantly verbs describing a certain behavior (e.g., to borrow, to report, to mention, to swap) and nouns (e.g., acceptance, alibi, verdict, tumor).

Morality materials described actions that would be perceived as either moral or immoral, whereas emotion materials would describe a neutral or a negative event. Finally, 40 neutral filler items were constructed which contained no moral or emotional content as well as no inconsistencies, and were similar in length to the experimental items (e.g., Context sentence: "Herr Krüger hat kein aktuelles Telefonbuch. Er braucht die Nummer seines Hausarztes." Target sentence: "Er ruft bei der Auskunft an, um an die Nummer zu gelangen." [Context sentence: Mr. Krüger does not possess an up-to-date phone book. He needs the telephone number of his general practitioner. Target sentence: He calls the directory enquiries service to find out the number.]). Following the presentation of the final word, for the morality blocks, the following question was displayed on the screen: “Ist das Verhalten moralisch akzeptabel?” [“Is the behavior morally acceptable?”], and for the emotion blocks: “Berührt dich das Gelesene?” [“Are you emotionally moved by the text?”].[4] Participants indicated their response (“Ja” [“Yes”] versus “Nein” [“No”]) by pressing the left or right arrow key on the computer keyboard.

[4] There is no one-to-one translation of the German word "berührt" that would match its emotional connotation, yet in our view "emotionally moving" comes close. Note that in contrast to the rating study described below, this term was not explicitly related to the state of arousal.

The randomization of items and conditions across participants was performed in the following way. The two different types of scenario (morality vs. emotion) were presented in the first versus second half of the experiment, and their order was counter-balanced across participants. For two consecutive participants, two lists were randomly generated such that each morality scenario appeared across the two lists either in the moral or the immoral condition, and each emotion scenario appeared either in the neutral or the negative condition. That is, the two participants received the same target sentence but with a different context in order to manipulate either the morality condition (moral vs. immoral) or the emotion condition (neutral vs. negative). Thus, for each participant, the 200-item list consisted of 40 moral and 40 immoral items, 40 neutral and 40 negative items, as well as 40 neutral filler items. The fillers were included in order to keep the procedure as similar as possible to the study of Leuthold et al. (2015) and to reduce a potential influence of extended local runs of immoral and negative items on ERPs. For instance, after the description of several immoral behaviors, participants might relax their judgment criteria and view immoral acts as more acceptable, which could potentially reduce the N400 effect (cf. Baetens, Van der Cruyssen, Achtziger, Vendekerckhove, & Van Overwalle, 2011).

Pre-test of materials. For the newly created emotion scenarios, we used a web-based questionnaire to assess the plausibility, valence, and arousal ratings of the materials. Altogether, we recruited 293 undergraduate students from the University of Tübingen (M = 23.6 years, 204 females). The 160 scenarios (80 items each with two conditions neutral/negative) were arranged in four lists each containing 40 randomly arranged scenarios plus the target sentence; each list was rated by no less than 66 participants. Participants were asked to rate on scales from 1-8 (a) how plausible they found the scenario ("Die beschriebene Situation ist ..." [The scenario described is ...]:

1 = sehr unrealistisch [very unrealistic] to 8 = sehr realistisch [very realistic]), (b) their "Erregungszustand" [arousal] in terms of how much they were emotionally moved by the scenario (1 = nicht ergreifend [not moved at all] to 8 = stark ergreifend [strongly moved]), and (c) the valence of the materials (1 = sehr negativ [very negative] to 8 = sehr positiv [very positive]). Two-tailed *t*-tests (cf. Table 2) showed that neutral and negative items were rated as being equally plausible (M = 6.01 vs. 6.14), $t(79) = 0.98$, $p = .33$. Furthermore, negative items were rated as being more negative (M = 2.39 vs. 5.07), $t(79) = 20.02$, $p < .001$, and more moving than neutral items (M = 5.00 vs. 3.51), $t(79) = 9.70$, $p < .001$.

For the morality materials, pre-tests for plausibility and morality were carried out using a web-based questionnaire (N = 55 participants). On a scale from 1 (sehr unmoralisch; sehr unrealistisch [very immoral; very unrealistic]) to 8 (sehr moralisch; sehr realistisch [very moral; very realistic]), moral items were rated as being morally more acceptable than immoral items (M = 5.99 vs. 2.52), $t(79) = 25.93$, $p < .001$, and also as being slightly more plausible (M = 6.21 vs. 5.15), $t(79) = 3.88$, $p < .001$.

Additionally, valence and arousal ratings for the morality materials were collected from a fresh group of participants (N = 40). On a scale from 1 (sehr negativ; nicht ergreifend [very negative; not emotionally moving]) to 8 (sehr positiv; ergreifend [very positive; emotionally moving]), moral items were rated as more positive (M = 5.44 vs. 2.60), $t(79) = 16.35$, $p < .001$, and less moving (M = 3.79 vs. 4.34), $t(79) = 3.20$, $p < .01$, than immoral items.

To compare valence, arousal, and plausibility scores across the two sets of materials, these rating scores were separately analyzed using ANOVAs with factors material and condition. Given the above-reported analysis of condition effects, significant results will only be reported for the main effect of material and the interaction

Table 2. Characteristics and rating data of morality and emotion materials.

	Morality		Emotion	
	Moral (SE)	Immoral (SE)	Neutral (SE)	Negative (SE)
Cloze	43 % (4 %)	38 % (4 %)	40 % (4 %)	47 % (4 %)
Semantic Relatedness	0.46 (0.01)	0.45 (0.01)	0.40 (0.01)	0.41 (0.01)
Word Frequency	59.46 (9.72)		58.31 (6.98)	
Word length	8.66 (0.21)		7.51 (0.19)	
Plausibility	6.21 (0.03)	5.15 (0.04)	6.14 (0.11)	6.01 (0.11)
Valence	5.44 (0.13)	2.60 (0.09)	5.07 (0.12)	2.39 (0.07)
Arousal	3.79 (0.13)	4.34 (0.11)	3.51 (0.13)	5.00 (0.12)
Morality	5.99 (0.04)	2.52 (0.03)		

Note. Means and standard errors (in brackets) of the rating data were calculated for each material and condition. Plausibility, valence and arousal ratings concern the whole scenario inclusive target sentence with 1 (= *very unrealistic*, *very negative*, *not touched*, *very immoral*) to 8 (= *very realistic*, *very positive*, *strongly touched*, *very moral*). Only morality materials were rated for the degree of morality with 1 (= *very immoral*) to 8 (= *very moral*). Word frequency (per million) and word length concern the critical words, cloze probability (as percentages) and semantic relatedness was calculated for the critical word in relation with a context.

of material and condition. For plausibility, the main effect of material, $F(1, 316) = 15.05$, $p < .001$, and the Material x Condition interaction were significant, $F(1, 316) = 23.88$, $p < .05$, as both neutral and negative emotion items were more plausible than immoral items, all ps $< .001$, but were not more plausible than moral items, all ps $> .21$. For arousal ratings, the Material x Condition interaction was significant as well, $F(1, 316) = 14.34$, $p < .001$, indicating a stronger condition effect for emotion than morality materials. Finally, for valence ratings, the main effect of material was significant, $F(1, 316) = 7.48$, $p < .01$, due to a lower valence score for emotion compared to morality items.

Moreover, all materials were analyzed with regard to critical word frequency, cloze probability, and semantic relatedness. For calculating word frequencies, we chose the SUBTLEX-DE corpus (Brysbaert et al., 2011). Two words were not listed in the

corpus. The frequencies (per million) of the remaining critical words did not differ between materials (morality vs. emotion: M = 59.46 vs. 58.31), $t(314) = 0.10$, $p = .92$.

To determine cloze probability, participants were presented with both the context and the target sentence without the critical word, which they were asked to fill in. Due to an error, no cloze probability scores were obtained for two moral items and one immoral item. Cloze probability did not reliably differ between the moral (M = 0.43) and the immoral condition (M = 0.38), $t(155) = 1.04$, $p = .30$, and also not between the neutral (M = .40) and the negative condition (M = 0.47), $t(158) = 1.36$, $p = .18$. There were also no significant differences in cloze probability for the critical words of morality materials (M = .40) and of emotion materials (M = 0.43), $t(315) = 0.73$, $p = .46$.

Finally, we calculated semantic relatedness as the cosine similarity between the context and the critical word with the LSAfun package in R and the German dewak100k_lsa corpus as semantic space (Günther, Dudschig, & Kaup, 2015) based on Latent Semantic Analysis (LSA) (Landauer, Foltz, & Laham, 1998). Seven critical words of the morality materials and five critical words of the emotion materials were not listed in the corpus. In a separate analysis of semantic relatedness scores for the two sets of materials, there was no significant difference between moral (M = 0.46) versus immoral (M = 0.45) sentences, $t(144) = 0.09$, $p = .93$, and also not between neutral (M = 0.40) versus negative (M = 0.41) sentences, $t(148) = 0.83$, $p = .41$. However, the comparison between materials revealed a higher semantic relatedness score for morality than emotion materials (M = 0.45 vs. 0.40), $t(294) = 3.44$, $p < .001$.

Procedure. After electrode application, participants were seated in an electrically shielded booth in front of a 21-in. computer monitor (60 Hz) at a viewing distance of 65 cm (maintained by a chin rest). Experimental materials (context, words) were presented at the center of the screen in white 16-point Helvetica font on a black

background using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) running under MATLAB (2012b), on an Apple Mac Mini (OS 10.7.). Participants were instructed to avoid any eye, head, and jaw movements and to maintain fixation at the center of the screen during word-by-word presentation. Furthermore, they were instructed to read the stories attentively, and to perform the respective judgments by pressing the appropriate response key.

For each of the morality and emotion materials, a practice block containing three trials preceded the experimental items that were presented in a total of four blocks of 25 items each. Blocks were separated by a short break that was controlled in its duration by the participant. Participants started a trial block by pressing the space bar. Then, the context was displayed for a minimum duration of 1,500 ms. When participants had read the context sentences, they initiated the word-by-word presentation of the target sentence by pressing the space bar, which started with the presentation of a fixation point for 1000 ms. Then, each word was displayed centrally for 300 ms, with a 200-ms blank interval between successive word presentations. After the offset of the final word and a blank interval of 1,100 ms, the presentation of the decision screen followed, stating the mapping of judgments (yes-no) to response keys. This mapping was constant within a given participant but counterbalanced across participants.

Electrophysiological measures. Electroencephalographic (EEG) activity was recorded continuously without online low-pass filtering from 72 Ag-AgCl electrodes using a BIOSEMI Active-Two DC-amplifier system with a sampling rate of 512 Hz for EEG and electrooculogram (EOG). All EEG/ERP analyses were performed using available MATLAB toolboxes (EEGLAB: Delorme & Makeig, 2004; FieldTrip: Oostenveld, Fries, Maris, & Schoffelen, 2011) and custom MATLAB scripts (for details, see Dudschig, Mackenzie, Strozyk, Kaup, & Leuthold, 2016). The analysis epoch started

200 ms prior to the onset of the critical word and lasted for 1,700 ms. For preprocessing purposes, signals from all EEG channels were off-line recalculated to an average reference and high-pass filtered (Butterworth filter, 0.1 Hz, 12 dB/oct). (Ocular) artifacts were then removed and EEG data were corrected (for a similar procedure, see Nolan, Whelan, & Reilly, 2010). As in Dudschig et al., a predefined z-score threshold of ±3 was used to identify outliers relating to channels, epochs, independent components, and single-channels in single-epochs. Firstly, epochs containing extreme values in single electrodes (e.g., amplifier blockings, values larger ±1000 μV in any electrode) were removed, as were trials containing values exceeding ±75 μV in multiple electrodes that were unrelated to eye movements. Secondly, z-scored variance measures were calculated for all electrodes, and noisy EEG electrodes (z-score > ±3) were removed if their activity was uncorrelated to EOG activity. Thirdly, a spatial independent components analysis (ICA) based on the infomax algorithm (Bell & Sejnowski, 1995) was performed on the "cleaned" EEG data set and ICA components reflecting ocular activity (blinks and horizontal eye movements) were removed from this data set (*M*[removed components] = 3.4). Fourthly, previously removed noisy channels (*M* = 2.35, range = 0 to 5) were interpolated in the ICA-cleaned EEG data set using the average EEG activity of adjacent uncontaminated channels within a specified distance (4 cm, ~ 3-4 neighbours per electrode) in order to ensure a full electrode array for each participant. The mean number of trials remaining (*M* = 36.75 out of 40; range = 21-40, median = 38.0) per condition was not reliably different across conditions, all *p*s > .40.

Data analysis. For artifact-free trials, the signal at each electrode site was averaged separately for each experimental condition, time-locked to the onset of the critical word, and low-pass filtered (Butterworth filter, 30 Hz, 36 dB/oct). In addition, all EEG channels were recalculated to an average mastoid reference as in Leuthold et al.

(2015) and aligned to a 200-ms baseline prior to the onset of the critical word. To facilitate comparison across studies, similar to previous moral and emotion comprehension studies, mean ERP amplitudes were determined for the following time ranges: 200 to 250 ms (P200; as in Leuthold et al, 2015; Van Berkum et al., 2009), 300 to 500 ms (N400; as in Delaney-Busch & Kuperberg, 2013; Leuthold et al,, 2015), and 500 to 700 ms (LPP, as in Delaney-Busch & Kuperberg, 2013; Holt et al., 2009; and similar to Van Berkum et al., 2009). Since P200 effects are typically larger over anterior midline electrodes (e.g., Bohan et al., 2012; Leuthold et al., 2015), whereas N400 and LPP effects usually show a more pronounced centroparietal distribution (e.g., Delaney-Busch & Kuperberg, 2013; Holt et al., 2009; Leuthold et al., 2015; Van Berkum et al., 2009), midline electrodes were pooled to form an anterior (AFz, Fz, FCz) and a posterior region-of-interest (ROI; CPz, Pz, POz).

Statistical analyses of reaction times and ERP amplitudes were performed by means of repeated measures analyses of variance (ANOVA). The analysis of the binary yes-no judgments were analyzed using a logit model as recommended by Jaeger (2008), implemented via the glmer function within the lme4 R-package (Bates, Maechler, Bolker, & Walker, 2014). Separate glmer model fitting procedures were implemented for morality and emotion materials. The model was specified with fixed effects of condition, valence, and arousal with random intercepts for participants and items (i.e., answer ~ condition + valence + arousal + (1|participants) + (1|items)). For all statistical analyses, the significance level was set to alpha = .05.

Complementing the standard, condition-dependent ERP analysis, judgment-dependent ERP analyses were conducted as mentioned in the introduction. That is, we measured ERP amplitudes in waveforms averaged for moral items that were judged as acceptable (yes-response) and for immoral items judged as unacceptable (no-response)

as well as for neutral items that were judged as not moving (no-response) and negative items judged as moving (yes-response). It is worth mentioning that a possible limitation of this judgment-dependent analysis relates to the fact that, in contrast to the standard analysis, items might not be perfectly matched (i.e., in terms of contexts and critical words presented) across the respective experimental conditions. The ANOVA performed on these ERP amplitude data will be reported after the standard ERP analysis.

Results

Behavioral measures

Separate logistic regression analyses were performed for the emotion and morality materials to determine the impact of condition, valence, and arousal for the respective binary judgments. Moral items were more often judged as acceptable than immoral items (84.91 vs. 17.77%, p <.001), and negative items were judged more often as being emotionally moving than neutral items (68.03 vs. 34.29%, p <.001). For moral materials, there was a significant effect of condition (β = -3.23, *SE* = 0.28, *Wald Z* = -11.53, p < .001) and valence (β = 0.38, *SE* = 0.08, *Wald Z* = 4.69, p < .001). These results suggest that the likelihood of yes-responses ("acceptable") decreased for immoral items, and increased for more positively rated items.

For emotion materials, there was a significant effect of condition (β = -0.48, *SE* = 0.23, *Wald Z* = -2.06, p < .05), valence (β = -0.30, *SE* = 0.07, *Wald Z* = -4.19, p < .001) and arousal (β = 1.15, *SE* = 0.07, *Wald Z* = 17.15, p < .001), indicating that the likelihood of yes-responses ("moving") decreased for neutral items and for more positively rated items but mainly increased for more arousing items.

The separate ANOVAs performed on reaction time (RT) data yielded faster responses to immoral than moral items (1370 vs. 1603 ms), $F(1, 27) = 6.43$, $p < .05$, η_p^2

= .19 for the morality materials. For emotion materials, there was a trend for faster responses to negative than neutral items (1010 vs. 1127 ms), $F(1, 27) = 3.80$, $p = .06$, $\eta_p^2 = .12$.

Electrophysiological measures - condition-dependent ERP results.

ERP waveforms averaged according to the rating-based item classification (as determined by the pre-tests discussed above) are shown in Figure 1. For the ERP data, we performed separate ANOVAs for morality and emotion materials on mean ERP amplitudes at midline electrodes with factors *condition* (moral vs. immoral or emotional-neutral vs. negative) and *ant-post* (anterior vs. posterior) ROI.

For both types of materials, analyses of mean ERP amplitudes coincided with an overall main effect of ant-post, indicating an anterior positivity for the early P200 time window (200-250 ms), and a posterior positivity for the subsequent time windows (300-500 ms, 500-700 ms). For the sake of brevity, we refrain from reporting the respective main effects of ant-post (all Fs(1, 27) < 5.77, all ps < .05) in the following.

200-250 ms (P200). In this time window, there were no reliable condition effects neither for morality materials, all Fs < 1.07, ps > .31, nor for emotion materials, all Fs(1, 27) < 2.39, ps > .13.

300-500 ms (N400). Mean ERP amplitudes for morality materials yielded a Condition x Ant-Post interaction, $F(1, 27) = 5.51$, $p < .05$, $\eta_p^2 = .17$, reflecting a trend for a more negative-going ERP waveform for immoral than moral items over anterior electrodes (3.43 vs. 4.54 μV), $F(1, 27) = 3.87$, $p = .06$, but not over posterior electrodes (4.93 vs. 4.90 μV), $F(1, 27) = 0.01$, $p = .94$.

For emotion materials ERP amplitudes were more positive-going for negative than neutral items (6.33 vs. 4.17 μV), $F(1, 27) = 13.30$, $p < .001$, $\eta_p^2 = .33$, but the Condition x Ant-Post interaction was not significant, $F(1, 27) = 2.84$, $p = .10$.

500-700 ms (LPP). In this time window, the reliable Condition x Ant-Post interaction for morality materials, $F(1, 27) = 6.46$, $p < .05$, indicated a more negative-going waveform for immoral versus moral items for the anterior ROI (4.97 vs. 6.19 μV), $F(1, 27) = 4.61$, $p < .05$, but not for the posterior ROI (7.20 vs. 7.30 μV), $F = 0.03$, $p = .86$.

Finally, mean ERP amplitudes for emotion materials yielded a significant Condition x Ant-Post interaction, $F(1, 27) = 5.71$, $p < .05$, $\eta_p^2 = .17$. Further testing indicated a larger positivity for negative than neutral items for the posterior ROI (10.04 vs. 8.11 μV), $F(1, 27) = 7.26$, $p < .05$, but not for the anterior ROI (6.36 vs. 6.12 μV), $F = 0.09$, $p = .77$.

Judgment-dependent ERP results.

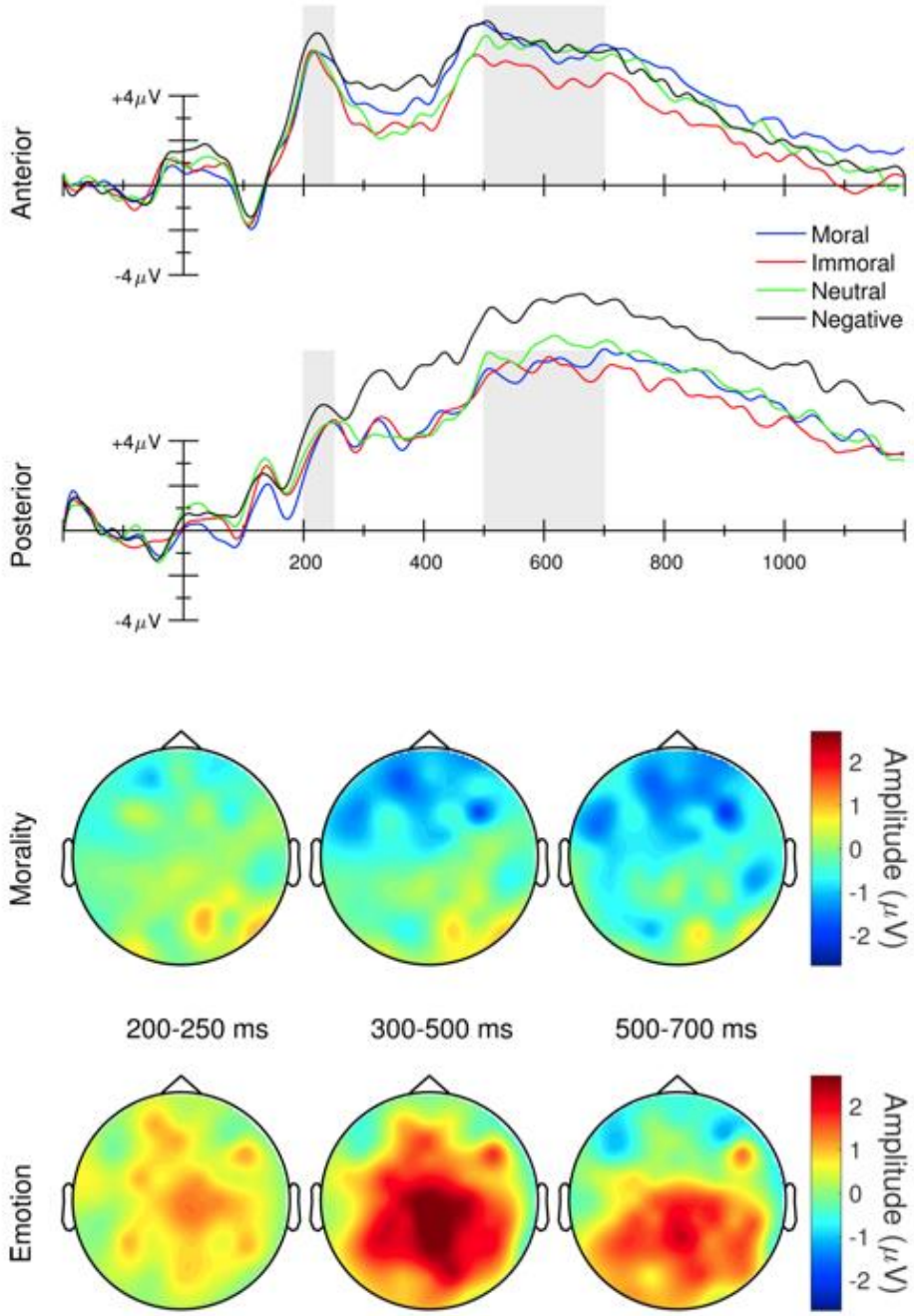
ERP waveforms averaged corresponding to the judgment-dependent analysis are shown in Figure 2. Again, for both types of materials, the analyses of mean ERP amplitudes showed main effects of ant-post (all Fs(1, 27) < 6.36, all ps < .05), indicating an anterior positivity for the early P200 time window, and a posterior positivity for the later time windows.

200-250 ms. For morality materials, the ANOVA of mean judgment-dependent ERP amplitudes with variables answer (yes vs. no) and ant-post (anterior vs. posterior) produced no significant effects, all Fs(1, 27) < 1.53, ps > .22.

The ANOVA for emotion materials showed a trend for the main effect of answer, $F(1, 27) = 3.91$, $p = .058$, $\eta_p^2 = .13$, but no interaction effect, $F < 0.01$, $p > .97$, due to a more positive-going ERP waveform for no-responses than yes-responses (6.44 vs. 5.19 μV).

300-500 ms. For morality materials, the Condition x Ant-Post interaction was significant, $F(1, 27) = 7.06$, $p < .05$, $\eta_p^2 = .21$, indicating a trend towards a more negative-

Figure 1. Upper panel: Condition-dependent grand average ERP waveforms elicited at anterior and posterior midline electrodes time-locked to the onset of the critical word for morality and emotion materials in Experiment 1. Positivity is plotted upwards. Lower panel: Spline-interpolated topographic map of mean ERP difference waveform for the 200-250 ms, 300-500 ms, and 500-700 ms time window in Experiment 1. Top panel: emotion condition (negative minus neutral). Bottom panel: morality condition (immoral minus moral).



going ERP waveform for no-responses than yes-responses for the anterior ROI (3.55 vs. 4.72 μV), $F(1, 27) = 3.55$, $p = .07$, but not for the posterior ROI (5.13 vs. 5.48 μV), $F = 0.04$, $p = .84$.

For emotion materials, the ERP positivity was larger for yes-responses than no-responses (7.63 vs. 4.17 μV), $F(1, 27) = 14.30$, $p < .001$, $\eta_p^2 = .35$. The answer effect tended to be stronger over posterior than anterior electrodes as indicated by the trend for the Condition x Ant-Post interaction, $F(1, 27) = 3.73$, $p = .064$, $\eta_p^2 = .12$ (cf. Figure 2).

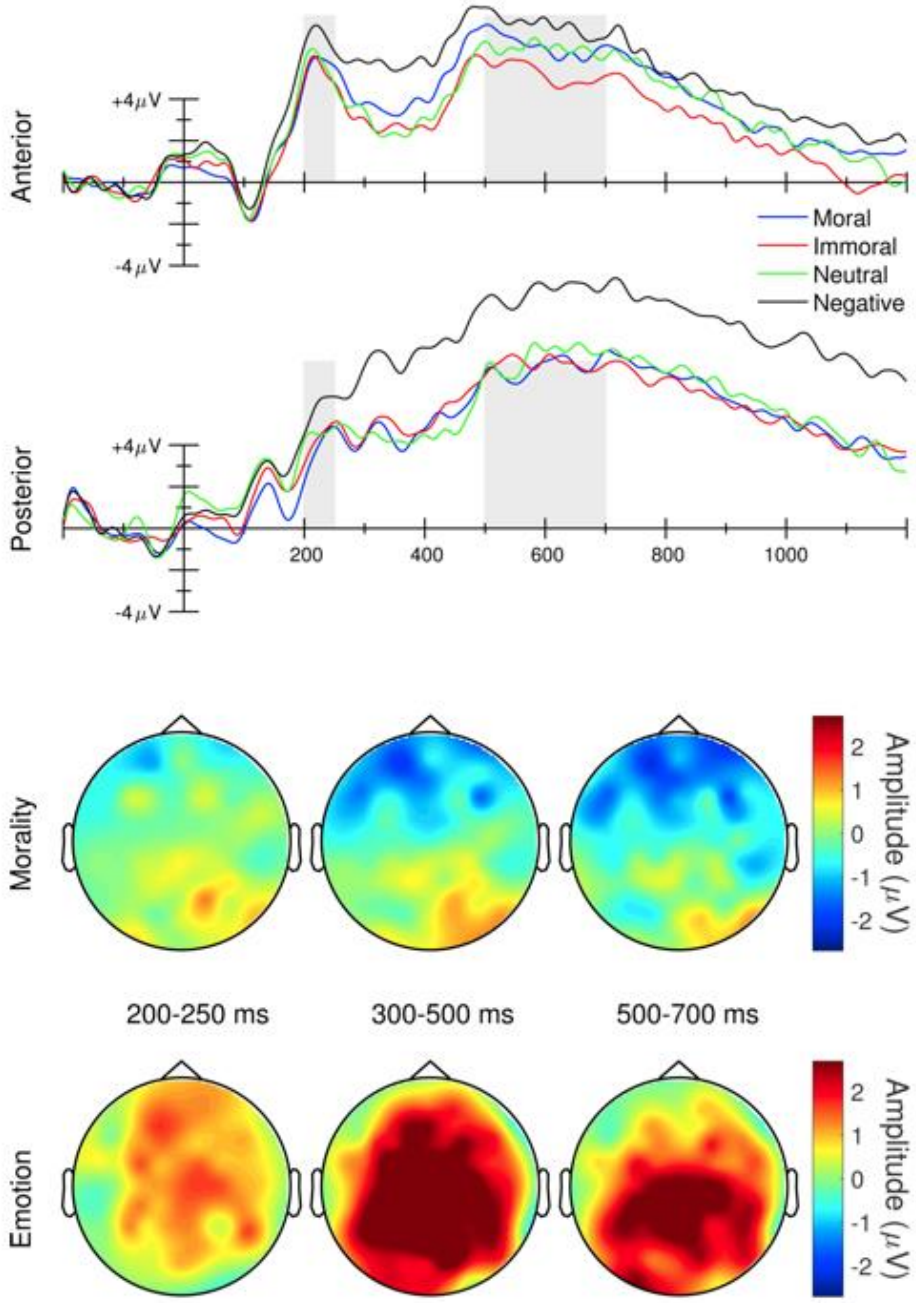
500-700 ms. For morality materials, the Condition x Ant-Post interaction was significant, $F(1, 27) = 8.32$, $p < .01$, $\eta_p^2 = .24$, indicating a trend for a more negative-going ERP waveform for no-responses than yes-responses for the anterior ROI (5.22 vs. 6.53 μV), $F(1, 27) = 3.96$, $p = .057$, but not for the posterior ROI (7.91 vs. 7.71 μV), $F(1, 27) = 0.11$, $p = .73$.

For emotion materials, the Answer x Ant-Post interaction was significant, $F(1, 27) = 6.85$, $p < .05$, $\eta_p^2 = .20$, due to a reliably larger positivity for yes-responses than no-responses for the posterior ROI (10.73 vs. 7.61 μV), $F(1, 27) = 9.19$, $p < .01$, but not for the anterior ROI (6.05 vs. 5.10 μV), $F(1, 27) = 1.44$, $p = .24$.

Discussion

In Experiment 1 participants performed different judgments depending on material type, either focusing on the moral acceptability of someone's behavior for morality scenarios or whether they were emotionally moved by the described emotion scenarios. In line with our expectations, behavioral data showed that moral items were very frequently judged as acceptable and immoral items as unacceptable, with less than 18 % of the items being judged by participants in a way that was inconsistent with the classifications that were based on the results of the pre-test (e.g., judging an item of the immoral con-

Figure 2. Upper panel: Judgment-dependent grand average ERP waveforms elicited at anterior and posterior midline electrodes time-locked to the onset of the critical word for morality and emotion materials in Experiment 1. Positivity is plotted upwards. Lower panel: Spline-interpolated topographic map of mean ERP difference waveform for the 200-250 ms, 300-500 ms, and 500-700 ms time window in Experiment 1. Top panel: emotion condition (yes minus no). Bottom panel: morality condition (yes minus no).



dition as morally acceptable). Crucially, the logistic regression analysis indicated that, in addition to the pre-classified condition variable, also valence influences morality judgments, in line with views that emotional aspects of the scenarios contribute to moral decision making (e.g., Greene et al., 2001; Haidt, 2001). For emotion items, subjective judgments of items did accord slightly less well with the pre-classified neutral vs. negative item classification (about 67%). A possible reason for this lower consistency is suggested by the logistic regression analysis results, which indicated that mainly rating-based arousal scores and to a lesser extent valence scores for each item were influencing the affective yes-no judgments in addition to the pre-classified condition variable. This is also plausible given the rating results for emotion materials, indicating that some neutral scenarios received positive valence ratings. There was also a moderately positive correlation indicating increasing arousal ratings with positive valence for these neutral items (r = .32). Finally, another reason could be that some of the sentence final emotional words were valenced, and thus yes-responses to neutral items might also reflect a word-based valence effect.

The finding of shorter RTs for immoral and negative items indicates that these items were more salient than moral and neutral items, as also suggested by the rating study results. In addition, we observed faster responses for emotional than moral materials, suggesting that moral judgments involve a more time-consuming decision process. However, based on this result alone, we cannot exclude the possibility that affect-related processes contribute to moral judgments. In summary, the behavioral data clearly indicate that participants performed the different judgment tasks appropriately.

In terms of the ERP results, a first key finding relates to the ERP analysis for emotion materials, which showed a larger posterior than anterior ERP positivity from 300-700 ms, as expected. Given its topographic distribution and time-course, and the

fact that its amplitude was larger for negative than neutral materials, we take this positivity to reflect the LPP. It is worth noting that the judgment-dependent analysis of ERP amplitudes produced the same results as the standard analysis of ERP amplitudes. Hence, we take the larger LPP to negative than neutral items to reflect an emotion effect. This inference seems justified given the pre-test results for emotion materials. That is, negative and neutral emotion items differed with respect to their valence and arousal but not regarding their linguistic features (cloze probability, semantic relatedness, critical word frequency). Similar to Van Berkum et al. (2009), we view it as unlikely that the present LPP emotion effect reflects a decision-related P300 effect; we return to this issue in the General Discussion. Together, and in line with similar reports in the literature (e.g., Delaney-Busch & Kuperberg, 2013; Holt et al., 2009), we therefore take the present LPP findings as reflecting stronger affective processing of negative than neutral items during discourse comprehension.

Crucially, and in contrast to our hypothesis that morality items undergo an implicit affective evaluation as indicated by an LPP effect (cf. Leuthold et al., 2015), no reliably enhanced ERP positivity was observed for immoral as compared to moral items. Rather, a more negative-going ERP amplitude for immoral than moral scenarios appeared from 500-700 ms (cf. Figure 1). Although the direction of this amplitude effect is in line with the centroparietal N400 effect reported by Van Berkum et al. (2009), its topographic distribution is not. That is, the present morality effect in the 500-700 ms time window showed an anterior rather than the classic centroparietal N400 distribution, it occurred later, and was also more sustained. In the General Discussion, we will evaluate possible explanations for this anterior ERP negativity effect.

In summary, the ERP findings from Experiment 1 indicate that evaluative processing of immoral items elicited a larger anterior ERP negativity than moral items,

whereas negative emotional items triggered a larger LPP than neutral items, suggesting that the different materials differ with regard to their cognitive versus affective processing. This difference in processing might be attributed to the fact that participants performed different tasks to the two types of materials. In the following experiment, we will therefore test whether the evaluation of moral content, as indicated by the anterior negativity, is task-dependent by asking participants to perform emotional judgments for morality materials as well.

Experiment 2

In Experiment 2, participants saw the same morality and emotional scenarios as in Experiment 1, but in this case judged all materials as to whether they were emotionally moved by them. We reasoned that focusing on the evaluation of the emotional content of morality materials might change their online processing in such a way that their affective analysis is prioritized (e.g., Lai et al., 2012; see also Holt et al., 2009). Like in Experiment 1, we analyzed the binary yes-no judgments using logistic regression analyses. This also allowed us to examine whether task demands influence the processing of morality materials and whether emotion judgments for both morality and emotion materials take into account the same affective item dimensions. In this case, a larger LPP should be elicited by both immoral and negative emotional items compared to moral and neutral emotional items. That is, the larger anterior negativity observed for morality materials in Experiment 1 should be absent.

Methods

Participants. Thirty right-handed native German speakers from the University of Tübingen participated for course credits or payment. Data from one participant were

excluded from the analyses due to less than 30% of trials per condition remaining after artifact rejection and from one participant due to excessive alpha activity. Because one behavioral data set was lost due to a technical problem, 27 participants entered the behavioral data analysis and 28 participants (M = 23.0 years, 20 females) contributed data to the ERP analysis.

Materials, procedure, and design. Experiment 1 was identical to Experiment 2 concerning all methodological aspects except that participants now performed yes-no responses to both moral and emotion materials with regard to the question: "Berührt Sie das Gelesene?" ["Are you emotionally moved by the presented text?"].

Data analysis. During EEG preprocessing, the number of ICA components removed for cleaning the EEG data set was M = 3.9, and the number of previously removed noisy channels that were interpolated in the ICA-cleaned EEG data set was M = 1.7 (range = 0 to 5). Following artifact rejection, the mean number of trials remaining per condition (M = 37.50 trials out of 40; range = 27-40, median = 39.0) was not reliably different across conditions, all ps > .29.

Binary yes-no judgments were analyzed using a logistic regression analysis identical to Experiment 1. Also, in addition to the standard ERP analysis, we measured ERP amplitudes in waveforms averaged for moral and neutral items that were judged as not emotionally moving and for immoral and negative items judged as moving. These judgment-dependent ERP results are reported at the end of the results section.

Results

Behavioral measures

As in Experiment 1, separate logistic regression analyses were performed for the binary emotion judgments to emotion and morality materials. For morality materials, immoral

items were more often judged as emotionally moving than moral items (71.85 vs. 44.30 %, p <.001). There were significant effects of condition (β = 0.83, *SE* = 0.22, *Wald Z* = 3.71, p < .001), valence (β = -0.15, *SE* = 0.06, *Wald Z* = -2.28, p < .05) and arousal (β = 0.94, *SE* = 0.08, *Wald Z* = 11.13, p < .001). These results suggest that the likelihood of yes-responses ("moving") increased for immoral items and for more arousing items but slightly decreased for more positively rated items.

For emotion materials, negative items were more frequently judged as moving compared to neutral items (72.69 vs. 35.13 %, p <.001). There was a significant effect of valence (β = -0.18, *SE* = 0.08, *Wald Z* = -2.17, p < .05) and arousal (β = 1.34, *SE* = 0.08, *Wald Z* = 16.77, p < .001), indicating that the likelihood of yes-responses ("moving") slightly decreased for more positively rated items but mainly increased for more arousing items.

The ANOVA performed on RT yielded no reliably faster responses to immoral than moral items (1044 vs. 1157 ms), $F(1, 26) = 2.82$, $p = .11$, $\eta_p^2 = .10$. For emotion materials, RT was faster for negative than neutral items (953 vs. 1054 ms), $F(1, 26) = 6.15$, $p < .05$, $\eta_p^2 = .19$.

Electrophysiological measures - condition-dependent ERP results.

ERP waveforms averaged according to the rating-based item classification are shown in Figure 3. For both types of materials, the waveform was characterized by an anterior P200 (200-250 ms), and a broadly distributed positivity between 300-500 ms that tended to be posteriorly distributed in the late LPP time window (500-700 ms). As before, main effects of topography will not be discussed in the following.

200-250 ms (P200). Mean ERP amplitudes for morality materials in this time interval were not reliably influenced by experimental conditions, all Fs < 0.31, ps > .58.

For emotion materials, the Condition x Ant-Post interaction was significant, $F(1, 27) = 4.81$, $p < .05$, $\eta_p^2 = .15$, indicating a larger positivity for negative than neutral items for the posterior ROI (4.32 vs. 2.75 μV), $F(1, 27) = 13.87$, $p < .001$, and as a trend for the anterior ROI (7.03 vs. 6.28 μV), $F(1, 27) = 3.33$, $p = .08$.

300-500 ms (N400). Mean ERP amplitudes for morality materials yielded a Condition x Ant-Post interaction, $F(1, 27) = 5.88$, $p < .05$, $\eta_p^2 = .18$. Further testing indicated a trend for a more positive-going ERP waveform for immoral than moral items for posterior ROIs (5.63 vs. 4.91 μV), $F(1, 27) = 3.73$, $p = .06$, but no reliable effect for anterior ROIs (4.85 vs. 4.46 μV), $F < 0.01$, $p = .98$.

For emotion materials, the significant Condition x Ant-Post interaction, $F(1, 27) = 20.17$, $p < .001$, $\eta_p^2 = .43$, indicated that the condition effect was more pronounced for the posterior ROI. However, the ERP positivity was reliably larger for negative than neutral items over both anterior (6.69 vs. 5.67 μV) and posterior midline electrodes (7.84 vs. 4.88 μV), all Fs$(1, 27) > 6.57$, $p < .05$.

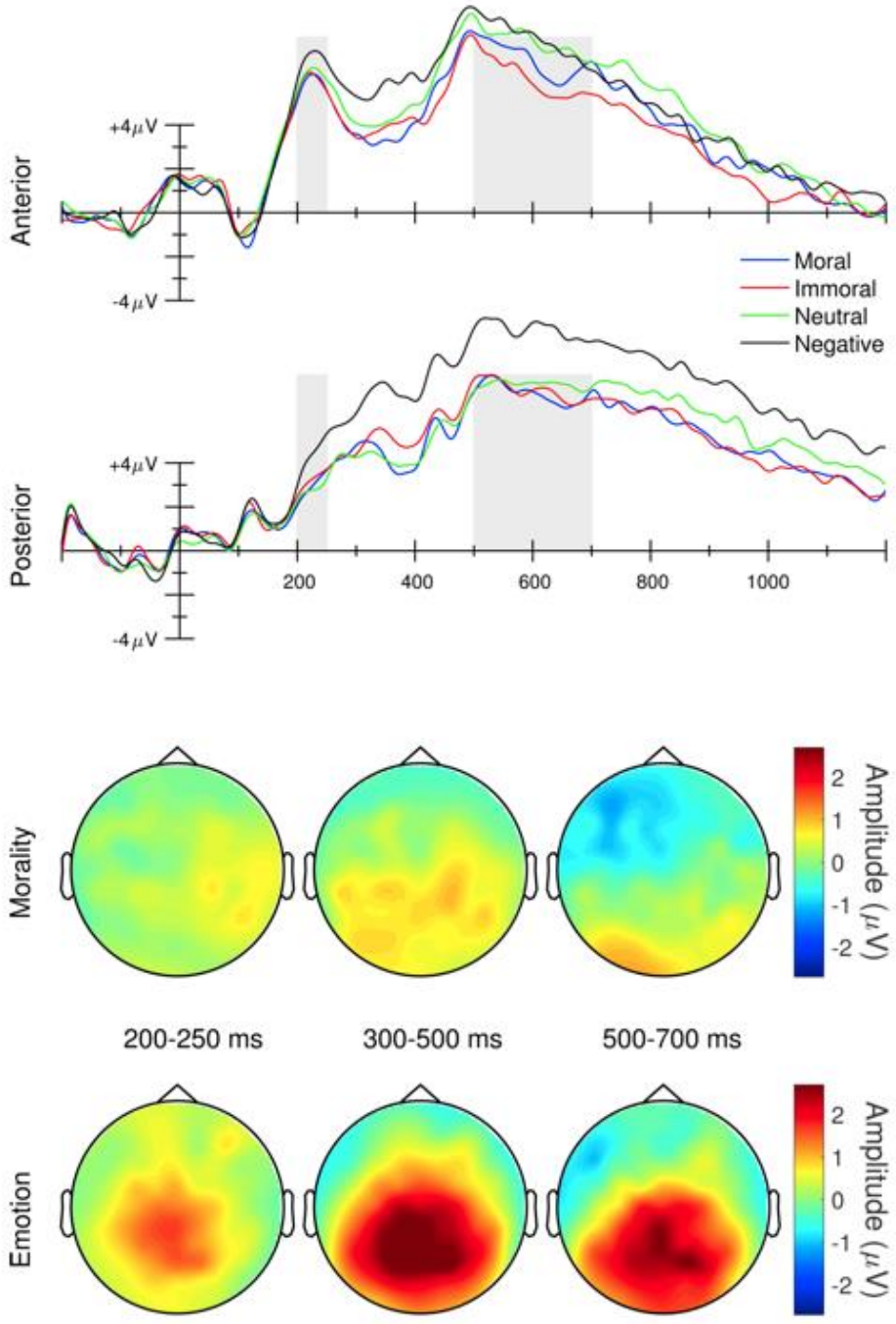
500-700 ms (LPP). For the morality materials, in this time window only the interaction of Condition x Ant-Post was significant, $F(1, 27) = 5.72$, $p < .05$, $\eta_p^2 = .17$, but further testing revealed no reliable effects, all Fs$(1, 27) < 2.93$, ps $\geq .10$.

Finally, analyses of emotion materials showed a significant Condition x Ant-Post interaction, $F(1, 27) = 26.32$, $p < .001$, $\eta_p^2 = .49$, due to a larger positivity for negative than neutral items for the posterior ROI (10.09 vs. 7.57 μV), $F(1, 27) = 35.50$, $p < .001$, but not for the anterior ROI (7.98 vs. 7.80 μV), $F(1, 27) = 0.20$, $p = .66$.

Judgment-dependent ERP results.

ERP waveforms averaged according to the judgment-dependent classification are shown in Figure 4.

Figure 3. Upper panel: Condition-dependent grand average ERP waveforms elicited at anterior and posterior midline electrodes time-locked to the onset of the critical word for morality and emotion materials in Experiment 2. Positivity is plotted upwards. Lower panel: Spline-interpolated topographic map of mean ERP difference waveform for the 200-250 ms, 300-500 ms, and 500-700 ms time window in Experiment 2. Top panel: emotion condition (negative minus neutral). Bottom panel: morality condition (immoral minus moral).



200-250 ms. The ANOVA for judgment-dependent ERPs of morality materials revealed no reliable condition effects, all *F*s(1, 27) < 1.69, *p*s > .20. For emotion materials, there was a trend for a larger positivity for yes-responses vs. no-responses (5.25 vs. 4.43 μV), $F(1, 27) = 3.09$, $p = .09$.

300-500 ms. In this time window, ERP amplitudes of morality materials yielded a larger positivity for yes-responses than no-responses (5.68 vs. 4.45 μV), $F(1, 27) = 4.66$, $p < .05$, $\eta_p^2 = .15$.

For emotion materials, the Answer x Ant-Post interaction was significant, $F(1, 27) = 11.79$, $p < .01$, $\eta_p^2 = .30$, due to an enlarged positivity for yes-responses than no-responses for the posterior ROI (7.65 vs. 4.27 μV), $F(1, 27) = 38.38$, $p < .001$, compared to the anterior ROI (6.33 vs. 4.91 μV), $F(1, 27) = 5.16$, $p < .05$.

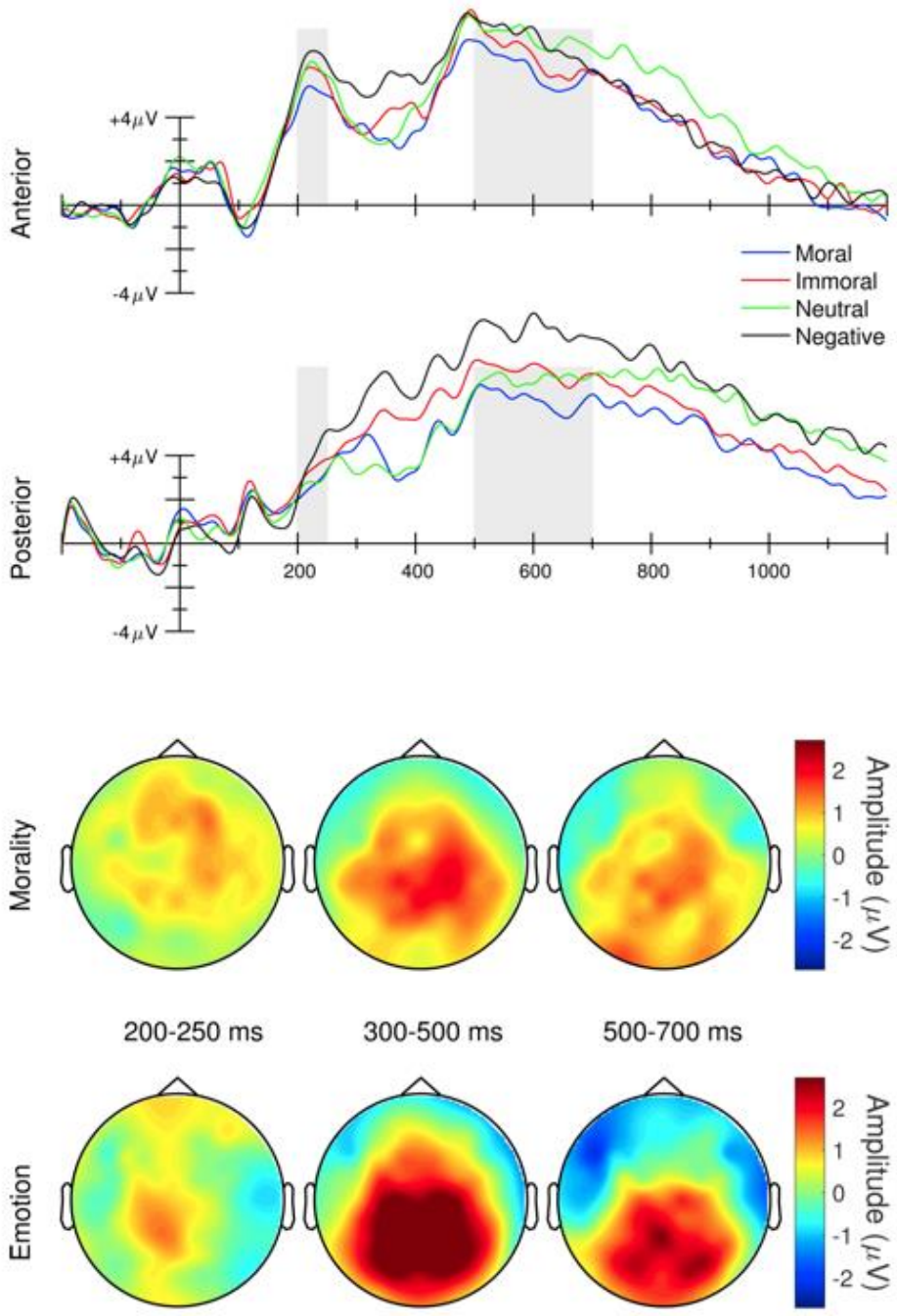
500-700 ms. ERP amplitudes in this subsequent time window were not influenced by experimental condition for morality materials, all *F*s(1, 27) < 2.35, *p*s > .13.

For emotion materials, the Answer x Ant-Post interaction was significant, $F(1, 27) = 13.75$, $p < .001$, $\eta_p^2 = .34$; further testing an enlarged positivity for yes-responses than no-responses for the posterior ROI (9.74 vs. 7.49 μV), $F(1, 27) = 7.75$, $p < .01$, but not for the anterior ROI (7.66 vs. 7.53 μV), $F = 0.04$, $p = .85$.

Discussion

The behavioral data analysis showed that as in Experiment 1, responses were faster for negative than neutral items, again lending support to the conclusion that the former items are emotionally more salient. In addition, the RT analysis indicated that moral and immoral items did not reliably differ with regard to the speed of emotion judgments, whereas they did for morality judgments in Experiment 1.

Figure 4. Upper panel: Judgment-dependent grand average ERP waveforms elicited at anterior and posterior midline electrodes time-locked to the onset of the critical word for morality and emotion materials in Experiment 2. Positivity is plotted upwards. Lower panel: Spline-interpolated topographic map of mean ERP difference waveform for the 200-250 ms, 300-500 ms, and 500-700 ms time window in Experiment 2. Top panel: emotion condition (yes minus no). Bottom panel: morality condition (yes minus no).



Importantly, response behavior differed with regard to the pre-classified item category for morality and emotion materials. Therefore, as in Experiment 1, ERP amplitudes were analyzed dependent on the pre-classified item categories (standard analysis) and dependent on the actual judgments.

First, however, it is important to note that the logistic regression analysis results for morality materials indicated that affective judgments were influenced not only by condition but mainly by rating-based arousal scores and to a smaller extent by valence scores for each item. Thus, in conjunction with the logistic regression analyses results for emotion materials, it appears that emotion judgments are strongly influenced by arousal and less so by valence. It is hence understandable that moral as compared to neutral items were judged more frequently as moving (44.30 vs. 35.13 %; $p < .001$). This finding is plausible given the fact that arousal rating results for morality materials indicated a smaller difference between moral and immoral items than neutral and negative items, which might also account for the absence of a reliable RT effect for the morality materials.

In the standard ERP analysis, replicating the LPP findings from Experiment 1, there was a larger posterior positivity for negative than neutral items from 300-700 ms, which we again take to reflect the LPP. This finding is consistent with the earlier conclusion that the LPP effect reflects the (affective) evaluation of motivationally significant stimuli (e.g., Hajcak et al, 2012). Moreover, rather than an anterior negativity (500-700 ms) as observed in Experiment 1, immoral compared to moral items tended to elicit a larger ERP positivity in the 300-500 ms time window over posterior midline electrodes. Importantly, corroborating this LPP effect, the judgment-dependent ERP analysis revealed a larger posterior positivity in the 300-500 ms time window for both immoral and negative items judged as moving as compared to moral and neutral items

judged as non-moving. Thus, LPP findings from Experiment 2 indicate that the affective evaluation of incoming linguistic information occurs not only for emotion materials but also morality materials.

General discussion

In two ERP experiments we investigated the nature and time course of evaluative processing of short morality (moral vs. immoral) and emotion (neutral vs. negative) scenarios using a discourse comprehension paradigm. Participants judged whether they found the described moral situation either morally acceptable or not and the emotional situation as moving or not (Experiment 1), or made emotional judgments to both types of scenarios (Experiment 2). Assuming that affective evaluations are triggered by both morality and emotion materials, we expected that critical words would trigger an early enhanced LPP (starting at ~300 ms) for immoral compared to moral scenarios and for negative compared to neutral emotion materials, irrespective of the judgment task. However, if performing moral judgments (Experiment 1) shifts the focus to the cognitive-semantic processing of moral content, we assumed that an N400 effect might instead be triggered by morality materials.

Crucially, we obtained behavioral evidence for the task-dependent processing of morality materials and also that the specific emotional characteristics (valence, arousal) of emotion and morality materials influenced participants' judgments. Specifically, moral acceptability judgments were slower for moral than immoral items, whereas the speed of emotional judgments did not reliably differ. The latter judgments were also performed faster than the morality judgments. On the one hand, this outcome suggests that a more complex and hence time-consuming cognitive decision process underlies moral than emotional decision making, at least for the materials used in this

study. On the other hand, it also indicates that the way readers process information about the persons and events described in the text, that is, which information they focus on and evaluate, depends on their specific goals. The additional finding of faster judgment responses to emotionally negative and immoral items than neutral and moral items in Experiment 1 might be attributed to the fact that the former items are more salient. Finally, in both experiments participants answered the respective judgment questions as expected, in the majority of cases. That is, immoral items were judged as less acceptable (Experiment 1) and more moving (Experiment 2) than moral ones and negative items were judged as emotionally more moving than neutral ones. Still, binary judgment behavior differed as compared to the pre-classified item category for both morality and for emotion materials. These findings suggest that participants adopted response criteria that did not fully accord with the rating-based classification of items. Whereas the rating study suggested that immoral compared to moral items have higher mean valence and mean arousal scores, it is clear that there is no perfect separation of moral and immoral conditions with regard to these emotion dimensions at the level of individual items as outlined earlier. Moreover, deciding whether a scenario is morally acceptable or not as well as being either moving or not might involve the processing of stimulus aspects different from those defining their moral content alone. This assumption is supported by logistic regression analysis results. That is, moral judgments were influenced by valence but not arousal, whereas emotion judgments were mainly driven by differences in arousal rather than valence for both morality and emotion materials. Together, behavioral findings indicate that participants followed task instructions and, more importantly, that processing was influenced by the task and the specific moral and emotional content of materials, which is why ERP amplitudes were also analyzed

dependent on both the pre-classified item categories (standard analysis) and dependent on the actual judgments.

A first key ERP finding concerns the larger LPP for negative than neutral emotional scenarios, starting after about 300 ms and lasting at least up to 700 ms after the onset of the critical word. It is also important to note that this LPP effect replicated across two independent experiments using the same materials and tasks. Crucially, target sentences were identical for negative and neutral items and the discourse contexts were only moderately constraining regarding the critical word. Hence, the observed LPP effects reflect a discourse-based influence, and are not the result of mere lexical differences between target words or expectancy-driven processes that would be indicated by the N400 or the P300 components. In accord with similar previous research (e.g., Fields & Kuperberg, 2012; Holt et al., 2009; see also Fischler & Bradley, 2006), we therefore take this long-lasting LPP effect to indicate the more intense affective evaluation of negative than neutral items.

In this respect, the present work extends previous ERP studies examining discourse-based emotion effects using contexts that were either strongly constraining, and hence presumably triggered emotion congruity effects as indicated by the N400 (e.g., León et al., 2010; Leuthold et al., 2012), or varied the critical words across emotion conditions (e.g., Delaney-Busch & Kuperberg, 2013; Fields & Kuperberg, 2016; Holt et al., 2009; León et al., 2010). For instance, Delaney-Busch and Kuperberg (2013) found a larger LPP (500-700 ms) to pleasant and unpleasant emotion words irrespective of the valence of the preceding emotional discourse context, whereas the N400 effect was absent. They interpreted this finding in terms of the affective primacy hypothesis (Storbeck & Clore, 2007) and proposed that for emotional contexts the affective processing of incoming information dominates over semantic processing. The present

LPP effect in conjunction with the absence of an N400 effect accords with this view, suggesting that participants focused on the processing of the emotional rather than the semantic content in the present affective judgment task (e.g., Lai et al., 2012). Together, the present ERP findings for emotion materials narrow the identified research gap concerning the investigation of emotional language comprehension by demonstrating that an LPP indicating more intense affecting processing is also observed when discourse contexts determine the emotional meaning of identical critical (emotion) words in target sentences. However, the functional interpretation of the LPP as reflecting the affective processing of linguistic input is still a matter for further research (see below).

Importantly, we reasoned that if evaluative-affective categorization (as indicated by the LPP) contributes to moral judgments, then we should also see a larger LPP for immoral than moral items, as in previous studies using a similar approach (Leuthold et al., 2015; Van Berkum et al., 2009). In fact, such an enhanced LPP for immoral compared to moral items was present over posterior electrodes from 300-500 ms in Experiment 2. However, before discussing potential implications of this ERP effect, it is important to consider the alternative possibility that it reflects an N400 effect. In this case, one would have to assume that a larger N400 to moral than immoral items overlaps with the positive-going ERP waveform, thereby producing a larger ERP positivity to immoral than moral items. For instance, Holt et al. observed in their study a larger N400 to negative and positive words compared to neutral words, but only if participants passively read for comprehension. When they evaluated the emotional content, however, this N400 amplitude modulation was obscured by the overlapping LPP. This possibility would require that moral as compared to immoral items produce a cost at the level of lexico-semantic processing or when accessing semantic memory, as

it is typically the case for incongruent items with low cloze probability or low LSA scores. Yet, the present materials used identical target sentences, which precludes the influence of word-based effects. Moreover, an N400 effect due to material differences at sentence- and discourse-level is not supported, since moral and immoral items did not differ with regard to cloze probability and LSA scores. Also, Leuthold et al. (2015) found an LPP effect and no sign of an N400 when using a passive text comprehension task for which Holt et al. (2009) found an N400 effect to context-incongruent emotion materials. Together, we view it unlikely that the present ERP effect is due to N400 component overlap and rather reflects a genuine LPP effect. Thus, it appears that participants not only judged immoral items as emotionally more moving than moral ones but also that these items underwent more intense affective processing. However, we did not find a larger LPP for immoral than moral items when moral judgments were required in Experiment 1. This finding accords with other discourse comprehension studies in that the LPP, and hence affective processing of linguistic input, is modulated by various variables, including the specific discourse context and task demands (e.g., Delaney-Busch & Kuperberg, 2013; Fields & Kuperberg, 2016; Holt et al., 2009; Xiang & Kuperberg, 2015).

Critically, the moral judgment task had an impact on online processing, as suggested by the ERP findings of Experiment 1, in which an anterior negativity (rather than the LPP), differed in amplitude across morality conditions.[5] Before discussing the

[5] When analyzing the present judgment-dependent midline ERP amplitudes for moral materials in Experiment 1 and Experiment 2 for the 300-500 ms time window, the ANOVA with the between-subjects factor Task (moral judgment vs. reading) and the repeated measurement factors condition (moral vs. immoral) and ant-post (anterior vs. posterior) revealed a significant Task x Condition interaction, $F(1, 54) = 4.67$, $p < .05$, which was also significant for the anterior ROI, $F(1, 54) = 4.52$, $p < .05$. Moreover, analysis of midline ERP amplitudes for moral materials in Experiment 1 in combination with those of Leuthold et al. (2015) for the 300-500 ms time window, the analogous ANOVA with the between-subjects factor Task (moral judgment vs. reading) and the repeated measurement factors condition and ant-post revealed a significant Task x Condition interaction, $F(1, 54) = 6.47$, $p < .05$. This was due to the fact that ERP amplitudes were more negative-going for immoral than moral items in the moral judgment task but a reverse amplitude effect in the reading task of Leuthold et al. (2015). This is

possible functional significance of this negative ERP deflection in more detail, it is helpful to first rule out a possible alternative explanation in terms of ERP component overlap. Specifically, since N400 and LPP effects are known to be similarly distributed over the scalp, there remains a possibility that simultaneously triggered LPP and N400 effects attenuate each other, with the N400 effect showing up only over anterior electrodes. However, we consider this rather unlikely for the following reasons. First, the present anterior negativity effect was more sustained than typical N400 effects. Second, there were only relatively small LPP effects for morality materials in Experiment 2, despite the fact that emotional judgments were required, which are known to increase LPP effects in comparison to a passive comprehension task (e.g., Holt et al., 2009). Third, cloze probability for critical words and the target sentence (as well as semantic similarity) was the same for moral and immoral items, thereby minimizing possible (predictive) sentence-level and word-based effects on information processing which are known to trigger a posteriorly distributed N400 effect. Finally, what mattered in our materials were the moral implications of the events being described, whereas the posterior N400 effect in Van Berkum et al.'s (2009) study was triggered by explicit moral statements that were value-incongruent rather than congruent.

We observed that the present immoral compared to moral items elicited a tentatively larger negative-going deflection over anterior electrodes from roughly 300 to 700 ms after critical word onset. Of course, since this morality effect on the anterior negativity was unexpected, it is important to replicate this ERP effect in future studies and to elaborate its potential functional interpretation. In the following, we present such a possible interpretation based on other discourse comprehension studies that also found

an interesting result since it also indicates that the ERP amplitude effect to moral items depends on the task.

an anterior negativity (Baggio, van Lambalgen, & Hagoort, 2008; Xiang & Kuperberg, 2015). In these studies, the anterior negativity was taken to index language-related working memory demands, that is, when alternative but likely text inferences have to be simultaneously maintained or integrated within the discourse or situation model (cf. Zwaan & Radvansky, 1998). Specifically, we therefore speculate that when explicit moral judgments are required, this might impact on the processing of scenarios and the updating of the discourse model in such a way that readers maintain in working memory for a short while after critical word input both the moral and immoral action (for a similar reasoning, see Xiang & Kuperberg, 2015). Put differently, it is possible that working memory load and the demands on integrating linguistic information into the discourse model is higher in the case of immoral than moral items, giving rise to the enlarged anterior negativity.

Certainly, assuming that the present anterior negativity effect might relate to working memory functions would imply that cognitive-semantic processing plays a role when explicit moral judgments are required. By contrast, when participants merely read the same moral materials for comprehension instead of performing an explicit moral judgment task (Leuthold et al., 2015), a larger LPP was elicited by immoral than moral items, which we took to reflect the affective evaluation of morality materials.[6] Thus, it is evident that discrepant ERP patterns result, indicative of cognitive (anterior negativity) and affective processing (LPP), when explicit moral judgments are required rather than when the moral content is implicitly processed. Such a task-dependent impact on moral information processing is in line with fMRI evidence indicating that cognitive processes are more dominant when the task requires explicit moral judgments

[6] We would like to note that in a current text comprehension study (in preparation) using the same materials but a passive reading task, we replicated the larger LPP in the 300-500 ms time interval for immoral than moral items and also for negative than neutral items.

than merely the passive processing of moral content and vice versa (Sevinc & Spreng, 2014).

Open issues

An open issue concerns the question of whether, and in which way, the LPP is related to the P300 component. For instance, it is known that the amplitude of the centroparietal P300 is inversely related to the prior and also the subjective probability of a given stimulus event, task demands, and the significance of stimulus input (e.g., Johnson, 1988. With regard to ERP studies using emotional discourse contexts to study person perception (Bartholow, Fabiani, Gratton, and Bettencourt, 2001; Van Duynslaeger, Van Overwalle, & Verstraeten, 2007), it is interesting to note that a larger centroparietal ERP positivity has been found to sentence-final words describing a trait-consistent ('[...] gave his wife a slap') than a trait-inconsistent behavior ('[...] gave his mother a kiss') following a short passage of text describing a person (e.g., as being hostile). Assuming that readers construct a situation model in working memory about the persons and events described in the text, in line with theories about the mental processes reflected by the P300 (cf. Donchin & Coles, 1988; Nieuwenhuis, Aston-Jones, & Cohen, 2005), one might then assume that a larger P300 (or LPP) is triggered if this model needs updating, as in the case of inconsistent language input. With regard to the impact of emotional stimulus characteristics, the more recent locus coeruleus (LC)-P300 theory (e.g., Nieuwenhuis et al., 2005) might provide an integrative framework for the interpretation of the P300 and the emotion-related LPP, since it assumes that the centroparietal positivity reflects a phasic, LC-mediated enhancement of cortical activity not only after unexpected but also after motivationally relevant and salient stimuli.

It is also an open issue whether the integration of linguistic information into the discourse or situation model is reflected by ERP negativities rather than ERP

positivities. Thus, the N400 has also been related to the demands of integrating linguistic input into a situation model (e.g., Nieuwland & Van Berkum, 2006; Filik & Leuthold, 2008, 2013). Moreover, we speculated above that the present anterior negativity might also reflect such integration demands. Together, it remains an important task to further examine the cognitive and affective processes that are more specifically reflected by various ERP components (P300, LPP, N400, and anterior negativity) typically observed in discourse comprehension studies.

Conclusions

In conclusion, the present study provides evidence for the assumption that the processing of morality scenarios depends on the specific task performed by participants. Specifically, for explicit moral judgments, immoral items elicited a larger anterior negativity than moral items, indicating the enhanced cognitive processing of moral content. By contrast, an LPP effect similar to that observed for negative compared to neutral emotional items was elicited for emotion judgments, indicating the affective categorization of incoming information during discourse comprehension. Future research would need to take into account the potential impact of task demands when elucidating the nature of the potential cognitive and affective processes contributing to moral evaluations and decisions.

References

Avramova, Y. R., & Inbar, Y. (2013). Emotion and moral judgment. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*, 169-178.

Baetens, K., Van der Cruyssen, L., Achtziger, A., Vandekerckhove, M., & Van Overwalle, F. (2011). N400 and LPP in spontaneous trait inferences. *Brain Research, 1418*, 83-92.

Baggio, G., Van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language*, *59*, 36-53.

Bargh, J. A., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*, 104-128.

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12*, 197-204.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014b). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-13.

Bell, A.J., & Sejnowski, T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*, 1129–1159.

Bohan, J., Leuthold, H., Hijikata, Y., & Sanford, A. J. (2012). The processing of good-fit semantic anomalies: An ERP investigation. *Neuropsychologia, 50*, 3174-3184.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications

for the choice of frequency estimates in German. *Experimental Psychology, 58*, 412-424.

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language, 122*, 211-226.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11*, 97-104.

Delaney-Busch, N., & Kuperberg, G. (2013). Friendly drug-dealers and terrifying puppies: Affective primacy can attenuate the N400 effect in emotional discourse contexts. *Cognitive, Affective, & Behavioral Neuroscience, 13*, 473-490.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*, 9-21.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 357-374.

Dudschig, C., Mackenzie, I. G., Strozyk, J., Kaup, B., & Leuthold, H. (2016). The sounds of sentences: Differentiating the influence of physical sound, sound imagery, and linguistically implied sounds on physical sound processing. *Cognitive, Affective, & Behavioral Neuroscience, 16*, 940-961.

Eder, A. B., Leuthold, H., Rothermund, K., & Schweinberger, S. R. (2012). Automatic response activation in sequential affective priming: An ERP study. *Social, Cognitive, and Affective Neuroscience, 7*, 436-445.

Ferretti, T. R., Singer, M., & Patterson, C. (2008). Electrophysiological evidence for the time-course of verifying text ideas. *Cognition, 108*, 881-888.

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage, 62,* 562-574.

Fields, E. C., & Kuperberg, G. R. (2016). Dynamic effects of self- relevance and task on the neural processing of emotional words in context. *Frontiers in Psychology*, 6:2003.

Filik, R., & Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology, 45*, 554-558.

Filik, R., & Leuthold, H. (2013). The role of character-based knowledge in online narrative comprehension: Evidence from eye movements and ERPs. *Brain Research, 1506*, 94-104.

Filik, R., Hunter, C. M., & Leuthold, H. (2015). When language gets emotional: Irony and the embodiment of affect in discourse. *Acta Psychologica, 156*, 114-125.

Fischler, I., Bloom, P. A., Childers, D. G., Arroyo, A. A., & Perry, N. W. (1984). Brain potentials during sentence verification: Late negativity and long-term memory strength. *Neuropsychologia, 22*, 559-568.

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in Brain Research, 156*, 185-203.

Günther, F., Dudschig, C., & Kaup, B. (2015). LSAfun - An R package for computations based on Latent Semantic Analysis. *Behavior Research Methods, 47*, 930-944.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293*, 2105-2108.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438-441.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review, 108*, 814.

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology, 35*, 129-155.

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of ERP components* (pp. 441-474). New York, NY: Oxford University Press.

Herring, D. R., Taylor, J. H., White, K. R., & Crites Jr, S. L. (2011). Electrophysiological responses to evaluative priming: the LPP is sensitive to incongruity. *Emotion, 11*, 794-806.

Holt, D. J., Lynn, S. K., & Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *Journal of Cognitive Neuroscience, 21,* 2245-2262.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*, 887-900.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434-446.

Johnson, R.Jr. (1988). The amplitude of the P300 component of the event-related potential: review and synthesis. In: Ackles, P.K., Jennings, J.R., Coles, M.G.H.,

editors. *Advances in Psychophysiology*, Vol. 3. Greenwich: JAI Press, pp. 69–138.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, 36 ECVP Abstract Supplement.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, 1146*, 23-49.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284.

Lai, V. T., Hagoort, P., & Casasanto, D. (2012). Affective primacy vs. cognitive primacy: Dissolving the debate. *Frontiers in Psychology, 3*, 243.

León, I., Díaz, J. M., de Vega, M., & Hernández, J. A. (2010). Discourse-based emotional consistency modulates early and middle components of event-related potentials. *Emotion, 10*, 863-873.

Leuthold, H., Filik, R., Mackenzie, I. G., & Murphy, K. (2012). The on-line processing of socio-emotional information in prototypical scenarios: Inferences from brain potentials. *Social, Cognitive, and Affective Neuroscience, 7*, 457-466.

Leuthold, H., Kunkel, A., Mackenzie, I. G., & Filik, R. (2015). Online processing of moral transgressions: ERP evidence for spontaneous evaluation. *Social, Cognitive, and Affective Neuroscience, 10*, 1021-1029.

Moll, J., de Oliveira-Souza, R., Bramati, I. E., & Grafman, J. (2002). Functional networks in emotional moral and nonmoral social judgments. *NeuroImage, 16*, 696-703.

Morris, J. P., Squires, N. K., Taber, C. S., & Lodge, M. (2003). Activation of political attitudes: A psychophysiological examination of the hot cognition hypothesis. *Political Psychology, 24*, 727-745.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J.D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin, 131*, 510-532.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience, 18*, 1098-1111.

Nolan, H., Whelan, R., & Reilly, R. (2010). FASTER: fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods, 192*, 152-162.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 156869.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437-442.

Sevinc, G., & Spreng, R. N. (2014). Contextual and perceptual brain processes underlying moral cognition: a quantitative meta-analysis of moral reasoning and moral emotions. *PloS One, 9*, e87427.

Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cognition and Emotion, 21*, 1212-1237.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657-671.

Van Berkum, J. J. A., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science, 20*, 1092-1099.

Van de Meerendonk, N., Kolk, H. H. J., Chwilla, D. J., & Vissers, C. Th. W. M. (2009). Monitoring in language perception. *Language and Linguistics Compass, 3*, 1211–1224.

Van Duynslaeger, M., Van Overwalle, F., & Verstraeten, E. (2007). Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Social, Cognitive, and Affective Neuroscience, 2*, 174-188.

Xiang, M., & Kuperberg, G. (2015). Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience, 30*, 648-672.

Zhang, Q., Lawson, A., Guo, C., & Jiang, Y. (2006). Electrophysiological correlates of visual affective priming. *Brain Research Bulletin, 71*, 316-323.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162-185.

3.3.2 Study 2: Kunkel, A., Mackenzie, I.G., Filik, R., & Leuthold, H. (in prep.). *Embodied affective processing of moral and emotional content during discourse comprehension: A peripheral psychophysiological study*

Authors involved:

- Prof. Dr. Hartmut Leuthold (HL)
- Dr. Ian G. Mackenzie (IGM)
- Prof. Dr. Ruth Filik (RF)
- Angelika Kunkel (AK)

List of author contributions

- Study conception and design: HL, AK, and IGM
- Material preparation: AK
- Programming: IGM
- Acquisition of data: AK and student assistant
- Analysis: IGM and AK
- Interpretation of data: HL, AK, IGM, and RF
- Writing paper: AK, HL, IGM, and RF

Embodied affective processing of moral and emotional content during discourse comprehension: A peripheral psychophysiological study

Angelika Kunkel, Ian Grant Mackenzie

Department of Psychology, University of Tübingen, Germany

Ruth Filik

School of Psychology, University of Nottingham, United Kingdom

Hartmut Leuthold

Department of Psychology, University of Tübingen, Germany

Address correspondence to:

Angelika Kunkel
Department of Psychology
Eberhard Karls University of Tübingen
Schleichstrasse 4
72076 Tübingen
Germany

Email: angelika.kunkel@uni-tuebingen.de
Phone: +49 (0)7071 29-75590

Abstract

Recently, we showed that when participants read about daily moral transgressions or emotional events a larger late positive potential (LPP) was found, which was taken to indicate the increased affective processing of immoral than moral as well as negative than neutral scenarios (Kunkel, Filik, Mackenzie, & Leuthold, 2018; Kunkel, Mackenzie, Filik, & Leuthold, under review). Using facial electromyogram (fEMG) and electrodermal activity (EDA), Kunkel et al. (under review) examined whether the same emotion and morality scenarios trigger embodied responses if participants read them for comprehension. Results showed no reliable fEMG differences in either corrugator or zygomaticus activity and also not in phasic EDA responses. Niedenthal, Winkielman, Mondillon and Vermeulen (2009) suggested that embodied processing is task-dependent and only apparent if required by an emotion-related task. Therefore, the present experiment examined whether fEMG, EDA, and the electrocardiogram (ECG) were influenced by the same scenarios if participants perform an emotional judgment task. Despite participants showing clear fEMG effects to standard emotional pictures and words, emotion materials elicited no reliable fEMG effects in all three muscles. For morality materials, fEMG corrugator, levator, and zygomaticus activity was reliable influenced by immoral compared to moral scenarios indicating negative affect. Changes in heart rate of the ECG were only differentially influenced by negative compared to neutral scenarios but not for morality materials. EDA was completely insensitive to the present stimuli. We conclude that for the present materials, affective simulation takes place only in facial muscle responses when an explicit, emotional-related task is performed.

Keywords: Moral, emotion, affective evaluation, embodiment, discourse comprehension, fEMG, EDA, ECG

Language is a powerful tool to induce emotions (cf. Sanford & Emmott, 2012). For instance, we might experience emotions when reading in a narrative about a character who is tortured or who falls in love. But which mental processes underlie these emotional responses while reading a passage of text? One answer to this question builds on the grounded cognition framework (for a review, see Barsalou, 2008), assuming that comprehending a piece of emotional text involves the simulation and reactivation of previously experienced emotions (e.g., Niedenthal, 2007). Recent text comprehension studies using materials that describe emotional events (e.g., "You and your lover embrace after a long separation.", Havas, Glenberg, & Rinck, 2007) provided behavioral evidence in accord with this view (see also Filik, Hunter, & Leuthold, 2015; Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010). However, whereas reading of short scenarios describing emotional events or daily moral transgressions elicit implicit affective evaluations as indicated by event-related brain potential (ERPs) (e.g., Leuthold, Kunkel, Mackenzie, & Filik, 2015, Kunkel, Filik, Mackenzie, & Leuthold, 2018), it remains unclear whether such materials also trigger embodied responses that are reflected by peripheral psychophysiological measures. Therefore, we will record the facial electromyogram (fEMG), the electrocardiogram (ECG), and electrodermal activity (EDA) as indices of embodied responses during emotional and moral discourse comprehension.

Previous studies suggest that affective embodiment depends on the salience and task-relevance of the emotional content (Havas et al., 2007; Niedenthal, Winkielman, Mondillon, & Vermeulen, 2009). For instance, Niedenthal et al. (2009) used fEMG recordings from different facial muscles to reveal embodied responses to isolated emotional words, building on the fact that the fEMG captures subtle facial muscle movements that are related to specific emotional expressions (cf. Dimberg, 1990; Van

Boxtel, 2010). That is, positive stimuli typically elicit stronger activation in the zygomaticus major (happiness) and negative stimuli elicit stronger activation in the corrugator supercilii (anger, fear) and the levator labii superioris (disgust) (Larsen, Norris, & Cacioppo, 2003; Niedenthal et al., 2009; Tan et al., 2012).[1] Crucially, Niedenthal and colleagues demonstrated specific embodied responses in facial EMG activity depending on the specific emotional word category (joy, anger, disgust, neutral), but only when the task required accessing the emotional word meaning.

Yet, when readers are presented vignettes that are enriched with contextual information, embodied responses might also occur in passive reading for comprehension tasks. Thus, in a study of Bartholow, Fabiani, Gratton, and Bettencourt (2001), participants were asked to form (intentional) trait inferences about a protagonist who was briefly described in the context (e.g., someone being friendly and courteous versus rude and hostile). Early corrugator activity (100-300 ms) was stronger to target sentences final emotion words that described trait-inconsistent than trait-consistent negative behavior (e.g., "Pheldar gave his wife a *slap.*"), whereas such an fEMG effect was absent for target sentences describing positive behaviors (e.g., "Pheldar gave his mother a kiss."). It appears that embodied responses are rapidly elicited when readers evaluate an emotionally salient negative behavior of another person. Thompson, Mackenzie, Leuthold, and Filik (2016) showed that the comprehension of both criticism and praise differentially influenced fEMG responses triggered by critical emotion words at the end of a literal versus ironic utterance. They presented short scenarios like "Susie texted Linda to say that she hadn't been to the gym at all that week. Linda texted her back to say: You're so motivated.". Corrugator (frowning) activity was reduced and zygomaticus (smiling) activity enhanced for ironic compared to literal criticism,

[1] In the following, we refer to theses facial muscles as zygomaticus, corrugator, and levator, respectively.

whereas reverse fEMG effects were observed for ironic compared to literal praise. Moreover, phasic EDA activity, providing an index of sympathetic arousal of the underlying sudomotor nerve activity (Dawson, Schell, & Filion, 2007), was stronger when an emoticon was present than not. Thus, negative and positive (language) inputs elicited embodied responses. However, since the studies of Bartholow and colleagues (2001) and of Thompson et al. (2016) both used emotional target words, it remains unclear whether embodied responses are also observed when the specific discourse context determines the emotional meaning of a relatively neutral critical word in the target sentence.

It is noteworthy that evidence in favor of embodied affective responses during linguistic comprehension has also been provided in studies concerned with the processing of moral transgressions. For example, Cannon, Schnall and White (2011) recorded the fEMG to short auditory statements describing either moral or immoral behavior as defined by moral foundation theory (cf. Haidt & Graham, 2007; e.g. harm-bad: "Someone tortured a stray cat" vs. harm-good: "Someone volunteered at a soup kitchen"). Participants were asked to rate item morality on a 7-point scale ("How negative or positive was this behavior?"), which in our view reflects an emotional valence rather than a morality rating. fEMG activity varied depending on the specific moral foundation (harm/care, fairness/reciprocity, authority/respect, ingroup/loyalty, purity/sanctity) and valence (good vs. bad) of the statement. Crucially, corrugator activity was found to be stronger for bad behaviors described in harm statements which the authors took to indicate anger, whereas levator activity was stronger for statements including purity and fairness transgressions, indicating disgust. In accord with the view that different moral foundations elicit specific moral emotions (Haidt & Graham, 2007), Cannon et al. concluded that facial muscle activation indicates the specific emotional

response elicited by a given statement. In accord with this conclusion, Krumhuber, Tsankova, and Kappas (2018) reported stronger levator activity during the reading of vignettes describing social-cultural norm violations as compared to neutral vignettes, whereas corrugator activity and EDA showed no reliable experimental effects. The authors took levator activity to indicate disgust when participants read about the described behavior, suggesting that norm transgressions entail an moral, and therefore, an emotional component. However, since participants were asked to rate the items for their emotionality, it is reasonable to assume that participants actively processed the emotional content of items while reading them. Together, based on the fEMG results of Cannon et al. (2011) and Krumhuber et al. (2018), it appears that embodied responses are triggered by the moral content of text materials, at least, if participants are asked to focus on the emotional meaning of the linguistic input.

Most recently, Kunkel, Mackenzie, Filik, and Leuthold (under review) took a different approach to investigate the brain processes and embodied responses elicited by daily moral transgressions as well as emotional events. They presented prototypical morality and emotion scenarios that introduced a protagonist situated in either immoral or moral and negative or neutral circumstances. Crucially, only the moderately constraining contexts varied between the respective material-specific conditions, whereas the word-by-word presented target sentences were identical. Here, the critical word disambiguated the situation in terms of an either moral versus immoral behavior or a neutral (or mildly positive) versus negative event. This procedure allowed Kunkel et al. to control for word-based effects that might have played a role in previous studies reviewed above, while examining in separate experiments the affective processing of moral and emotional content within the brain and the body's periphery. In the first experiment, the ERP was characterized by a larger late posterior positivity (LPP) to

immoral than moral as well as to negative than neutral scenarios. This replicated similar LPP effects when participants explicitly judged the emotional content as inferred from the discourse context (Kunkel et al., 2018) and when they merely read the respective moral scenarios (Leuthold et al., 2015). As in these previous studies, Kunkel and colleagues (under review) took the LPP effect to indicate the increased affective processing of incoming moral and emotional linguistic information in discourse comprehension. In the second experiment of Kunkel et al. (2018), fEMG and EDA were recorded while participants read the same morality and emotion scenarios (experimental study) and also in a subsequent control study presenting emotional pictures and emotional words. In the control study, negative compared to neutral pictures and words elicited stronger corrugator activity, indicating an embodied facial response to these emotional materials. However, negative versus neutral and also immoral versus moral items did not differentially influence fEMG activity during discourse comprehension. In contrast to the fEMG, EDA was generally insensitive to the emotional content of stimuli. Together, Kunkel et al. (under review) found no evidence for affective simulation in fEMG activity and EDA when participants merely read for comprehension and when the same relatively neutral target words were employed in morality and emotion materials, respectively.

Given that embodied responses during comprehension were observed when some kind of emotion judgment was required to materials (Cannon et al., 2011; Krumhuber et al., 2018; Niedenthal et al., 2009), the aim of the present study was to investigate whether embodied responses are elicited by the emotional and moral materials of Kunkel et al. (under review) during discourse comprehension when participants perform an emotion-related judgement task. In addition to the fEMG and EDA, the ECG was recorded to identify potential changes in heart rate (HR) elicited by

scenarios describing everyday moral transgressions and situations with negative emotional content as compared to moral and neutral (or mildly positive) scenarios. HR to arousing emotional pictures is initially decelerating, indicating orienting and attention towards the stimulus, and then accelerating, reflecting preparation for action (e.g., Bradley, Codispoti, Cuthbert, & Lang, 2001). Moreover, Weis and Herbert (2017) observed in the first four seconds after the onset of a pronoun-noun phrase (e.g., "Meine Freude" ["my happiness"]) an increased HR for emotional than neutral phrases. In order to check the sensitivity of the psychophysiological measures, the discourse comprehension task was followed by the control experiment of Kunkel et al. (under review), in which participants judged the emotional picture and word stimuli.

Generally, we hypothesized that the emotional judgment task will increase the emotional salience of incoming linguistic input for both types of materials. The affective processing of these materials might hence also result in differential peripheral body responses. More specifically, we predict that the higher arousal value of immoral compared to moral and negative compared to neutral scenarios will be reflected in stronger phasic EDA activity. Moreover, since negative and immoral items are characterized by their negative valence, fEMG activity in the corrugator and levator muscles should be larger for immoral compared to moral and for negative compared to neutral scenarios, whereas fEMG zygomaticus activity should show a weaker or no differential valence effect. Finally, regarding the ECG, we predicted an HR decrease for immoral compared to moral and for negative compared to neutral scenarios.

Method

Participants. 48 native German speakers from the University of Tübingen received course credits or payment for participating (M_{age} = 23.33 years, SD_{age} = 3.98 years, 42 females, 46 right-handed).

Materials and design. For the discourse comprehension experiment, materials were taken from Kunkel et al. (2018) and consisted of 80 items each for the morality and for the emotion scenarios (see Table 1 for examples; the full set is available from the first author). All scenarios consisted of two parts. The first part consisted of two or three sentences describing the context, and the second part was the target sentence containing the critical word. The same target sentence was used for moral and immoral materials, and the same held true for neutral and negative emotional materials in order to eliminate possible sentence-level and word-based effects (with the context varying across conditions, see Table 1). Moral materials described actions that would be perceived as either moral or immoral, whereas emotional materials would describe a neutral or a negative event. In addition, 40 filler items were used which contained no inconsistencies and were similar in length to the experimental items. See method section of Kunkel et al, (2018) for relevant item characteristics and rating data.

Presentation and randomization of items and conditions across participants was identical to that of Kunkel et al. (2018). For two consecutive participants, two lists were randomly generated such that each morality scenario appeared across the two lists either in the moral or the immoral condition, and each emotion scenario appeared either in the neutral or the negative condition. Thus, for each participant, the 200-item list consisted of 40 moral and 40 immoral items, 40 neutral and 40 negative items, as well as 40 neutral filler items. All items were randomly presented.

For the control experiment, materials were taken from Experiment 2 of Kunkel et al. (under review) consisted of 40 items each for pictures and words, which were matched for arousal and valence. See method section for Experiment 2 of Kunkel et al. (under review) for relevant item characteristics and rating data. Presentation and randomization was identical to Kunkel et al. (under review).

Table 1. Example for moral materials with context for moral and for immoral items, for emotional materials with context for neutral and for negative Items, as well as the respective target sentences containing the critical word (in italics).

Moral Materials	Moral	Herr Zimmermann arbeitet in einem Chemiekonzern. Seit einigen Wochen gibt es Sicherheitsprobleme aufgrund fahrlässigen Verhaltens von Kollegen, sodass Menschen-leben gefährdet sind. (Mr. Zimmermann works at a chemical company. For several weeks, there have been safety issues caused by colleagues' careless behavior, putting human lives at risk.)
	Immoral	Herr Zimmermann arbeitet in einem Chemiekonzern als Sachbearbeiter in der Personalabteilung. Er ist ehrgeizig und will rascher aufsteigen als seine Kollegen. (Mr. Zimmermann works in human resources of a chemical company. He is ambitious and aims at being promoted faster than his colleagues.)
	Target Sentence	Im monatlichen Gespräch berichtet er seinem Chef deren *Fehler*. (In a monthly commitment talk he points out to his boss their *mistakes*.)
Emotion Materials	Neutral	Sarah trainierte den ganzen Nachmittag im Fitness-Studio und geht zum Abschluss in die Sauna, wonach sie sich wohlig gut fühlt. (Sarah had been at the gym all afternoon before completing her work out with a visit to the sauna, after which she feels pleasantly well.)
	Negative	Sarah hat einen sehr anstrengenden Job. Sie muss dafür seit mehreren Monaten jeden Tag um 5 Uhr aufstehen und den ganzen Tag hart arbeiten. (Sarah has a very demanding job. For several months she has been getting up at 5 am and working very hard all day.)
	Target Sentence	Sie ist sehr *erschöpft*. (She is very *exhausted*.)

Following the presentation of the final word a question was displayed on the screen: “Berührt dich das Gelesene?” [“Are you emotionally moved by the text?”]. Participants were asked to rate on a scale from 1 to 8: 1 = *nicht berührt* [*not emotionally moved*] to 8 = *stark berührt* [*strongly emotionally moved*].

Procedure. After electrode application, participants were seated in an electrically shielded booth in front of a 21-in. computer monitor (100 Hz) at a viewing distance of 65 cm (maintained by a chin rest). Experimental materials (scenarios, words, pictures) were presented at the center of the screen in white 16-point Helvetica font on a black background using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) running under MATLAB (2018a), on Ubuntu 16.04 LTS. Participants were instructed to avoid any head movements and to maintain fixation at the center of the screen during word presentation. Furthermore, they were instructed to read and watch attentively, and to answer the question by pressing the appropriate response key.

The experiment started with the instructions followed by a practice block containing three trials. Experimental trials were presented in a total of 20 blocks of 10 items each, separated by a short break. Both instructions and breaks were controlled in their duration by the participants pressing the space bar. For a trial, the context was displayed for a minimum duration of 1,500 ms. When participants had read the context sentences, they initiated the word-by-word presentation of the target sentence by pressing the space bar, which started with the presentation of a fixation point for 1000 ms. Then, each word was displayed centrally for 300 ms, with a 200-ms blank interval between successive word presentations. The question was asked immediately after presenting the final word, and was displayed until participants pressed a key to indicate their arousal rating. Keys ranged from one through eight on the computer keyboard.

Participants were instructed in writing on the screen about the control experiment, which contained eight blocks of 10 experimental trials each, separated by a self-paced break controlled in its duration by the participant. Pictures and words were randomly presented within a block of trials. Every stimulus was displayed for 500 ms with a variable inter stimulus interval of six to eight seconds. The question was asked immediately after presenting the picture or word.

Psychophysiological measures. For fEMG activity, three pairs of surface electrodes were positioned on the zygomaticus major (smiling, cheek), corrugator supercilii (frowning, inside brow), and levator labii superioris (wrinkling/ raising top lip, nose), respectively. Surface electrodes were placed on the left side of the face following Fridlund and Cacioppo (1986). Surface electrodes for ECG activity were placed on the left side of the hip and under the right collarbone. EDA electrodes were placed on the distal phalanges of the middle and index finger of the non-dominant hand. All psychophysiological measurements were recorded continuously from pairs of Ag-AgCl electrodes using a BIOSEMI Active-Two amplifier system with a sampling rate of 1024 Hz.

Data analysis. An emotional judgement on an eight-point scale was required in the present experiment. As in Kunkel et al. (2018), it is again conceivable that participants may apply decision criteria that lead, at least sometimes, to judgements that are inconsistent with the morality and emotion classification of the materials. That is, some items preclassified as moral might be judged as emotionally moving, or some items preclassified as negative might be judged as not emotional moving, and vice versa. Therefore, to check whether participants' responses were consistent to preclassifications, the response scale was divided into binary responses with ratings

ranging from 1 to 4 as *not/less emotionally moving*, and responses ranging from 5 to 8 as *emotionally moving.*

Participants' response behavior was analyzed by using linear mixed effects modeling, implemented by the lmer function within the lme4 R package (Bates, Maechler, Bolker, & Walker, 2014). We fitted a model with condition as fixed effect and with random intercepts for both participants and items. We obtained *p*-values by likelihood ratio tests comparing the model with and without the fixed effect term of interest. Separate lmer models were performed for morality and emotion materials to determine the impact of condition for participants' judgements.

All fEMG analyses were performed identical to Kunkel et al. (under review) using code adapted from FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). fEMG pre-processing involved applying a band-pass filter of 20-500 Hz to the continuous data as recommended by Van Boxtel (2010). fEMG data were epoched around the onset of the critical word (-2 to 8 seconds). The differential activity between each electrode at the zygomaticus major, corrugator supercilia, and levator labii superioris was calculated separately. This differential activity at each muscle group was then rectified. Artifacts were defined as trials containing fEMG activity exceeding 250 μV within the 10 second epoch. Artifacts (< 1% of trials) were removed. Artifact-free trials were aligned to a 250-ms baseline prior to the onset of the critical word. The analysis epoch consisted of eight sequential time-windows of 250 ms duration beginning at critical word onset. EMG activity was expressed as a percentage of baseline activity. An additional outlier exclusion procedure involved converting the percentage change from baseline to a z-standardized score, separately for each participant, condition, muscle group, and time-window. Observations with a z-score value greater than 3 were

removed. This procedure removed approximately 1.5 % of data points from the subsequent analyses.

All EDA analyses were performed using the Ledalab MATLAB toolbox (Benedek & Kaernbach, 2010, available from http://www.ledalab.de). The continuous EDA signal was downsampled to 32 Hz and a low-pass filter (5 Hz) was applied. The resulting signal was decomposed by continuous decomposition analysis (CDA; Benedek & Kaernbach, 2010). This method extracts the phasic information underlying the skin conductance response and aims at retrieving the signal characteristics of the underlying sudomotor nerve activity. Integrated phasic EDA response within an epoch 1-8 seconds after the critical word were analyzed. A minimum amplitude threshold of 0.05 μS was applied for the detection of EDA responses.

For the ECG, changes in HR were measured second-by-second in beats per minute (BPM) and were calculated as a change from baseline. The continuous ECG signal was epoched around the onset of the critical word (-1 to 8 seconds) and consisted of eight sequential time-windows of one second duration beginning at critical word onset. Baseline was set from -1 to 0 before the onset of the critical word. All fEMG, ECG, and EDA analyses were performed using custom MATLAB scripts.

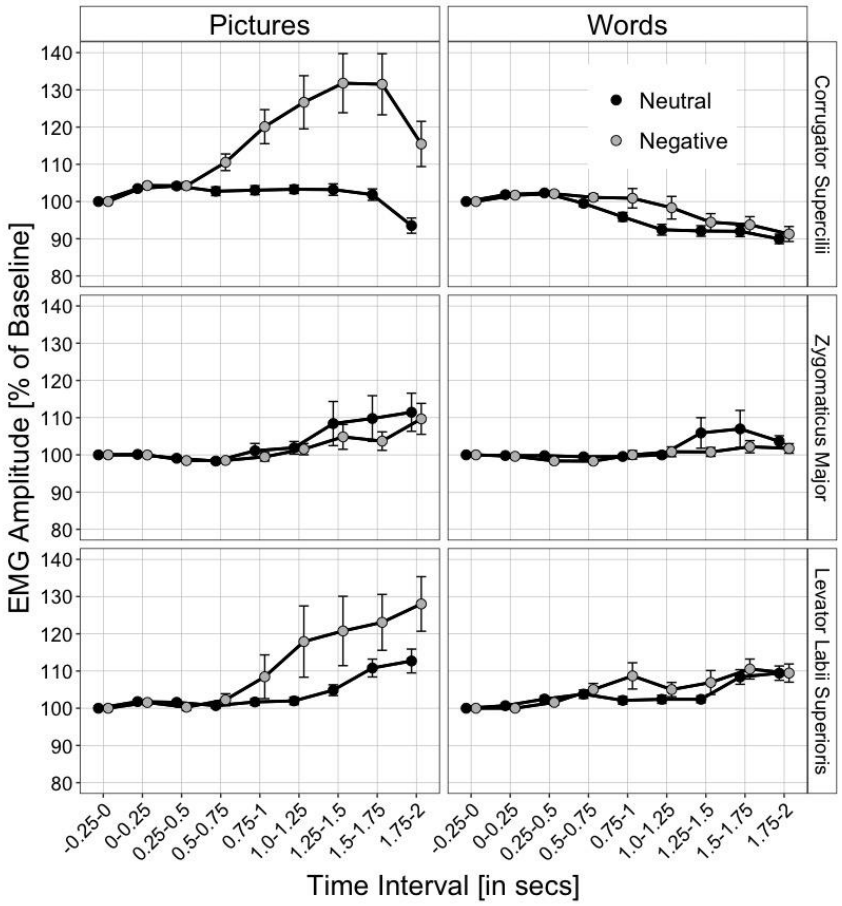
We performed separate linear mixed effect modelling for fEMG, ECG and EDA responses with time window and condition as fixed effects. As random effects, we included intercepts for both participants and items. Again, significance was tested by using the Likelihood Ratio Test. The models are described in more detail in the respective results sections. For all analyses the significance level was set to alpha = .05.

Results

Behavioral measures. Immoral scenarios were judged as more emotionally moving than moral scenarios (4.62 vs. 3.32), and negative scenarios were judged as more emotionally moving than neutral scenarios (4.50 vs. 2.90). Furthermore, with regard to a binary response categorization, immoral items were more often rated as emotionally moving than were moral items (53.70% vs. 28.43%), and negative items were more often rated as being emotional moving than were neutral items (52.08% vs. 20.57%). For both materials, LME analysis corroborated these results given that the model including condition as fixed effect (condition + (1 | participants) + (1 | item)) fitted the data better than the model without condition as fixed effect, χ^2s(1) > 65.46, *p*s < .001.

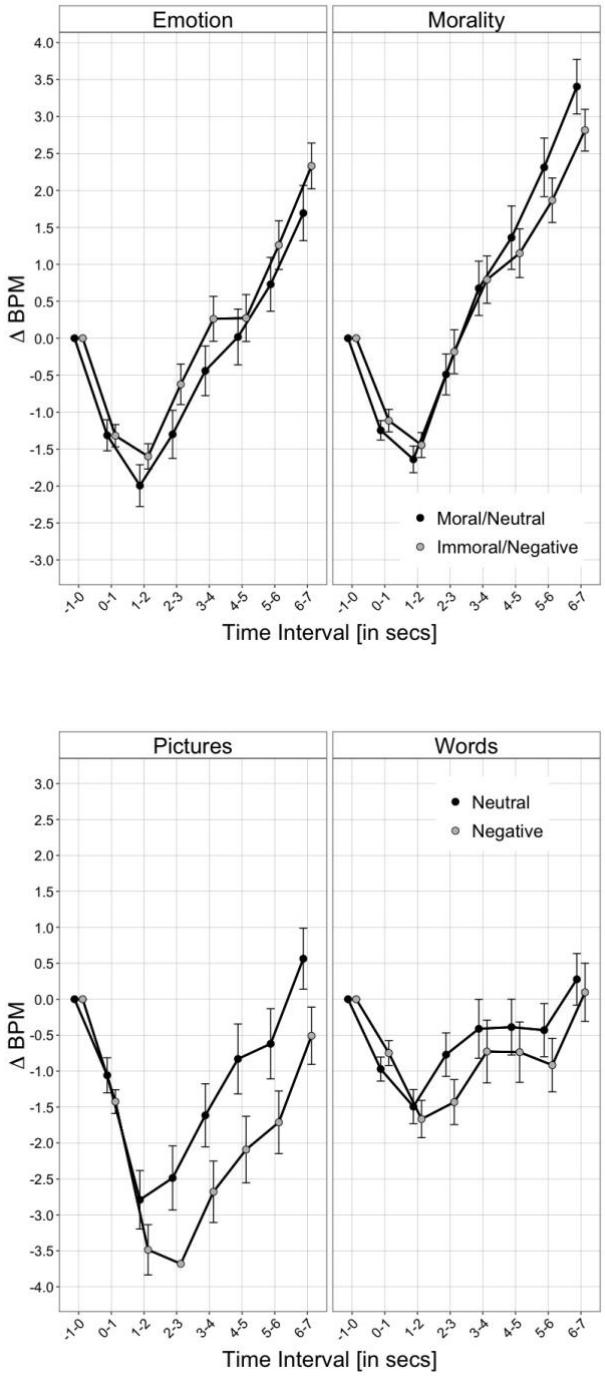
Control experiment. Mean fEMG activity was analyzed separately for pictures and words and for corrugator, zygomaticus, and levator muscle, respectively (cf. Figure 1). The LME models included condition (neutral vs. negative) and time window (eight levels: 0-250, 250-500, …, 1750-2000 ms) as fixed effects and participants and items as random effects. For words and pictures, corrugator and levator activity fitted best to the LME model including the Condition × Time window + (1 | participants) + (1 | item), all χ^2s(8) > 23.29, *p*s < .01, indicating an increased fEMG activity for negative compared to neutral items. For words and pictures, zygomaticus activity was not significantly influenced by condition, all χ^2s(1) < 1.98, *p*s > .16. For zygomaticus activity, the model including time window (Time window + (1 | participants) + (1 | item) fitted the data better than the null model, all χ^2s(7) > 36.46, *p*s < .001, indicating a continuous increase of zygomaticus activity for both conditions.

Figure 1. fEMG corrugator, levator, and zygomaticus response in the control experiment as a function of material (picture vs. word), condition (neutral vs. negative), and time window. Error bars reflect standard error of the mean.



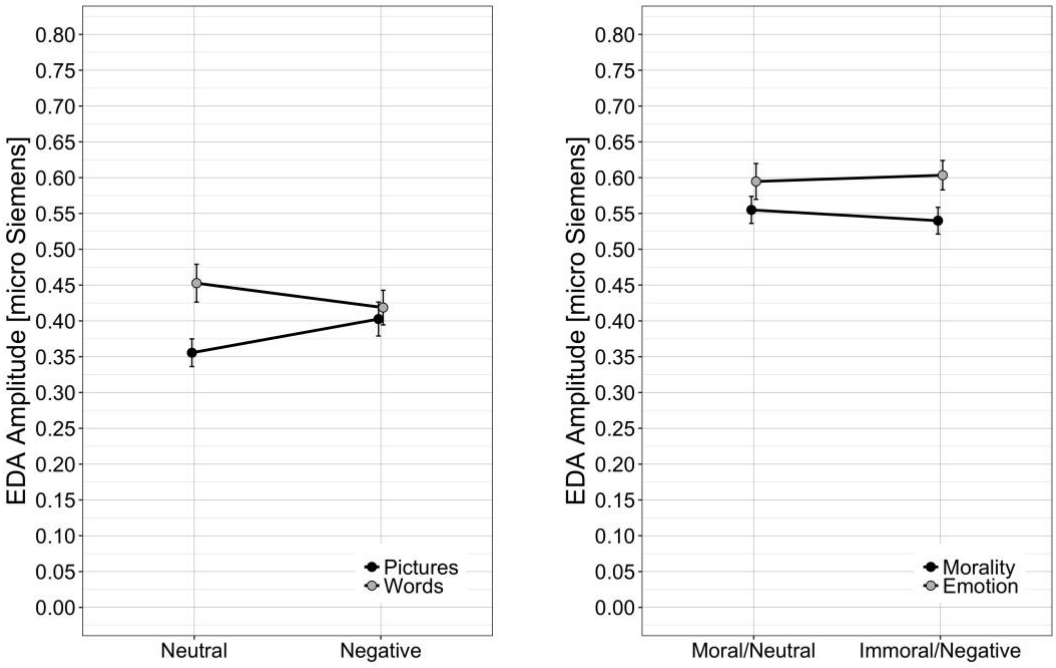
The statistical analysis of HR was calculated separately for pictures and words (cf. Figure 2). The full LME included condition (neutral vs. negative) and time window (eight levels: 0-1, 1-2, …, 6-7 sec) as fixed effects and participants and items as random effects. For words, the LME model including condition and time window as fixed effects (Condition + Time window + (1 | participants) + (1 | item) did not fit better than the model without condition, $\chi^2(1) = 1.47$, $p = .22$, indicating that HR was not sensitive to the emotional word condition. By contrast, for pictures the LME model including condition fitted the data better than that without condition, $\chi^2(1) > 8.04$, $p < .01$, indicating a stronger HR decrease for negative compared to neutral pictures.

Figure 2. ECG response in the control and discourse comprehension experiment as changes of BPM as a function of material (picture vs. word, emotion vs. morality), condition (neutral vs. negative, moral vs. immoral), and time window. Error bars reflect standard error of the mean.



For integrated phasic EDA responses, the LME model including condition as fixed effect (condition + (1 | participants) + (1 | item) as compared to the model without condition showed a trend for pictures, $\chi^2(1) = 2.78$, *p*s = .10, due to a higher EDA amplitude for negative compared to neutral pictures (Figure 3). There was no significant influence of condition for words, $\chi^2(1) = 1.11$, $p = .29$.

Figure 3. EDA amplitudes in μS for the control and discourse comprehension experiment as a function of material (picture vs. word, emotion vs. morality) and condition (neutral vs. negative, moral vs. immoral). Error bars reflect standard error of the mean.



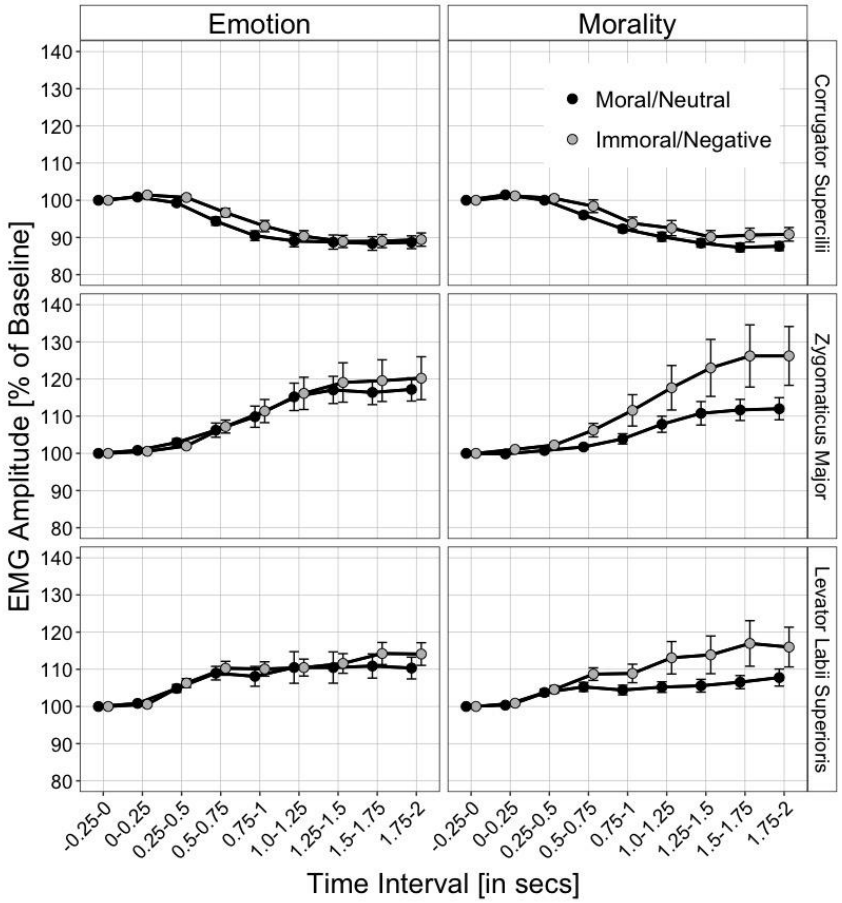
Discourse comprehension experiment. Mean fEMG activity was analyzed separately for morality and emotion materials and for corrugator, levator, and zygomaticus muscle, respectively (cf. Figure 4). The full LME included condition (moral vs. immoral, neutral vs. negative) and time window (eight levels: 0-250, 250-

500, …, 1750-2000 ms) as fixed effects and participants and items as random effects. For morality materials and for all three muscles, fEMG activity fitted best to the LME model (Condition × Time window + (1 | participants) + (1 | item)) compared to the model without condition as fixed effect, all χ^2s(8) > 19.25, $p < .05$, indicating an increased activation for immoral compared to moral scenarios for all three muscles. For emotion materials, the LME model including condition and time window as fixed effect (Condition + Time window + (1 | participants) + (1 | item)) did not fit better than the model without condition, all χ^2s < 2.11, ps > .15, indicating that fEMG was not sensitive to condition. For all three muscles, the model including Time window + (1 | participants) + (1 | item) fitted best, all χ^2s(6) > 188.97, $p < .001$, indicating a continuous increase of levator and zygomaticus activity and a continuous decrease of corrugator activity.

The statistical analysis of HR was calculated separately for morality and emotion materials (cf. Figure 2). The full LME model included condition (neutral vs. negative, moral vs. immoral) and time window (eight levels: 0-1, 1-2, …, 6-7 sec) as fixed effects and participants and items as random effects. For emotion materials, changes of HR fitted better to the LME model including condition and time window as fixed effects (Condition + Time window + (1 | participants) + (1 | item)) compared to the model without condition, $\chi^2(1) = 4.14$, $p < .05$, indicating that the HR decrease was stronger for neutral compared to negative sentences. For morality materials, the LME model (Condition × Time window + (1 | participants) + (1 | item)) compared to the model without condition as fixed effect indicated that HR produced a trend, $\chi^2(7) = 13.98$, $p = .05$, indicating that moral scenarios had a stronger initial decrease followed by a stronger increase of HR.

For integrated phasic EDA responses, the LME analysis showed for both morality and emotion materials no condition effect, all χ^2s < 0.38, ps > .53.

Figure 4. fEMG corrugator, levator, and zygomaticus response in the discourse comprehension experiment as a function of material (emotion vs. morality), condition (neutral vs. negative, moral vs. immoral), and time window. Error bars reflect standard error of the mean.



Discussion

The aim of the present study was to examine embodied responses during discourse comprehension of incoming emotional information when participants perform an emotional judgement task. Participants were asked to rate whether they were emotionally moved by morality and emotion scenarios. We hypothesized that if emotional text comprehension involves affective simulations (e.g., Niedenthal, 2007), this should be indicated by increased fEMG activity, changes in heart rate as measured by ECG, or phasic EDA responses to both negative compared to neutral and immoral

compared to moral scenarios. Additionally, in order to assess the sensitivity of fEMG, ECG, and EDA measures, participants performed a control experiment with picture and word stimuli that typically evoke differential responses in these physiological measures. In the following, the results of the control experiment will be discussed first.

EDA was not significantly sensitive to our manipulations of the control experiment as there was only a trend for an increased EDA amplitude for negative compared to neutral pictures. We hence conclude that even when an affective judgement is demanded, EDA is relatively insensitive in revealing an activation within the autonomic nervous system as elicited by emotional pictures and words. In line with the literature, ECG showed a stronger initial decrease of heart rate for negative than neutral pictures, indicating a stronger, presumably automatically mediated, orienting response after emotional stimuli (Bradley & Lang, 2000). Already during this initial decrease, stimulus valence differentially influenced ECG activity, that is, negative stimuli produced a stronger HR decrease than neutral picture stimuli. Accordingly, this result provides evidence for the view that particularly negative pictures produce a stronger embodied response, reflecting the more aversive nature of these stimuli. This was not the case for words, presumably because negative words are less arousing than negative pictures (e.g., Weis & Herbert, 2017). Nevertheless, the ECG response appears to be sufficiently sensitive to reveal embodied responses to relatively strong emotional stimuli.

With regard to embodied facial responses, the control experiment indicated that fEMG corrugator and levator activities were both stronger to negative than neutral items, hence effectively indicating an embodied repose in terms of negative affect (e.g., Larsen et al., 2003). In contrast, negative versus neutral pictures and words did not differentially influence fEMG zygomaticus activity. This particular result corresponds to those of Tan

et al. (2012) and of our previous experiment (e.g., Kunkel et al., under review), suggesting that zygomaticus activity is not sensitive to negative affect. Together, the control experiment suggests that particular fEMG corrugator and levator activity are reliable markers of embodied responses to negative emotional stimuli, whereas EDA and ECG measures appear less sensitive.

This impression was also supported by the results of the discourse comprehension experiment. Thus, the zero-effects in the analysis of EDA provide no evidence for the occurrence of autonomic arousal to emotional stimuli that differ in their arousal value during emotional discourse comprehension. Moreover, whereas the analysis of the ECG indicated an initial decrease of HR followed by a later increase for both materials, it was only for emotion materials that the change in HR was influenced by the different scenarios. That is, HR decreased less strongly initially and was hence generally higher for negative compared to neutral scenarios. This result corresponds to that of Weis and Herbert (2017), who observed a stronger HR increase to emotional than neutral phrases. Like these researchers, we assume that the HR change for emotional materials reflects their higher emotional arousal in anticipation of action (see also Bradley et al., 2001). The absence of a differential HR response to immoral versus moral items might be due to their lower arousal value, as previously found in the rating study of Kunkel et al. (2018). In contrast to HR, fEMG was nod differentially influenced by negative versus neutral items during emotional discourse comprehension. This zero-effect is unexpected and would need to be replicated in future studies before being interpreted as indicating the absence of a valence-based embodiment effect.

Crucially, fEMG revealed evidence in favor of embodied responses to relatively neutral target words during the comprehension of moral content. That is, as expected, levator and corrugator activities were sensitive to moral transgressions, showing

stronger activity for immoral compared to moral scenarios. More specifically, we take these fEMG effects to indicate that moral transgressions elicit negative affect. These results extend those of previous studies concerned with moral decision making (Cannon et al., 2011; Krumhuber et al., 2018). Cannon and colleagues (2011) used short audible statements describing either moral or immoral behavior and Krumhuber and colleagues presented text vignettes describing social-cultural norm-consistent or inconsistent behavior. However, these materials differed across conditions with regard to word- and sentence-based characteristics. In contrast, the present materials minimize word- and sentence-based influences on the affective processing of moral content. Hence, our fEMG findings provide more clear-cut evidence in favor of embodied affective responses during discourse comprehension of moral scenarios.

However, and in contrast to our expectations, zygomaticus fEMG activity was stronger for immoral compared to moral items. Since increased zygomaticus activity is usually taken to indicate positive affect and moral transgressions are assumed to elicit negative affect, as supported by rating study results (Kunkel et al., 2018) and present levator and corrugator findings, one may wonder how this specific fEMG result can be interpreted. First, it is important to note that in the case of isolated emotional word stimuli, zygomaticus activity was found to be a largely insensitive indicator of negative valence (e.g., Larsen et al., 2003; Tan et al., 2012). With this in mind, one possibility is that the present larger zygomaticus activity for immoral than moral items reflects cross-talk from muscles that are associated with negative effect (e.g., Unz & Schwab, 2005; Tassinary, Cacioppo, & Vanman, 2007). In support of this idea, we found present levator and zygomaticus activity to be positively correlated (r = 0.53). Moreover, and in line with this assumption, Lang and colleagues (Lang, Greenwald, Bradley, & Hamm, 1993) found zygomaticus activity to picture stimuli rated as disgusting. Also, it has been

suggested that when more complex scenes are described, zygomaticus activity might be interpreted as reflecting a wry, sarcastic, or smirking expression (‘t Hart, Struiksma, Van Boxtel, & Van Berkum, 2018). Hence, it appears difficult to unambiguously infer a specific valence from zygomaticus activity. Together with the clear-cut fEMG results from the control experiment, we therefore suggest that the simultaneously increased activation of levator, corrugator, and zygomaticus for immoral compared to moral scenarios indicates an embodied response due to the negative valence of immoral items. Of course, future studies should investigate further the boundary conditions under which the zygomaticus is co-activated with the corrugator and levator during emotional and moral discourse comprehension.

Previously, we used the same morality and emotion materials in a reading-for-comprehension task (Kunkel et al., under review) and found no reliable EDA and fEMG effects indicative of embodied responses to incoming emotional language. We therefore speculated that the zero-effects in fEMG and EDA measures might be due to the fact that participants performed an emotion-unrelated task (Bradley et al., 2001). Based on the present results, this assumption can be clearly rejected as far as EDA is concerned. Across our studies and across both control and comprehension experiments, EDA was by and large insensitive to emotional versus neutral or immoral versus neutral stimuli. We therefore conclude that at least for the materials used by us, which might be characterized by less severe arousal differences as compared to those used in former studies (e.g., Bradley et al., 2001; Larsen et al., 2003), EDA can be ruled out as a sensitive marker of embodied emotional responses. The picture is different regarding fEMG and ECG measures, which were reliably influenced by negative compared to neutral stimuli in the control experiment, and which were also differentially sensitive to immoral versus moral and to negative versus neutral scenarios, respectively. Here, future

studies on discourse comprehension should continue measuring theses psychophysiological variables in order to better understand the factors that modulate their joint or differential occurrence. For instance, other explicit tasks (e.g., valence or amorality judgment tasks) than the one used here might also enhance the emotional salience of morality or emotional scenarios (e.g., Cannon et al., 2011; Niedenthal et al., 2009). Moreover, not only the arousing value or the valence of the situation per se might be critical for triggering an embodied response during discourse comprehension as indicated by ECG and fEMG, respectively, but also how the situation is affectively evaluated and perceived (cf. 't Hart et al., 2018; Van Berkum, 2018a, 2018b).

In conclusion, the present fEMG findings extend previous work (Kunkel et al., under review) by providing evidence for a possible role of affective simulations during the discourse-based processing of morality materials when an explicit, emotional-related judgement task is demanded. Moreover, ECG findings suggested that also negative versus neutral emotional materials lead to differential responses of the body. Finally, given the different ERP patterns when participants judge the moral acceptability of behaviors described in the scenarios versus when they read for comprehension (Kunkel et al., 2018; Leuthold et al., 2015), for future studies would still need to investigate in which way embodied simulations during discourse comprehension depend on the specific task demands.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Reviews of Psychology, 59,* 617-645.

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12*, 197-204.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version, 1*, 1-23.

Benedek, M. & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*, 80-91.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1*, 276-298.

Bradley, M. M., & Lang, P. J. (2000). Emotion and motivation. *Handbook of Psychophysiology, 2*, 602-642.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Cannon, P. R., Schnall, S., & White, M. (2011). Transgressions and expressions: Affective facial muscle activity predicts moral judgments. *Social Psychological and Personality Science, 2*, 325-331.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159-181). New York, NY: Cambridge University Press.

Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology, 27,* 481–494.

Filik, R., Hunter, C. M., & Leuthold, H. (2015). When language gets emotional: Irony and the embodiment of affect in discourse. *Acta Psychologica, 156*, 114-125.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*, 567-589.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98-116.

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic use of botulinum toxin-A affects processing of emotional language. *Psychological Science, 21,* 895-900.

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14*, 436-441.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perceptio,* 36 ECVP Abstract Supplement.

Krumhuber, E. G., Tsankova, E., & Kappas, A. (2018). Examining subjective and physiological responses to norm violation using text-based vignettes. *International Journal of Psychology, 53,* 23-30.

Kunkel, A., Filik, R., Mackenzie, I. G., & Leuthold, H. (2018). Task-dependent evaluative processing of moral and emotional content during comprehension: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience, 18*, 389-409.

Kunkel, A., Mackenzie, I. G., Filik, R., & Leuthold, H. (under review). Implicit evaluative processing of moral and emotional content during discourse comprehension.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*, 261-273.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology, 40*, 776-785.

Leuthold, H., Kunkel, A., Mackenzie, I. G., & Filik, R. (2015). Online processing of moral transgressions: ERP evidence for spontaneous evaluation. *Social, Cognitive, and Affective Neuroscience, 10*, 1021-1029.

Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*, 1002-1005.

Niedenthal, P. M., Winkielman, P., Mondillon, L., & Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology, 96*, 1120-1136.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011, 1.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision, 10*,437-442.

Sanford, A. J., & Emmott, C. (2012). *Mind, Brain and Narrative*. Cambridge University Press.

Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., & Traue, H. C. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating emotions. *Journal of Ambient Intelligence and Humanized Computing, 3*, 3-10.

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2007). The skeletomotor system: Surface electromyography. In: Cacioppo JT, Tassinary LG, Berntson GG (eds). *Handbook of psychophysiology*, 3rd edn. Cambridge University Press, New York, pp. 267–302.

't Hart, B. T. H., Struiksma, M. E., Van Boxtel, A. H., & Van Berkum, J. J. (2018). Emotion in Stories: Facial EMG Evidence for Both Mental Simulation and Moral Evaluation. *Frontiers in Psychology*, *9*, 613.

Thompson, D., Mackenzie, I. G., Leuthold, H., & Filik, R. (2016). Emotional responses to irony and emoticons in written language: evidence from EDA and facial EMG. *Psychophysiology*, *53*, 1054-1062.

Unz, D. C., & Schwab, F. (2005). Viewers viewed: Facial expression patterns while watching TV news. *Emerging Communication*, *7*.

Van Berkum, J. J. A. (2018a). Language comprehension, emotion and sociality: aren't we missing something?" *in Oxford Handbook of Psycholinguistics*, edsM. G. Rueschemeyer and S.-A.Gaskell (Oxford: OUP).

Van Berkum, J. J. A. (2018b). Language comprehension and emotion: Where are the interfaces, and who cares. *Chapter Handbook of Neurolinguistics*.

Van Boxtel, A. (2010). Facial EMG as a tool for inferring affective states. In A. J. Spink, F. Grieco, O. E. Krips, L. W. S. Loijens, L. P. J. J. Noldus, and P. H. Zimmerman (Eds.), *Proceedings of measuring behavior* (pp.104-108). Wageningen, The Netherlands: Noldus Information Technology.

Weis, P. P., & Herbert, C. (2017). Bodily Reactions to Emotional Words Referring to Own versus Other People's Emotions. *Frontiers in Psychology*, *8*, 1277.

Acknowledgments

The work reported in this paper was supported by a grant of the German Research Foundation (DFG) awarded to HL (LE 1035/4-1).

Reprint requests should be sent to Angelika Kunkel, Department of Psychology, Eberhard Karls University Tübingen, Schleichstr. 4, 72076 Tübingen, Germany, or via e-mail: angelika.kunkel@uni-tuebingen.de.

3.3.3 Study 3: Kunkel, A., Mackenzie, I.G., Filik, R., & Leuthold, H. (under review). *Implicit evaluative processing of moral and emotional content during discourse comprehension*

Authors involved:

- Prof. Dr. Hartmut Leuthold (HL)
- Dr. Ian G. Mackenzie (IGM)
- Prof. Dr. Ruth Filik (RF)
- Angelika Kunkel (AK)

List of author contributions

- Study conception and design: HL and AK
- Programming: IGM
- Acquisition of data: AK, bachelor students, and student assistant
- Analysis: IGM, AK, and HL
- Interpretation of data: HL, AK, IGM, and RF
- Writing paper: AK, HL, IGM, and RF

Implicit evaluative processing of moral and emotional content during discourse comprehension

Angelika Kunkel, Ian Grant Mackenzie

Department of Psychology, University of Tübingen, Germany

Ruth Filik

School of Psychology, University of Nottingham, United Kingdom

Hartmut Leuthold

Department of Psychology, University of Tübingen, Germany

Address correspondence to:

Angelika Kunkel
Department of Psychology
Eberhard Karls University of Tübingen
Schleichstrasse 4
72076 Tübingen
Germany

Email: angelika.kunkel@uni-tuebingen.de
Phone: +49 (0)7071 29-75590

Abstract

Recently, using event-related potentials (ERPs), we showed that when participants read about daily moral transgressions, the task that they performed determined the processing of incoming linguistic information (Kunkel, Filik, Mackenzie, & Leuthold, 2018). Specifically, when explicit moral judgments were required a larger anterior negativity indicated increased cognitive processing for immoral than moral scenarios, whereas for emotional judgments a larger late posterior positivity (LPP) indicated increased affective processing. The present two experiments examined which brain processes and embodied responses contribute to discourse comprehension if no task is required. Target sentences from negative vs. neutral emotional scenarios and from moral vs. immoral scenarios were presented using rapid serial visual presentation while ERPs (Experiment 1), and facial electromyogram (fEMG) and electrodermal activity (EDA, Experiment 2) were recorded. In Experiment 1, the LPP (300-500 ms) was larger for negative and immoral compared to neutral and moral scenarios. In Experiment 2, fEMG and EDA results were not sensitive to experimental manipulations in either morality or emotion materials, whereas participants showed clear fEMG effects in corrugator activity to standard emotional pictures and words. We conclude that for the present task and materials, discourse comprehension involves the top-down affective evaluation (LPP) of incoming linguistic moral and emotional information but not the embodied simulation of emotion (fEMG, EDA).

Keywords: Moral, emotion, affective evaluation, LPP, discourse comprehension, fEMG, EDA, embodiment

A fundamental aspect of the mind's functioning is that objects or events are initially evaluated and categorized with regard to their valence as either good or bad (e.g., Storbeck & Clore, 2007; Zajonc, 1984), before being reanalyzed by explicit and slower evaluative cognitive processes (cf. Cunningham & Zelazo, 2007). Recently, using event-related potentials (ERPs), we showed that when participants read about daily moral transgressions or emotional events, and explicitly judged their emotional content as inferred from the discourse context, a larger late positive potential (LPP) was found to immoral than moral as well as negative than neutral scenarios (Kunkel, Filik, Mackenzie, & Leuthold, 2018). In line with previous ERP studies (for reviews, see Fischler & Bradley, 2006, Hajcak, Weinberg, MacNamara, & Foti, 2012), we took this LPP effect to indicate the increased affective processing of incoming linguistic information during discourse comprehension. However, it is currently unclear which brain processes contribute to emotional discourse comprehension if no explicit task is required, and also whether embodied responses are involved. Therefore, in the present study we conducted two experiments to investigate (a) whether, in a simple passive reading task, both morality and emotion scenarios would still elicit such discourse-based affective evaluations, as indicated by the LPP, and (b) whether these evaluations would be grounded in embodied emotional responses, as shown by facial electromyographic (fEMG) and electrodermal activity (EDA).

The LPP has been found to sensitively reflect the automatic processing of the emotional content of pictorial stimuli (e.g., Gable, Adams, & Proudfit, 2015; Hajcak, MacNamara, Foti, Ferri, & Keil, 2013; Ito & Cacioppo, 2000; for a review see Hajcak, MacNamara, & Olvet, 2010) and also of linguistic input such as emotional words, sentences, and texts (for reviews, see Citron, 2012; Fischler & Bradley, 2006; Hajcak et al., 2012). LPP effects are known to depend on the emotional features of the critical item

(e.g., valence, arousal) and to be more pronounced when participants judge the emotional content rather than an emotion-irrelevant stimulus dimension (e.g., in a semantic classification) (cf. Fischler & Bradley, 2006). For instance, word-based emotion effects in the LPP were found to be stronger in an affective judgment task than a passive reading task (Holt, Lynn, & Kuperberg, 2009; see also Fields & Kuperberg, 2016). Moreover, Bartholow, Fabiani, Gratton, and Bettencourt (2001), asked participants to intentionally infer either positive or negative character traits from a passage of text (e.g., that someone is a friendly person). A larger late centroparietally distributed ERP positivity (450-1150 ms), or LPP, was observed after final affective words in the test sentence that described trait-inconsistent (". . . gave his wife a slap") rather than trait-consistent behavior (". . . gave his mother a kiss"). However, it has also been found that if a neutral rather than an emotional context preceded the target sentence during a comprehension task, a larger N400 component for emotion-congruent than incongruent conditions preceded the LPP (Delaney-Busch & Kuperberg, 2013; Holt et al., 2009), reflecting the semantic processing demands during meaning construction (for a review, see Kutas & Federmeier, 2011). Also, in ERP studies using contexts that were strongly constraining, a larger N400 was obtained for discourse-incongruent than congruent emotional words in the target sentence (e.g., León, Díaz, de Vega, & Hernández, 2010; Leuthold, Filik, Murphy, & Mackenzie, 2012). Crucially, in the above ERP studies on emotional comprehension, except for Leuthold et al. (2012), the critical words in the target sentence differed across emotion conditions, thereby precluding a straightforward interpretation of ERP effects in terms of discourse-based emotion inferences.

To our knowledge, Kunkel et al. (2018) were the first to investigate whether the emotional meaning of identical target sentences, and hence identical critical words,

are differentially processed when different emotions were inferred on the basis of only moderately constraining contexts. That is, emotional discourse comprehension was examined under conditions that excluded possible influences of differing emotion words (e.g., Delaney-Busch & Kuperberg, 2013; Holt et al., 2009) and also of emotional congruency effects due to the use of emotionally constraining contexts (e.g., León et al., 2010; Leuthold et al., 2012). More specifically, Kunkel et al. employed prototypical morality and emotion scenarios for which the protagonists and situations were introduced by the context sentences (for examples, see Table 1). Morality scenarios described actions that would be perceived as either morally acceptable, like helping a friend in need, or morally unacceptable, like adultery. For the sake of brevity, we use the labels moral and immoral scenarios in the following. Emotion materials described either a neutral (or mildly positive) compared to a negative event or situation.

For both morality and emotion materials, the context varied between the respective material-specific conditions. It was followed by the rapid serial visual presentation (RSVP) of the target sentence, which was identical for moral versus immoral scenarios and for negative vs. neutral scenarios. Thus, the critical word of the target sentence was embedded in a meaningful context and disambiguated the situation in terms of either a moral vs. immoral or neutral vs. negative outcome. Crucially, one major advantage of this procedure is that resulting ERP effects reflect discourse-based influences on information processing rather than influences related to the target word. Participants performed different judgment tasks while ERPs were recorded. In the first experiment, participants were asked to judge the described behavior as morally acceptable or unacceptable in the case of morality scenarios or as emotionally moving or not in the case of emotion scenarios. In the second experiment, participants made emotion-related judgments for both types of scenarios. The ERP results for emotion

materials showed a long-lasting LPP emotion effect from 300-700 ms in both experiments, suggesting a more intense affective evaluation of negative than neutral items. For the moral judgment task, immoral items elicited a larger anterior negativity than moral items about 500-700 ms after critical word onset, which was assumed to reflect the differential semantic processing of moral and immoral materials. However, morality scenarios elicited an LPP effect in the 300-500 ms time window when participants judged the emotional content of these scenarios, indicating that immoral items underwent more intense affective processing than moral items. Together with offline valence and arousal rating results and the analysis of binary judgments for the respective materials, these findings accord with the view that moral-laden, and specifically immoral materials, contain emotional information to which participants have quick access.

In relation to this, it is important to note that an ERP study conducted by Leuthold, Kunkel, Mackenzie, and Filik (2015) demonstrated that affective evaluations contributed to the processing of very similar morality scenarios during discourse comprehension when no explicit judgments were required. Unlike Kunkel et al. (2018), participants read the everyday morality scenarios simply for comprehension. As a control, morality scenarios were randomly intermixed with world knowledge violation scenarios (knowledge-consistent vs. inconsistent: e.g., a target sentence of "She receives as a dish a plate full of *snails* and white bread."), following a context that would make this statement either consistent or inconsistent with the participants' knowledge of the world. The ERP results for morality scenarios yielded an LPP effect starting at about 320 ms after the presentation of the critical word, similar to that observed by Kunkel et al. (2018) with emotion judgments. By contrast, world knowledge-inconsistent trials showed a larger N400 compared to consistent trials. Leuthold et al. (2015) concluded

that immoral items underwent more intense affective rather than semantic processing, that is, they assumed that the behaviors described in morality items were implicitly categorized as either good or bad.

Together, the studies of Kunkel et al. (2018) and Leuthold et al. (2015) suggest that incoming linguistic information is affectively evaluated in relation to the specific discourse context, and seemingly even so when no explicit judgment task is involved. However, it has been proposed that the global context such as task demands and the use of emotion-related versus emotion-neutral linguistic stimuli influence whether an affective versus semantic-cognitive analysis is prioritized (e.g., Delaney-Busch & Kuperberg, 2013; Lai, Hagoort, & Casasanto, 2012). Thus, it is conceivable that in Leuthold et al.'s (2015) study, the affective dimension of morality materials was relatively prominent due to the intermixed presentation of emotion-neutral world-knowledge materials. In contrast, Kunkel et al. (2018) presented emotion versus morality materials in different experimental halves, also when the same emotion judgments were required to both types of materials. The more sustained LPP emotion effect and larger emotion effect in terms of both valence and arousal for emotion than morality materials in the offline rating study (Kunkel et al., 2018) might suggest a higher affective salience for emotion than morality materials. Therefore, it remains an open issue whether morality materials would still undergo an implicit affective evaluation as indicated by the LPP (cf. Leuthold et al., 2015) in a passive reading-for-comprehension task with randomly intermixed morality and more salient emotion materials. Moreover, it is unclear whether *emotion* materials are implicitly evaluated in this task context, because previously an LPP effect for these materials was obtained when participants explicitly judged their affective content (Kunkel et al., 2018).

Assuming that immediate affective evaluations of language input contribute to emotional text comprehension, one might then further ask whether such comprehension is grounded in language-induced experiential simulations. That is, according to the embodied simulation account, the meaning of a stimulus is grounded in ordinary past sensorimotor experiences (e.g., Barsalou, 1999; Gibbs, 2003; Glenberg & Robertson, 1999; Kiefer & Pulvermüller, 2012). Thus, to comprehend a piece of text, it is assumed that a concrete representation of a sensory or motor experience is re-activated or, in other words, that such an experience is simulated during the processing of linguistic input. Crucially for the present purposes, research in embodied cognition has also provided evidence that emotional comprehension involves the (unconscious) reactivation and simulation of previously experienced emotions during meaning construction (Niedenthal, 2007). Using fEMG to reveal facial muscle movements related to certain emotional expressions (cf. Dimberg, 1990; van Boxtel, 2010), Niedenthal, Winkielman, Mondillon, & Vermeulen, (2009) demonstrated embodied effects when participants made affective judgments about the meaning of emotional words presented in isolation. More specifically, they found that positive- and negative-valenced words elicited stronger activation in facial muscles involved in smiling (e.g., zygomaticus major) and frowning (e.g., corrugator supercilii) than neutral words. No such facial EMG effects were obtained when participants performed an emotion-unrelated task. On the basis of these results Niedenthal and colleagues concluded that emotion words trigger language-based simulations in the emotion system but only if required by the task.

With regard to text comprehension, studies using behavioral approaches have provided evidence supporting the embodied processing of sentences describing emotion-laden events (Filik, Hunter, & Leuthold, 2015; Havas, Glenberg, & Rinck, 2007; Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010). To our knowledge,

there are only two studies to date reporting the involvement of affective simulations during discourse comprehension using peripheral psychophysiological measures (Bartholow et al., 2001; Thompson, Mackenzie, Leuthold, & Filik, 2016). As described earlier, Bartholow et al. (2001) examined the processes underlying intentional (explicit) trait inferences. In addition to their LPP findings, they reported stronger early fEMG activity (100-300 ms) from the corrugator supercilii to final affective words that described trait-inconsistent than trait-consistent negative behavior, whereas such an fEMG effect was absent for words describing positive behaviors. Although Bartholow and colleagues did not discuss their findings in the context of embodied cognition, their fEMG results nevertheless appear to suggest that an embodied response is rapidly elicited while readers evaluate an emotionally salient (negative) behavior of another person. More recently, Thompson et al. measured fEMG activity and also phasic EDA responses, as a measure of stimulus-induced arousal (Dawson, Schell, & Filion, 2007), while participants read statements conveying either ironic or literal criticism or praise (following a sentence that set up the context). In terms of facial EMG activity, they found evidence for reduced frowning (corrugator supercilii) and enhanced smiling (zygomaticus major) for ironic compared to literal criticism, but enhanced frowning and reduced smiling for ironic compared to literal praise, suggesting that irony weakens the emotional impact of language input. Moreover, phasic EDA responses were stronger when a meaning-disambiguating emoticon was present than not. These findings demonstrate the suitability of facial EMG and phasic EDA in revealing the embodied grounding of emotional comprehension during a passive reading task.

In summary, the first open issue is whether emotion and morality scenarios undergo an (implicit) affective evaluation at the level of brain processing when presentation is randomly intermixed within a passive reading for comprehension task.

The second open issue is whether discourse-based emotional comprehension involves experiential affective simulations or embodied responses.

Objectives of the present study

The aim of the present two experiments is to investigate in the absence of an explicit judgment task (a) whether similar to previous studies (Kunkel et al., 2018; Leuthold et al., 2015) incoming linguistic information is affectively evaluated as indicated by the LPP during emotional discourse comprehension (Experiment 1) and (b) to examine the potential embodied responses as indicated by fEMG and EDA (Experiment 2). For both experiments, we will use the same text materials as in Kunkel et al. (2018) but adopt the task and trial presentation procedure used by Leuthold et al. (2015). That is, ERPs, fEMG, and EDA will be recorded in a reading for comprehension task, in which scenarios describing everyday morality or emotional situations are randomly intermixed. ERP, fEMG amplitudes and EDA responses elicited by scenarios describing everyday moral transgressions will be compared to those elicited by moral scenarios, and situations with negative emotional content will be compared to those with neutral or mildly positive content. This will allow us to assess whether the respective materials differ with regard to their affective evaluation and associated (embodied) simulations within the emotion system. Generally, we hypothesized that for both types of materials, the emotional content of language input is affectively processed at the level of brain and peripheral body responses. The specific predictions will be discussed separately for each experiment.

Experiment 1

In Experiment 1, we examined the ERP in order to reveal affective (brain) processes during emotional discourse comprehension. First, we expect that similar to Kunkel et al.

(2018), negative compared to neutral scenarios will elicit a relatively long-lasting (300-700 ms) emotion effect as indicated by the LPP. Of course, it is possible that the LPP effect will be of smaller amplitude or not last as long lasting since no explicit affective judgments are demanded (cf. Fields & Kuperberg, 2016; Fischler & Bradley, 2006; Holt et al., 2009). Specifically, if morality materials are affectively evaluated even when embedded within the context of potentially more salient emotion materials, we expect to replicate the LPP findings of Leuthold et al. (2015).

Method

Participants. 33 native German speakers from the University of Tübingen received course credits or payment for participating. Data from five participants were excluded, due to excessive sweat artifacts (N = 2), excessive alpha activity (N = 2), or not having German as a first language (N = 1). For all analyses, we used the dataset from the remaining 28 right-handed participants (M_{age} = 23.86 years, SD_{age} = 4.66 years, 22 females).

Materials and design. Materials were taken from Kunkel et al. (2018) and consisted of 80 items each for the morality and for the emotion scenarios (see Table 1 for examples; the full set is available from the first author). All scenarios consisted of two parts. The first part consisted of two or three sentences describing the context, and the second part was the target sentence containing the critical word. The same target sentence was used for moral and immoral materials, and the same held true for neutral and negative emotional materials in order to eliminate possible sentence-level and word-based effects (with the context varying across conditions, see Table 1). Moral materials described actions that would be perceived as either moral or immoral, whereas emotional materials would describe a neutral or a negative event. In addition, 40 filler items were

Table 1. Example for moral materials with context for moral and for immoral items, for emotional materials with context for neutral and for negative items, as well as the respective target sentences containing the critical word (in italics).

Moral Materials	Moral	Herr Zimmermann arbeitet in einem Chemiekonzern. Seit einigen Wochen gibt es Sicherheitsprobleme aufgrund fahrlässigen Verhaltens von Kollegen, sodass Menschenleben gefährdet sind. (Mr. Zimmermann works at a chemical company. For several weeks, there have been safety issues caused by colleagues' careless behavior, putting human lives at risk.)
	Immoral	Herr Zimmermann arbeitet in einem Chemiekonzern als Sachbearbeiter in der Personalabteilung. Er ist ehrgeizig und will rascher aufsteigen als seine Kollegen. (Mr. Zimmermann works in human resources of a chemical company. He is ambitious and aims at being promoted faster than his colleagues.)
	Target Sentence	Im monatlichen Gespräch berichtet er seinem Chef deren *Fehler*. (In a monthly commitment talk he points out to his boss their *mistakes*.)
Emotion Materials	Neutral	Sarah trainierte den ganzen Nachmittag im Fitness-Studio und geht zum Abschluss in die Sauna, wonach sie sich wohlig gut fühlt. (Sarah had been at the gym all afternoon before completing her work out with a visit to the sauna, after which she feels pleasantly well.)
	Negative	Sarah hat einen sehr anstrengenden Job. Sie muss dafür seit mehreren Monaten jeden Tag um 5 Uhr aufstehen und den ganzen Tag hart arbeiten. (Sarah has a very demanding job. For several months she has been getting up at 5 am and working very hard all day.)
	Target Sentence	Sie ist sehr *erschöpft*. (She is very *exhausted*.)

used which contained no inconsistencies and were similar in length to the experimental items. Furthermore, true/false questions which were not related to emotional or moral content were used to check for comprehension (e.g., "Herr Zimmermann geht tauchen." [Mr. Zimmermann goes diving.] for the morality example provided in Table 1). Relevant item characteristics and rating data are presented in Table 2 (for more details, see Kunkel et al., 2018).

The presentation and randomization of items and conditions across participants was identical to that of Leuthold et al. (2015). For two consecutive participants, two lists were randomly generated such that each morality scenario appeared across the two lists either in the moral or the immoral condition, and each emotion scenario appeared either in the neutral or the negative condition. Thus, for each participant, the 200-item list consisted of 40 moral and 40 immoral items, 40 neutral and 40 negative items, as well as 40 neutral filler items. All items were randomly presented.

Procedure. After electrode application, participants were seated in an electrically shielded booth in front of a 21-in. computer monitor (100 Hz) at a viewing distance of 65 cm (maintained by a chin rest). Experimental materials were presented at the center of the screen in white 16-point Helvetica font on a black background using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) running under MATLAB (2017a), on Ubuntu 16.04 LTS. Participants were instructed to avoid any eye, head, and jaw movements and to maintain fixation at the center of the screen during word-by-word presentation. Furthermore, they were instructed to read the stories attentively, and to answer comprehension questions by pressing the appropriate response key. The experiment started with the instructions followed by a practice block containing three trials. Experimental trials were presented in a total of 20 blocks of 10 items each, separated by a short break. Both instructions and breaks were controlled in

Table 2. Characteristics and rating data of morality, emotion, picture, and word materials.

	Morality	Emotion			Pictures		Words	
	Moral (SE)	Immoral (SE)	Neutral (SE)	Negative (SE)	Neutral (SE)	Negative (SE)	Neutral (SE)	Negative (SE)
Word Frequency	59.46 (9.72)		58.31 (6.98)				13.78 (2.58)	13.61 (2.64)
Word length	8.66 (0.21)		7.51 (0.19)				6.15 (0.28)	6.40 (0.37)
Valence	5.44 (0.13)	2.60 (0.09)	5.07 (0.12)	2.39 (0.07)	5.04 (0.11)	2.37 (0.06)	0.00 (0.00)	-2.59 (0.04)
Arousal	3.79 (0.13)	4.34 (0.11)	3.51 (0.13)	5.00 (0.12)	3.40 (0.12)	6.10 (0.10)	2.42 (0.09)	4.28 (0.06)
Morality	5.99 (0.04)	2.52 (0.03)						

Note. Means and standard errors (in brackets) of the rating data were calculated for each material and control stimulus set and condition. Word frequency (per million) and word length concern the critical words and words of the BAWL-R. For valence and arousal ratings different scales are used. For morality and emotions materials, valence and arousal ratings concern the whole scenario inclusive target sentence rated on an 8-point scale (1 = *very negative, not touched*, 8 = *very positive, strongly touched*). Only morality materials were rated for the degree of morality (1 = *very immoral*, 8 = *very moral)*. Valence and arousal ratings for words were taken from the BAWL-R (Võ et al., 2009), with valence rated on a 7-point scale (-3 = *very negative*, 3 = *very positive*) and arousal rated on a 5-point scale (1 = *low arousal*, 5 = *high arousal*). Arousal and valence ratings for pictures were taken from the IAPS (Lang et al., 2005), both rated on a 9-point scale (valence, 1 = *negative*, 9 = *positive*; arousal, 1 = *high arousing*, 9 = *not arousing*).

their duration by the participants pressing the space bar. For a trial, the context was displayed for a minimum duration of 1,500 ms. When participants had read the context sentences, they initiated the word-by-word presentation of the target sentence by pressing the space bar, which started with the presentation of a fixation point for 1000 ms. Then, each word was displayed centrally for 300 ms, with a 200-ms blank interval between successive word presentations. For each item, an emotion-unrelated true/false question was constructed. In 20% of randomly selected trials, the corresponding question was asked immediately after presentation of the final word, and was displayed until participants pressed the 'x' or 'm' key on the computer keyboard to indicate their 'true' or 'false' response.

Electrophysiological measures. Electroencephalographic (EEG) activity was recorded continuously from 72 Ag-AgCl electrodes using a BIOSEMI Active-Two amplifier system with a sampling rate of 512 Hz for EEG and electrooculogram (EOG). All EEG/ERP analyses were performed using available MATLAB toolboxes (FieldTrip: Oostenveld, Fries, Maris, & Schoffelen, 2011) and custom MATLAB scripts. Following high-pass filtering (0.1 Hz, 12 dB/oct), EEG data were checked for artifacts and corrected, following a procedure similar to that described in Dudschig, Mackenzie, Strozyk, Kaup, and Leuthold (2016). There remained on average 96.11 % trials (out of 40; range = 26-40, median = 39) per condition; trial numbers did not differ between conditions, all *t*s(58) < 0.70, *p*s > .48.

Data analysis. The analysis epoch started 200 ms prior to the onset of the critical word and lasted until 1,500 ms after it. Off-line, all EEG channels were recalculated to an average reference[1] with channels F9, F10, M1, M2, IO1, and IO2

[1] An average reference was chosen due to noisy mastoid electrodes for one third of the participants. The same pattern of ERP amplitude results, however, was found when using an average mastoid reference.

excluded from this calculation. For artifact-free trials, the signal at each electrode site was averaged separately for each experimental condition, time-locked to the onset of the critical word, low-pass filtered (30 Hz, 36 dB/oct), and aligned to a 200-ms baseline prior to the onset of the critical word. Similar to our previous studies (cf. Kunkel et al., 2018; Leuthold et al., 2015), mean ERP amplitudes were determined for the following time ranges: 200 to 250 ms, 300 to 500 ms, and 500 to 700 ms. Since we were principally interested in emotion-related LPP modulations, similar to the posterior region of interest (ROI) of Kunkel et al. (2018), ERP signals were averaged across centro-parietal electrodes only (CP1, CPz, CP2, P1, Pz, P2, POz) to create a single region of interest (ROI).

For this single ROI, mean ERP amplitudes for emotion and morality materials were determined and subjected to separate repeated measures analyses of variance (ANOVA) with the factor *condition* (moral vs. immoral or neutral vs. negative). Separate statistical analyses were performed for morality and emotion materials since the critical target sentences for the two types of materials differed in wording, making a direct comparison of ERP effects triggered by emotion versus morality materials problematic. The significance level was set to alpha = .05.

Results

200-250 ms. There were no significant effects in this time interval, all *p*s > .13.

300-500 ms. The analysis of morality materials revealed a significant condition effect, $F(1, 27) = 6.95$, $p < .05$, $\eta_p^2 = .20$, due to a larger positivity for immoral than moral sentences (2.07 vs. 1.54 μV). The condition effect was also significant for emotion materials, $F(1, 27) = 5.82$, $p < .05$, $\eta_p^2 = .18$, due to larger positivity for negative than neutral sentences (2.22 vs. 1.62 μV).

500-700 ms. For morality materials, this analysis yielded a trend for the condition effect, $F(1, 27) = 3.14$, $p = .09$, $\eta_p^2 = .10$, due to a slightly larger positivity for immoral than moral sentences (2.74 vs. 2.40 μV). For emotion materials, the condition effect was significant, $F(1, 27) = 5.88$, $p < .05$, $\eta_p^2 = .18$, due to a larger positivity for negative than neutral sentences (3.27 vs. 2.69 μV).

In addition, to reveal whether there are task-dependent ERP amplitude differences, we first calculated average-reference ERP waveforms for Kunkel et al.'s (2018, Experiment 2) data and determined ERP amplitudes for the posterior ROI and same time intervals as in the present Experiment 1 for both emotion and morality materials. Then, we combined the respective ERP amplitude data sets and conducted separate ANOVAs for morality and emotion materials including the between-subjects factor of task (affective judgment vs. reading) and the repeated measurement factor condition (neutral vs. negative and moral vs. immoral).

200-250 ms. In this time window, only the analysis of emotion materials revealed a significant main effect of condition, $F(1, 54) = 11.41$, $p < .01$, $\eta_p^2 = .17$, and a Task x Condition interaction, $F(1, 54) = 4.19$, $p < .05$, $\eta_p^2 = .07$, due to a larger positivity for negative than neutral sentences in the affective judgment task only (1.02 vs. 0.24 μV), $F(1, 27) = 14.19$, $p < .001$.

300-500 ms. The analysis of morality materials revealed a larger positivity for immoral than moral sentences (2.07 vs. 1.64 μV), $F(1, 54) = 9.07$, $p < .01$, $\eta_p^2 = .14$, and this morality effect was not modulated by task, $F = 0.45$, $p = .51$. For emotion materials, there was a larger positivity for negative than neutral sentences (2.64 vs. 1.62 μV), $F(1, 54) = 38.46$, $p < .001$, $\eta_p^2 = .42$. The significant Task x Condition interaction, $F(1, 54) = 6.58$, $p < .05$, $\eta_p^2 = .11$, indicated a stronger emotion effect for the judgment than the reading task (negative minus neutral = 1.45 vs. 0.60 μV).

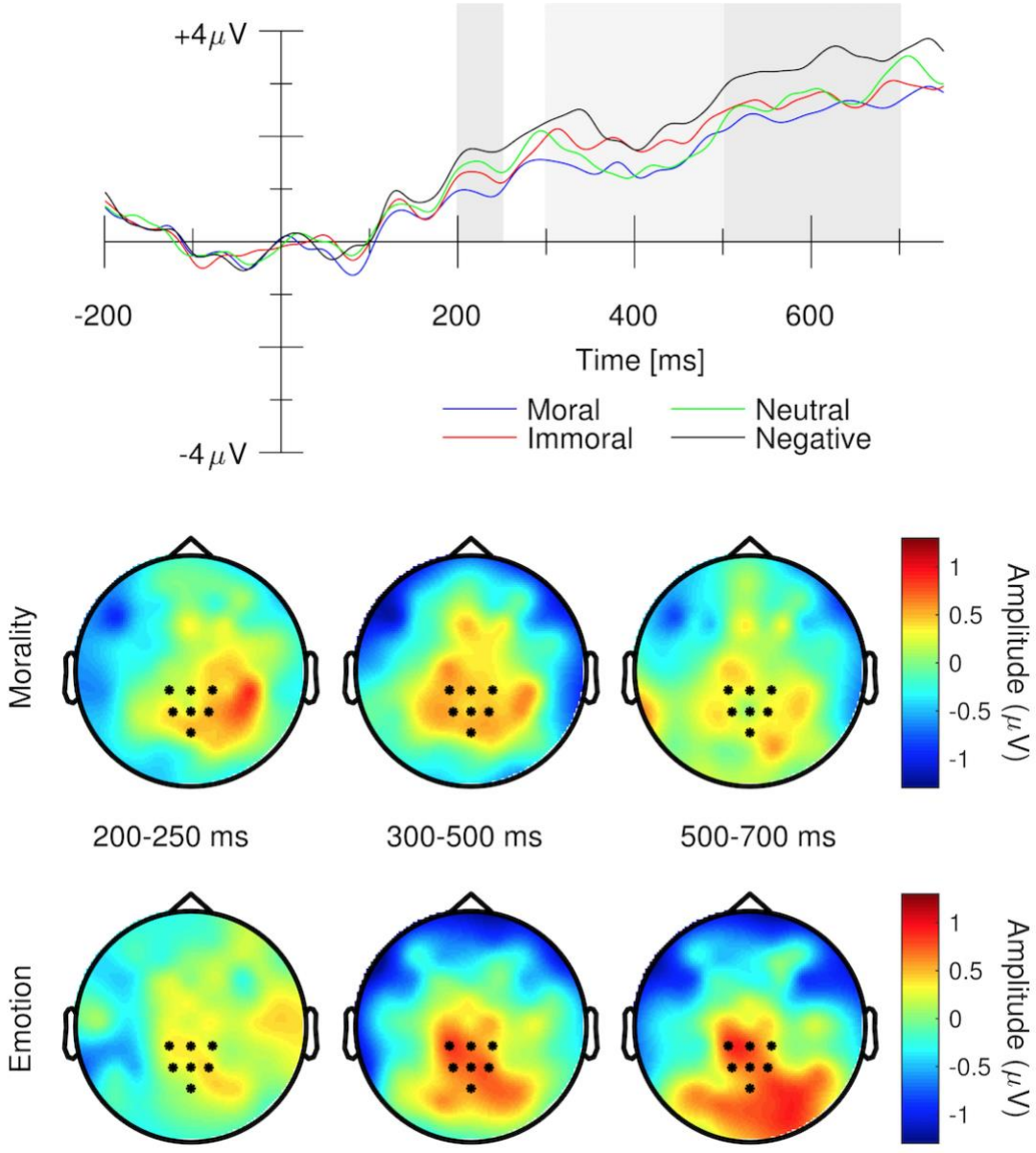
500-700 ms. It was only for emotion materials that the condition effect, $F(1, 54) = 35.30$, $p < .001$, $\eta_p^2 = .40$, and the Task x Condition interaction were significant, $F(1, 54) = 5.62$, $p < .05$, $\eta_p^2 = .09$. The interaction indicated a stronger emotion effect for the judgment task than the reading task (1.36 vs. 0.58 μV).

Discussion

In Experiment 1, participants read randomly intermixed morality and emotion scenarios for comprehension without an explicit judgment task. In line with our expectations, a larger LPP amplitude was obtained for negative compared to neutral scenarios and for immoral compared to moral scenarios, with the LPP effect starting at about 300 ms for both types of materials. Crucially, we replicated the LPP effect for morality materials when embedded amongst emotion-related scenarios instead of emotionally-neutral world knowledge violations (e.g., Leuthold et al., 2015). Therefore, the affective evaluation of morality scenarios appears to be independent of the global context (emotion-related vs. emotion-neutral) but dependent on task demands (moral vs. emotion judgment) (cf. Kunkel et al., 2018). The present LPP effect indicates that incoming information is emotionally (or motivationally) more salient for immoral compared to moral scenarios and is more rapidly

affectively processed during discourse comprehension. Interestingly, there was no differential LPP effect for the passive reading versus the affective judgment task. Thus, it appears that when no explicit moral judgments are required, the language processing system is generally biased towards the processing of the more salient emotional rather than the less salient cognitive content of morality scenarios.

Figure 1. Upper panel: Condition-dependent grand average ERP waveforms elicited at the centro-parietal ROI (averaged across CP1, CPz, CP2, P1, Pz, P2, POz), time-locked to the onset of the critical word for morality and emotion materials. Positivity is plotted upwards. Lower panel: Spline-interpolated topographic map of mean ERP difference waveform for the 200-250 ms, 300-500 ms, and 500-700 ms time window. Top panel: morality condition (immoral minus moral). Bottom panel: emotion condition (negative minus neutral).



This assumption is further corroborated by the presence of an LPP effect for emotion materials. Crucially, the larger LPP for negative than neutral items between 300 and 700 ms replicated the LPP effect observed by Kunkel et al. (2018) when explicit affective judgments were required. In accord with previous research (e.g., Fischler & Bradley, 2006; Holt et al., 2009), the LPP effect was of smaller magnitude for the passive reading than the affective judgement task. Still, the present LPP findings suggest that negative scenarios, due to their emotional salience, are affectively evaluated even if not demanded by the task, further corroborating the view that language input is affectively evaluated by default. Moreover, we found no evidence for the presence of an N400 effect preceding the LPP. That is, a larger N400 was observed in previous discourse-based comprehension studies if the target sentence's critical emotion word was incongruent (positive or negative word) rather than congruent (neutral) with the neutral discourse context (Delaney-Busch & Kuperberg, 2013; Holt et al., 2009). A similar N400 effect was obtained when discourse contexts were strongly constraining (e.g., León et al., 2010; Leuthold et al., 2012).

In summary, and in accord with similar previous research (Kunkel et al., 2018; Leuthold et al., 2015), we take the LPP to reflect the discourse-based affective evaluation of incoming linguistic information. Yet, it remains unclear whether the comprehension of emotional language input is grounded in language-induced affective simulations. The next experiment aims at investigating potential embodied responses elicited by the same morality and emotion scenarios.

Experiment 2

Experiment 2 measured peripheral psychophysiological signals (fEMG, EDA) while participants read the same morality and emotion scenarios as in Experiment 1 in order

to examine affective simulations during emotional discourse comprehension. That is, as in a previous study on irony comprehension of Thompson and colleagues (2016), we analyzed phasic EDA activity, providing an index of sympathetic arousal of the underlying sudomotor nerve activity (Dawson et al., 2007), and fEMG activity to reveal involuntary facial muscle movements, which have been related to the simulation of specific emotions (Niedenthal, 2007; Niedenthal et al., 2009). More specifically, fEMG activity from the corrugator supercilii (frowning) is taken to sensitively indicate negative but also positive affect, whereas fEMG zygomaticus major (smiling) activity indicates positive affect while being insensitive to negative affect (e.g., Larsen, Norris, & Cacioppo, 2003; Tan et al., 2012).

If the emotional content of incoming linguistic information triggers affective simulations, given that immoral compared to moral and negative compared to neutral scenarios are characterized by higher arousal value, we might expect this effect to be reflected in phasic EDA activity. Moreover, since negative and immoral items are characterized by their negative valence, fEMG activity in the corrugator muscle should be larger for immoral compared to moral and for negative compared to neutral scenarios. However, fEMG zygomaticus activity might show a weaker or no differential valence effect given that neutral and moral items are only slightly positive-valenced, as indicated by the offline rating study reported in Kunkel et al. (2018; see also Table 2).

Following the discourse comprehension task, participants performed a control experiment in which emotional and neutral words as well as pictures were presented in order to check the sensitivity of fEMG and EDA measurements. Negative and neutral pictures and words were selected from available databases (cf. Method section) in such a way that mean arousal and valence scores generally corresponded for picture and word

stimuli. Participants were asked to attentively process these stimuli and to answer occasional true/false questions.

Method

Participants. Forty native German speakers from the University of Tübingen received course credits or payment for participating (M_{age} = 22.32 years, SD_{age} = 3.44 years, 28 females, 32 right-handed).

Materials and design. For the discourse comprehension experiment, the same materials as in Experiment 1 were used. For the control experiment, materials consisted of 40 items each for pictures and words, which were matched for arousal and valence (cf. Table 2). 20 items each for neutral and negative pictures were taken from the international affective pictures system (IAPS, see Appendix; Lang, Bradley, & Cuthbert, 2005). Arousal and valence ratings on 9-point scales (valence, 1 = *negative* and 9 = *positive*; arousal, 1 = *high arousing* and 9 = *not arousing*) were taken from Lang et al. (2005). Negative pictures were rated as more negative than neutral pictures (2.37 vs. 5.04), $t(19) = 20.42$, $p < .001$, and as more arousing (6.10 vs. 3.40), $t(19) = 15.30$, $p < .001$.

20 items each for negative and neutral words were taken from the Berlin Affective Word List Reloaded (BAWL-R; Võ et al., 2009) and matched for frequency and length. Only nouns were selected by their valence and arousal ratings taken from the BAWL-R (cf. Appendix), with valence rated on a 7-point scale (-3 = *very negative*, 3 = *very positive*) and arousal rated on a 5-point scale (1 = *low arousal*, 5 = *high arousal*). Negative words were rated as more negative than neutral words (-2.59 vs. 0.00), $t(19) = 64.74$, $p < .001$, and as more arousing (4.28 vs. 2.42), $t(19) = 19.45$, $p < .001$.

Procedure. After electrode application, participants were seated in an electrically shielded booth in front of a 21-in. computer monitor (100 Hz) at a viewing distance of 65 cm (maintained by a chin rest). Experimental materials (scenarios, words, pictures) were presented at the center of the screen in white 16-point Helvetica font on a black background using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) running under MATLAB (2017a), on Ubuntu 16.04 LTS. Again, participants were instructed to avoid any head movements and to maintain fixation at the center of the screen during word-by-word presentation. Furthermore, they were instructed to read and watch attentively, and to answer control questions by pressing the appropriate response key.

The procedure of the emotional discourse comprehension experiment and the random presentation of morality and emotion items was identical to Experiment 1. Following this experiment, participants were instructed in writing on the screen about the control experiment, which contained eight blocks of 10 experimental trials each, separated by a short break controlled in its duration by the participant. Pictures and words were randomly presented within a block of trials. Every stimulus was displayed for 500 ms with a variable inter stimulus interval of five to eight seconds.

For each item a potential emotion-unrelated true/false question was constructed (e.g., “Das Tier auf dem Bild ist ein Schaf” [“The animal depicted is a sheep”]; “Das Wort besteht aus drei Silben” [“The word consists of three syllables”]). In 20% randomly selected trials, the corresponding question was asked immediately after presentation of a picture or word until participants indicated their ‘true’ or ‘false’ response by pressing the left or right arrow key on the computer keyboard. Response keys were changed compared to Experiment 1 since EDA was recorded from the

participants' non-dominant hand, therefore, the keys were operated by the index and middle finger of the dominant hand.

Psychophysiological measures. fEMG activity was recorded continuously from two pairs of Ag-AgCl electrodes using a BIOSEMI Active-Two amplifier system, one each for the zygomaticus major (smiling, cheek), and corrugator supercilii (frowning, inside brow) with a sampling rate of 1024 Hz. Surface electrodes were placed on the left side of the face following Fridlund and Cacioppo (1986).

EDA was continuously recorded from 2 Ag/AgCl electrodes using a BIOSEMI Active-Two amplifier system with a sampling rate of 1024 Hz. The electrodes were placed on the distal phalanges of the middle and index finger of the non-dominant hand.

Data analysis. All fEMG analyses were performed using code adapted from FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011) and custom MATLAB scripts. fEMG pre-processing involved applying a band-pass filter of 20-500 Hz to the continuous data as recommended by van Boxtel (2010). fEMG data were epoched around the onset of the critical word (-2 to 8 seconds). The differential activity between each electrode at the zygomaticus major and corrugator supercilii was calculated separately. This differential activity at each muscle group was then rectified. Artifacts were defined as trials containing fEMG activity exceeding 250 μV within the 10 second epoch. Artifacts (< 1% of trials) were removed. Artifact-free trials were aligned to a 250-ms baseline prior to the onset of the critical word. The analysis epoch consisted of eight sequential time-windows of 250 ms duration beginning at critical word onset. EMG activity was expressed as a percentage of baseline activity. An additional outlier exclusion procedure involved converting the percentage change from baseline to a z-standardized score, separately for each participant, condition, muscle group, and time-

window. Observations with a z-score value greater than 3 were removed. This procedure removed approximately 1.5 % of data points from the subsequent analyses.

All EDA analyses were performed using the Ledalab MATLAB toolbox (Benedek & Kaernbach, 2010, available from http://www.ledalab.de). The continuous EDA signal was downsampled to 32 Hz and a low-pass filter (5 Hz) was applied. The resulting signal was decomposed by continuous decomposition analysis (CDA; Benedek & Kaernbach, 2010). This method extracts the phasic information underlying the skin conductance response and aims at retrieving the signal characteristics of the underlying sudomotor nerve activity. Integrated phasic EDA response within an epoch 1-8 seconds after the critical word were analyzed. A minimum amplitude threshold of 0.05 μS was applied for the detection of EDA responses.

fEMG and EDA responses were statistically analyzed by means of repeated measures ANOVAs as described in more detail in the respective results sections. In the fEMG analyses, where appropriate, Greenhouse-Geisser corrected *p*-values and epsilon adjustments were applied.

Results

Control experiment. Mean fEMG activity was analyzed by means of separate repeated measures ANOVA for pictures and words and for corrugator supercilii and zygomaticus major muscles, respectively. The ANOVAs included the variables condition (neutral vs. negative) and time segment (8 levels: 0-250, 250-500, …, 1750-2000 ms). For words, fEMG corrugator activity was larger for negative than neutral words (102.00% vs. 99.99%), $F(1, 39) = 7.46$, $p < .01$, $\eta_p^2 = .16$. The main effect of time interval was significant, $F(7, 273) = 3.50$, $p < .05$, $\varepsilon = .38$, $\eta_p^2 = .08$, indicating a continuous decrease of fEMG activity over time. The Condition x Time interval

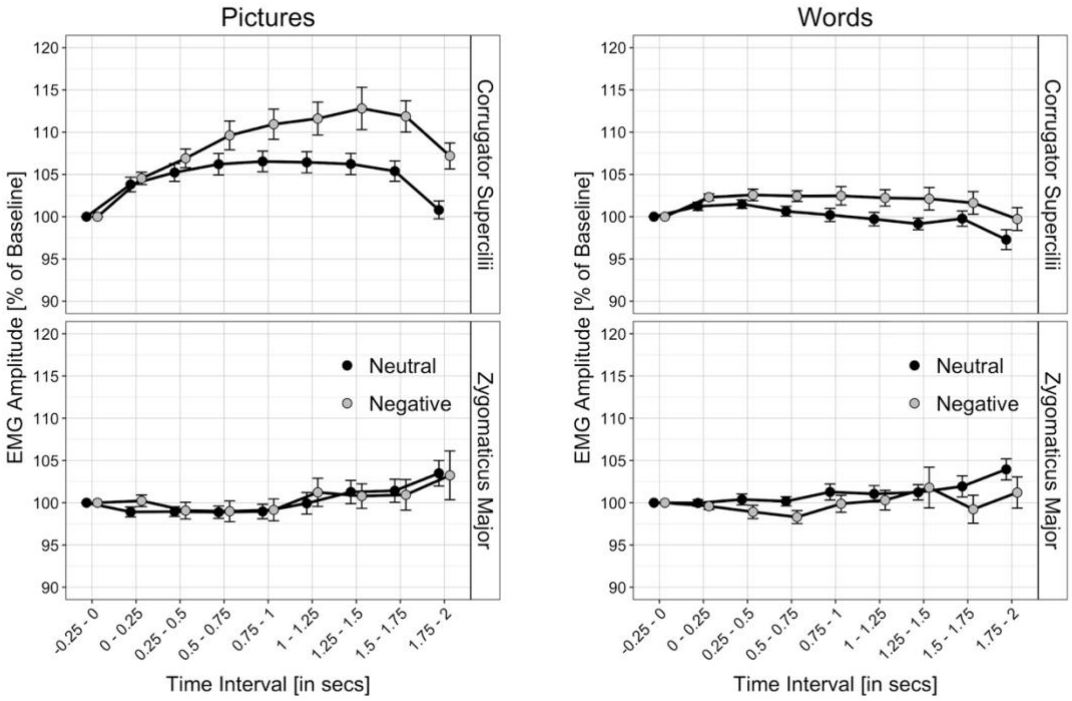
interaction was not significant, $F = 0.71$, $p = .58$. There were no significant effects for fEMG zygomaticus activity, all *p*s > .11.

For pictures, fEMG corrugator activity was larger for negative than neutral pictures (109.44% vs. 105.08%), $F(1, 39) = 16.12$, $p < .001$, $\eta_p^2 = .29$. The main effect of time interval was significant, $F(7, 273) = 8.83$, $p < .001$, $\varepsilon = .43$, $\eta_p^2 = .18$, indicating an initial increase and late decrease of fEMG activity. Furthermore, the Condition x Time interval interaction was significant, $F(7, 273) = 5.19$, $p < .01$, $\varepsilon = .46$, $\eta_p^2 = .12$, indicating an increasing condition effect between 250 and 2000 ms, all *F*s > 4.77, *p*s < .05. For fEMG zygomaticus activity, the main effect of time interval was significant, $F(7, 273) = 4.18$, $p < .05$, $\varepsilon = 31$, $\eta_p^2 = .10$, indicating a continuous increase of fEMG activity. However, the main effect of condition and the Condition x Time interval interaction were not significant, all *p*s > .68.

The statistical analysis of integrated phasic EDA responses corresponded to that of the fEMG data except that the variable time segment was dropped. Integrated phasic EDA responses differed neither between negative and neutral words (0.31 vs. 0.34 μS), $F(1, 39) = 0.65$, $p = .43$, nor between negative and neutral pictures (0.28 vs. 0.28 μS), $F(1, 39) = 0.21$, $p = .65$.

Discourse comprehension experiment. Mean fEMG activity was analyzed by means of separate repeated measures ANOVA for morality and emotion materials and for corrugator supercilii and zygomaticus major muscles, respectively. The ANOVAs included the variables condition (moral vs. immoral, neutral vs. negative) and time segment (8 levels: 0-250, 250-500, …, 1750-2000 ms). The significance level was set to alpha = .05.

Figure 2. fEMG corrugator and zygomaticus response in the control experiment as a function of material (picture vs. word), condition (neutral vs. negative), and time window. Error bars reflect standard error of the mean.

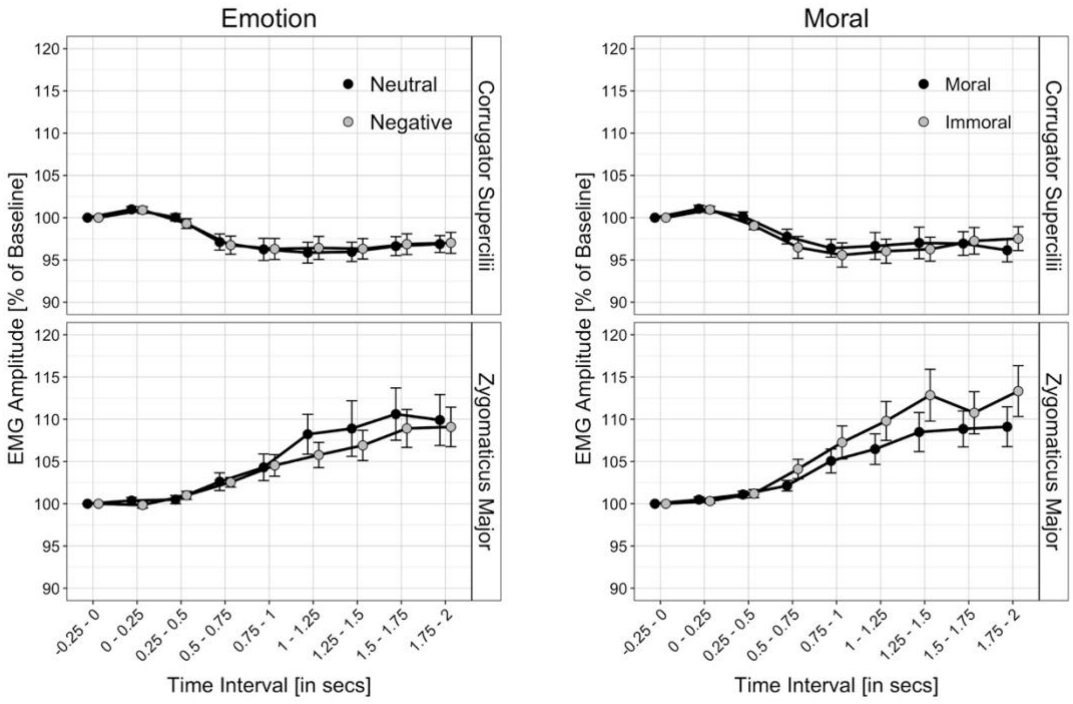


For morality materials, fEMG activity revealed no significant condition-related main or interaction effects, all *p*s > .12. For both fEMG corrugator and zygomaticus activity, only the main effect of time interval was significant, *F*s(7, 273) ≥ 6.60, *p*s < .01, εs ≤ .27, η_p^2s ≥ .14, indicating across time intervals a decrease and increase of fEMG activity, respectively (cf. Figure 3).

For emotion materials, the analysis of fEMG activity revealed no significant condition-related main or interaction effects, all *p*s > .53. As found for morality materials, only the main effect of time interval was significant for both fEMG corrugator and zygomaticus activity, *F*s(7, 273) ≥ 10.24, *p*s < .01, εs ≤ .33, η_p^2s ≥ .21, due to a decrease and increase of fEMG activity across time intervals, respectively (cf. Figure 3).

The ANOVA for integrated phasic EDA responses revealed no significant EDA differences between negative and neutral scenarios (0.36 vs. 0.36 μS), $F(1, 39) = 0.02$, $p = .88$, or between immoral and moral scenarios (0.33 vs. 0.35 μS), $F(1, 39) = 0.42$, $p = .52$.

Figure 3. fEMG corrugator and zygomaticus response in the comprehension experiment as a function of material (emotion vs. morality), condition (neutral vs. negative, moral vs. immoral), and time window. Error bars reflect standard error of the mean.



Discussion

Experiment 2 was identical to Experiment 1 except that fEMG and EDA instead of ERPs were measured in order to examine embodied responses to morality and emotion materials. Additionally, participants performed a control experiment to assess the

sensitivity of fEMG and EDA measures to negative and neutral stimuli that differed with respect to arousal and valence scores for both picture and word sets (cf. Table 2).

Crucially, in the control experiment, fEMG corrugator activity was larger for negative than neutral words irrespective of the analysis time window, whereas for pictures, this valence effect increased from 250 ms after stimulus onset up to 2000. In line with the study of Larsen et al. (2003), these findings suggest that the measurement of corrugator activity can be employed to infer the embodied response associated with the processing of negative compared to neutral stimuli even if participants merely attend to the stimuli by occasionally performing an emotion-unrelated task. In contrast, negative versus neutral words and pictures did not differentially influence fEMG zygomaticus activity in accord with similar previous reports demonstrating that corrugator activity but not zygomaticus activity sensitively reflects the differential processing of negative versus neutral pictures and words (e.g., Larsen et al., 2003; Tan et al., 2012). Interestingly, and in contrast to the findings of Niedenthal et al. (2009), corrugator activity was modulated by stimulus valence even though participants judged emotion-unrelated stimulus aspects. Therefore, we assume that the recording of fEMG corrugator activity is sufficiently sensitive to examine whether negative versus neutral scenarios as well as immoral versus moral scenarios also elicit embodied responses during emotional discourse comprehension.

However, and in contrast to two previous studies concerned with discourse-based emotional inferences (Bartholow et al., 2001; Thompson et al., 2016), negative versus neutral and also immoral versus moral items did not differentially influence fEMG activity of either the corrugator or zygomaticus muscles. That no reliable effect on fEMG zygomaticus activity was found is understandable given the results of the control experiment and reports in the literature, suggesting the relative insensitivity of

the zygomaticus to negative affect (e.g., Larsen et al., 2003; Tan et al., 2012). Since there was no sign of an fEMG corrugator effect, we are left to conclude that discourse-based emotional inferences for the present task and materials are unlikely to involve affective simulations.

The analysis of EDA did not provide any evidence for the occurrence of embodied responses to emotion stimuli differing in their arousal value. That is, in the comprehension experiment there were no differential EDA responses to the different emotion and morality materials. This zero-effect accords with findings of Thompson et al. (2016), who showed the phasic EDA response not to be influenced by emotional discourse comprehension. Also, the control experiment demonstrated that the emotional content of pictures or words did not automatically elicit such an activation within the autonomic nervous system. Previous research, however, indicated that EDA responses increased in magnitude with the degree of arousal of unpleasant or pleasant pictures that participants were asked to affectively judge (Bradley, Codispoti, Cuthbert, & Lang, 2001). Hence, we speculate that the present zero-effect in EDA is due to the only moderately different arousal values of our language materials and the fact that participants performed an emotion-unrelated task.

General Discussion

The aim of the present study was to advance our understanding of discourse-based emotion comprehension using a passive reading for comprehension task. Specifically, in a first experiment we investigated whether randomly intermixed emotion and morality scenarios underwent an (implicit) affective categorization at the level of brain processing. Based on the assumption that both emotion and morality materials are affectively evaluated even when the task did not demand so, we expected an enhanced

LPP for negative compared to neutral and also for immoral compared to moral scenarios (Experiment 1). A second experiment addressed the issue of whether discourse-based emotional inferences lead to embodied responses. We hypothesized that if emotional comprehension involves affective simulations (e.g., Niedenthal, 2007), this should be indicated by increased fEMG activity or EDA responses (Experiment 2) to both negative and immoral items compared to neutral and moral items.

Crucially, in Experiment 1 we obtained ERP evidence in accord with the assumption that incoming linguistic information is implicitly evaluated concerning its emotional content. First, LPP amplitude was larger for negative compared to neutral scenarios between 300-700 ms after the onset of the critical word in the target sentence. This particular finding replicates the LPP effect reported by Kunkel et al. (2018) when participants performed affective judgments to emotion materials that were presented in isolation rather than intermixed with morality materials. It is important to note that since the target sentences were identical for negative and neutral scenarios, the present LPP effect cannot be attributed to lexical differences of critical target words but rather demonstrates a discourse-based influence even when the task did not explicitly demand explicit affective judgments. Moreover, given that discourse contexts were only moderately constraining and critical words did not differ in cloze probability or their semantic relatedness (cf. Kunkel et al., 2018), it seems unlikely that the LPP effect is due to word-based expectancies. Therefore, we take the larger LPP to negative than neutral items to reflect the implicit affective categorization of linguistic input.

A second important ERP finding concerns the larger LPP between 300-500 ms, and as a trend between 500-700 ms, for immoral than moral items. This finding replicates the similar LPP effect reported by Leuthold et al. (2015) for a comprehension task in which morality materials were intermixed with emotion-neutral world knowledge

materials. Thus, it appears that immoral compared to moral scenarios contain more salient emotional information that is rapidly accessed and evaluated during discourse comprehension, at least when explicit moral judgments are not required (cf. Kunkel et al., 2018). It should be remembered that immoral and moral conditions used identical critical words and did not reliably differ regarding cloze probability and semantic relatedness. Therefore, as for emotion materials, we would rule out word-based lexical and also expectancy effects as an account of the LPP effect to morality items. Instead, we assume that morality items undergo an implicit affective categorization also when these items are presented together with affectively more salient emotion materials, extending the similar LPP findings of Leuthold et al. (2015) when morality items were intermixed with emotion-neutral world knowledge materials. Together, the LPP findings from Experiment 1 accord with those reported in our previous studies, in which participants read for comprehension (Leuthold et al., 2015) or explicitly evaluated the emotional content of language input (Kunkel et al., 2018), corroborating the view that incoming linguistic information is affectively evaluated by default.

Moreover, it is also worth mentioning that no N400 effects were observed. In previous discourse-based comprehension studies on emotion processing, a larger N400 was obtained when discourse contexts were strongly constraining and expectancies about a forthcoming emotion word were violated (e.g., León et al., 2010; Leuthold et al., 2012). This is also true for an ERP study conducted by Van Berkum, Holleman, Nieuwland, and Otten (2009) that was concerned with the processing of moral-laden statements. When such statements were inconsistent rather than consistent with the personally held values of participants (e.g., "If my child were homosexual, I'd find this *hard/easy* to accept"), a slightly larger N400 occurred followed by a larger LPP. Similarly, larger N400 and LPP components were observed to positive or negative

words following an incongruent (neutral) discourse context when participants read for comprehension (Delaney-Busch & Kuperberg, 2013; Holt et al., 2009). These N400 findings suggest that violation of emotion-related expectancies about forthcoming emotion words lead to incongruency effects that are associated with a higher demand for semantic processing. Having said this, one might wonder why Bartholow and colleagues (2001) in their study on intentional trait inferences during discourse comprehension found only a larger LPP but not N400 to emotion-related trait violations. We assume that the instruction to form explicit trait inferences might have changed task characteristics in such a way in this study that the processing of the emotional content of language input became prominent. Thus, compatible with our interpretation, the LPP observed by Bartholow and colleagues might reflect the discourse-based affective evaluation of trait-consistent and inconsistent scenarios.

With regard to our second research question, namely, whether discourse-based emotional comprehension involves affective simulations, the study of Bartholow and colleagues (2001) is relevant as well. They reported stronger fEMG corrugator activity immediately after presentation of the final affective word that described a trait-inconsistent rather than trait-consistent negative behavior. For words describing positive behaviors, such a trait-consistency effect was absent, in contrast to their LPP findings. In contrast, the present Experiment 2 revealed no reliable fMEG differences in either corrugator or zygomaticus activity and also not in phasic EDA responses, despite the fact that negative versus neutral and also immoral versus moral scenarios differed with regard to their rated valence and arousal. In the light of the findings of the control experiment, which revealed that fEMG corrugator activity reliably reflected the emotion word effect when participants performed an emotion-unrelated judgment task, it appears unlikely that the absence of an emotion-related effect on fEMG corrugator activity in

the comprehension experiment is due to insufficient sensitivity of this measure in revealing embodied responses. Rather, the findings in Experiment 2 appear to suggest that affective simulations are not involved during discourse-based processing of morality and emotion scenarios.

Of course, as Thompson and colleagues (2016) found both fEMG corrugator and zygomaticus activity to be modulated by the emotional processing of language input during irony comprehension, and also Bartholow et al. (2001) found larger fEMG corrugator activity to discourse-related emotion word violations, one might wonder which factors would explain the absence of fEMG effects in Experiment 2. Given the differences regarding tasks, procedures, and language materials used, it is difficult to come up with a straightforward explanation for the mixed outcomes across studies. However, we speculate that the salience of the emotional content of language materials might play a crucial role. For instance, it could be argued that emotional salience mattered in Bartholow et al.'s study, since they employed a task that might have promoted the explicit affective processing of discourse scenarios. Thus, it is an important goal for future research to test whether (embodied) affective simulations are activated if demanded by the task. For instance, if participants were asked to perform explicit affective judgments as in Kunkel et al. (2018), one might expect that embodied responses are elicited during the processing of negative or immoral compared to than neutral and moral items, indicating the involvement of affective simulations in emotional discourse comprehension.

Conclusions

In conclusion, the present ERP findings extend previous work (Kunkel et al., 2018; Leuthold et al., 2015) by providing clear ERP evidence for the implicit affective

evaluation of incoming emotional information even when no explicit affective judgment task is required. In addition, such an affective evaluation of moral transgression items was demonstrated also when embedded within other emotionally salient but morality-unrelated materials. Finally, we found no evidence for a possible role of affective simulations during the discourse-based processing of emotion and morality materials. It remains an important task for future studies to investigate whether embodied simulations during discourse comprehension depend on the salience and task-relevance of the emotional content of language materials.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660.

Bartholow, B. D., Fabiani, M., Gratton, G., & Bettencourt, B. A. (2001). A psychophysiological examination of cognitive processing of and affective responses to social expectancy violations. *Psychological Science, 12,* 197-204.

Benedek, M. & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190,* 80-91.

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion, 1,* 276-298.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433-436.

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language, 122,* 211-226.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences, 11,* 97-104.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159-181). New York, NY: Cambridge University Press.

Delaney-Busch, N., & Kuperberg, G. (2013). Friendly drug-dealers and terrifying puppies: Affective primacy can attenuate the N400 effect in emotional discourse contexts. *Cognitive, Affective, & Behavioral Neuroscience, 13,* 473-490.

Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology, 27,* 481–494.

Dudschig, C., Mackenzie, I. G., Strozyk, J., Kaup, B., & Leuthold, H. (2016). The sounds of sentences: Differentiating the influence of physical sound, sound imagery, and linguistically implied sounds on physical sound processing. *Cognitive, Affective, & Behavioral Neuroscience, 16*, 940-961.

Fields, E. C., & Kuperberg, G. R. (2016). Dynamic effects of self-relevance and task on the neural processing of emotional words in context. *Frontiers in Psychology, 6*, 2003.

Filik, R., Hunter, C. M., & Leuthold, H. (2015). When language gets emotional: Irony and the embodiment of affect in discourse. *Acta Psychologica, 156*, 114-125.

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in Brain Research, 156*, 185-203.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology, 23*, 567-589.

Gable, P. A., Adams, D. L., & Proudfit, G. H. (2015). Transient tasks and enduring emotions: the impacts of affective content, task relevance, and picture duration on the sustained late positive potential. *Cognitive, Affective, & Behavioral Neuroscience, 15*, 45-54.

Gibbs, R. W. (2003). Embodied experience and linguistic meaning. *Brain and Language, 84*, 1-15.

Glenberg, A.M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes, 28*, 1-26.

Hajcak, G., MacNamara, A., Foti, D., Ferri, J., & Keil, A. (2013). The dynamic allocation of attention to emotion: simultaneous and independent evidence from the late positive potential and steady state visual evoked potentials. *Biological Psychology, 92*, 447-455.

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology, 35*, 129-155.

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. *The Oxford handbook of event-related potential components*, *441*, 474.

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic use of botulinum toxin-A affects processing of emotional language. *Psychological Science, 21,* 895-900.

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14*, 436-441.

Holt, D. J., Lynn, S. K., & Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *Journal of Cognitive Neuroscience, 21*, 2245-2262.

Ito, T. A., & Cacioppo, J. T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology, 36*, 660-676.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex, 48,* 805-25.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, 36 ECVP Abstract Supplement.

Kunkel, A., Filik, R., Mackenzie, I. G., & Leuthold, H. (2018). Task-dependent evaluative processing of moral and emotional content during comprehension: An ERP study. *Cognitive, Affective, & Behavioral Neuroscience, 18*, 389-409.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621-647.

Lai, V. T., Hagoort, P., & Casasanto, D. (2012). Affective primacy vs. cognitive primacy: Dissolving the debate. *Frontiers in Psychology, 3*, 243.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). Technical report A-8. *University of Florida*.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology, 40*, 776-785.

León, I., Díaz, J. M., de Vega, M., & Hernández, J. A. (2010). Discourse-based emotional consistency modulates early and middle components of event-related potentials. *Emotion, 10*, 863-873.

Leuthold, H., Filik, R., Murphy, K., & Mackenzie, I. G. (2012). The on-line processing of socio-emotional information in prototypical scenarios: inferences from brain potentials. *Social Cognitive and Affective Neuroscience, 7*, 457-466.

Leuthold, H., Kunkel, A., Mackenzie, I. G., & Filik, R. (2015). Online processing of moral transgressions: ERP evidence for spontaneous evaluation. *Social, Cognitive, and Affective Neuroscience, 10*, 1021-1029.

Niedenthal, P. M. (2007). Embodying emotion. *Science, 316*, 1002-1005.

Niedenthal, P. M., Winkielman, P., Mondillon, L., & Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology, 96*, 1120-1136.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*,437-442.

Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cognition and Emotion, 21*, 1212-1237.

Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., & Traue, H. C. (2012). Repeatability of facial electromyography (EMG) activity over corrugator supercilii and zygomaticus major on differentiating emotions. *Journal of Ambient Intelligence and Humanized Computing, 3*, 3-10.

Thompson, D., Mackenzie, I. G., Leuthold, H., & Filik, R. (2016). Emotional responses to irony and emoticons in written language: evidence from EDA and facial EMG. *Psychophysiology, 53*, 1054-1062.

Van Berkum, J. J. A., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or Wrong? The Brain's Fast Response to Morally Objectionable Statements. *Psychological Science, 20*, 1092-1099.

van Boxtel, A. (2010). Facial EMG as a tool for inferring affective states. In A. J. Spink, F. Grieco, O. E. Krips, L. W. S. Loijens, L. P. J. J. Noldus, and P. H. Zimmerman (Eds.), *Proceedings of measuring behavior* (pp.104-108). Wageningen, The Netherlands: Noldus Information Technology.

Võ, M. L., Conrad, M., Kuchinke, L., Urton, K., Hofmann, M. J., & Jacobs, A. M. (2009). The Berlin affective word list reloaded (BAWL-R). *Behavior Research Methods, 41*, 534-538.

Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist, 39*, 117-123.

4 Discussion

Experimental research has shown that emotions play an important role in moral judgements (for a review see Avramova & Inbar, 2013). The exact function and the time point when they become relevant during the judgement process has not been clarified. The ERP study by Leuthold et al. (2015) provided initial evidence for affective processing during moral language comprehension. They examined the mental mechanisms during reading for comprehension of everyday morality scenarios. Assuming that moral violations are related to world knowledge violations that trigger an N400 at the discourse level (Filik & Leuthold, 2008; Hagoort et al., 2004), Leuthold et al. (2015) additionally presented vignettes of world knowledge violations in order to test for the presence of the N400 effect. In contrast to their expectations, ERPs showed a larger positivity for immoral as compared to moral scenarios. Leuthold et al. (2015) related this positivity to the LPP component which was taken to reflect an implicit evaluative categorisation process of incoming moral information (cf. Cacioppo et al., 1993; Ito & Cacioppo, 2000). The present doctoral thesis was motivated by theories of emotional language comprehension research and the study of Leuthold et al. (2015). Its aim was to venture into the research gap by trying to define the specific role of affective processing during moral language comprehension. In the following section, I will refer to three research questions the present doctoral thesis aimed to answer: First, does moral information undergo affective processing? Second, does moral content of incoming linguistic information trigger embodied processing? Third, does moral information processing differ from the processing of emotional information? Before tackling these questions in detail, I will summarise the studies conducted to answer these questions.

4.1 Summary of studies

In three studies it was tested whether incoming moral information undergoes an implicit affective evaluation. The paradigm of Leuthold et al. (2015) was used, yet, instead of presenting world knowledge vignettes, scenarios describing emotional events were employed. All three studies, investigated the processing of scenarios describing everyday moral content and scenarios describing emotional situations. It is important to note that the emotional and moral meaning of each scenario became apparent only after the target sentence was integrated into a preceding, ambiguous context. Participants were asked to perform explicit judgements (Study 1 and 2) or read for comprehension (Study 3). In order to assess on-line processing of moral and emotional content, behavioural (RT, accuracy) and psychophysiological measures (ECG, EDA, facial EMG, ERPs) were analysed. The major

aim was to reveal the neural mechanisms as well as embodied responses to morality and emotion scenarios in discourse comprehension.

Study 1 investigated in two ERP experiments the nature and time course of evaluative processing of short moral and emotional scenarios with regard to brain processes. Participants judged whether they found the described moral situation either morally acceptable or not and the emotional situation as either emotionally moving or not (Experiment 1), or made only emotional judgements to both types of scenarios (Experiment 2). As in previous studies using a similar approach (Leuthold et al., 2015; Van Berkum et al., 2009), it was assumed that moral and emotional materials trigger affective evaluations irrespective of the task. Critical words were expected to trigger an enhanced LPP for immoral as compared to moral scenarios as well as for negative as compared to neutral emotional materials, irrespective of the judgement task. However, the results of Study 1 indicated that only for emotional judgements the same LPP effect was found for both types of scenarios (Experiment 2). When making moral judgements on morality materials (Experiment 1), an anterior negativity for immoral as compared to moral items was found.

Study 2 served to investigate emotion simulation indicated by embodied changes in facial EMG, phasic EDA, and ECG triggered by the same morality and emotion scenarios from Study 1. The fEMG measurement included activity changes of the corrugator and levator muscle, indicating a negative affect, and of the zygomaticus muscle, indicating a positive affect (Larsen et al., 2003). Participants judged whether they found the described moral and emotional situation emotionally moving or not like in Experiment 1 of Study 1. To test the sensitivity of all three peripheral psychophysiological measures, a control experiment was conducted where participants were presented with emotional versus neutral pictures and words. Based on previous studies (for a review see Bradley & Lang, 2000), we expected clear emotional responses at least to visual stimuli. Phasic EDA was not sensitive to control and experimental manipulations. For pictures, the heart rate measure showed the expected pattern with a stronger initial decrease for negative as compared to neutral pictures (Bradley & Lang, 2000). Also for words of the control experiment and morality materials, an initial decrease of heart rate could be observed, but not differentiated between conditions. Moreover, in contrast to our expectations for emotion materials, heart rate decreased stronger for neutral than negative scenarios. As could be expected with regard to the fEMG, negative compared to neutral control items triggered increased activation of the corrugator and levator muscles. For morality materials, levator and corrugator were sensitive to moral transgressions in line with findings by Cannon et al. (2011) and

Krumhuber et al. (2018). Crucially, zygomaticus activity was stronger for immoral as compared to moral scenarios. For emotional materials, fEMG was not sensitive to the valence of emotional scenarios.

Study 3 was inspired by the results of Study 1, where the processing of incoming moral information differed depending on the task participants performed. It was still unclear what happens if participants merely read morality and emotion scenarios for comprehension without an explicit moral or emotional judgement task. Hence, Study 3 was motivated by the assumption that the global context could influence whether an affective versus semantic-cognitive analysis is prioritised (e.g., Delaney-Busch & Kuperberg, 2013; Lai et al., 2012). The global context can be considered here as the task demands or the use of emotion-related versus emotion-neutral linguistic stimuli among which morality scenarios are embedded. It was unclear whether emotion materials are still implicitly evaluated in a task context where materials were not explicitly judged according to their affective (Study 1 and 2) or moral content (Study 1). Moreover, Leuthold et al. (2015) presented morality scenarios in an emotion-irrelevant context, for which world knowledge materials and morality materials were presented in an intermixed manner. Thus, it is also unclear how morality materials are processed in a global emotion-related context where emotion and morality materials are intermixed. In two experiments, participants performed a passive reading-for-comprehension task with randomly intermixed morality and emotion materials while ERPs (Experiment 1), fEMG and EDA (Experiment 2) were recorded. Again, for Experiment 2, we additionally conducted a control experiment with the same pictures and words of Study 2 in order to test the sensitivity of the peripheral psychophysiological measures. The ERP results of Experiment 1 indicated an LPP effect for both, immoral and moral scenarios as well as for negative and neutral scenarios. The results of Experiment 2 revealed no reliable fMEG differences in either corrugator or zygomaticus activity and also not in phasic EDA responses, although corrugator activity was larger for negative than for neutral words and pictures in the control experiment.

4.2 Processing of moral and emotional materials

4.2.1 The processing of emotional information

In order to answer the first research question and to see whether moral information undergoes affective processing, it is necessary to first know how emotionally loaded materials are processed. Based on existing evidence (cf. Fischler & Bradley, 2006; Hajcak et al., 2010), the present doctoral thesis assumes that linguistic information, describing emotionally-

relevant events, is affectively processed and evaluated. It is important to note that, to my knowledge the present studies investigated for the first time whether the emotional meaning of identical target sentences is differently processed when emotions need to be inferred from the context. That is, emotional discourse comprehension was examined under conditions that excluded possible influences of differing emotion words (e.g., Delaney-Busch & Kuperberg, 2013; Holt et al., 2009), and also of emotional congruency effects due to the use of emotionally constraining contexts (e.g., León et al., 2010; Leuthold et al., 2012).

A first key ERP finding is that in three experiments (Study 1 and 3), the same LPP effect was replicated for emotional negative compared to neutral scenarios under different circumstances: when items were presented blockwise during an emotional judgement task, as well as when items were presented intermixedly within morality items during a reading-for-comprehension task. The observed LPP effect for negative compared to neutral scenarios is taken to reflect the discourse-based influence in emotional language comprehension rather than resulting from lexical or emotion related differences between target words. This is because the target sentences were identical for negative and neutral scenarios. The LPP effect in conjunction with the absence of an N400 effect accords with the view that participants focused on the processing of the emotional rather than the semantic content (e.g., Delaney-Busch & Kuperberg, 2013; Lai et al., 2012), indicating the enhanced affective evaluation of negative than neutral scenarios. In this respect, Study 1 and 3 extend previous ERP studies investigating discourse-based emotion effects.

One might argue, however, that the LPP is related to the P300 component, which has been taken to reflect working memory operations (e.g., Donchin & Coles, 1988). It is still an open issue whether, and in which way, the LPP is related to the P300 component. The P300 is typically associated with participants' attention to stimuli (Nieuwenhuis, Aston-Jones, & Cohen, 2005) and its amplitude is known to be inversely related to the prior and also the subjective probability of a given stimulus event, task demands, and its significance. The late ERP positivity found in the present studies does not, however, simply reflect stimulus novelty, low-level perceptual differences (since identical target words were used), or violations of expectation (target words do not differ with regard to cloze probability and semantic relatedness) (Hajcak et al., 2010). In accordance with previous research on emotional stimulus processing (Fields & Kuperberg, 2012), and displaying a different ERP pattern for world knowledge vignettes (Leuthold et al., 2015), I take the long-lasting positivity for negative scenarios to relate to the LPP rather than the P300 component, indicating the affective evaluation of stimuli (Cacioppo et al., 1993; Ito & Cacioppo, 2000).

A second goal was to answer the questions whether the processing of emotional materials is embodied and hence reflected by peripheral psychophysiological measures. That is, it was examined whether discourse-based comprehension of emotional language involves affective simulations. The control experiments of Study 2 and 3 indicated that negative compared to neutral words and pictures do not influence fEMG zygomaticus activity differently. This is in line with similar reports showing that zygomaticus is sensitive to positive stimuli only (e.g., Larsen et al., 2003; Tan et al., 2012). By contrast, corrugator and levator activity was larger for both negative than to neutral words and pictures. These results indicate the recording of corrugator and levator activity is sufficiently sensitive to examine embodied responses to negative stimulus input. Therefore, it should also be sensitive to embodied responses elicited during emotional discourse comprehension of negative versus neutral scenarios as well as immoral versus moral scenarios. Hence, the analysis of corrugator and levator activity can be employed to reveal embodied responses associated with the processing of negative or neutral stimuli, even if participants merely attend to the stimuli by occasionally performing an emotion-unrelated task as in Study 3. Phasic EDA was not sensitive at all to experimental manipulations, neither in the control experiment nor in the comprehension experiment. With the exception of the ECG results in Study 2, emotion materials revealed no reliable embodied responses.

In summary, the emotion text materials were indeed appropriate to trigger implicit affective evaluations with regard to brain processes, but not with regard to embodied simulation of concrete emotional states. For the present emotion materials, the results of peripheral psychophysiological measures provide no support for the view that the same embodied processes that are engaged while experiencing emotions were also involved in understanding emotions during language comprehension.

4.2.2 Does moral information undergo affective processing?

With regard to brain processes, this research question can not be answered in a straightforward manner. If morality materials are affectively processed, one would expect an LPP effect as found for emotion materials, because both types of materials would require the same mental mechanism. That is, the LPP should be larger for emotionally more salient immoral scenarios like for negative emotional scenarios.

As expected on the basis of a previous study using a similar approach and materials (Leuthold et al., 2015), an LPP effect was elicited by immoral compared to moral scenarios. This was found in an emotional judgement task (Study 1) as well as in a passive reading-

for-comprehension task (Study 3). Moreover, behavioural rating and experimental data showed that participants judged immoral items as more emotionally moving than moral ones. Together, these results indicate the evaluative-affective categorisation of incoming moral information during discourse comprehension. Accordingly, the LPP effect could reflect the higher emotional intensity of our immoral scenarios. Alternativly, this effect could be driven by the motivational salience of the items (Hajcak et al., 2010).

Experiment 2 of Study 1 demonstrated that, other than for moral scenarios, the LPP for immoral scenarios was absent if participants explicitly focused on the moral content of incoming information when moral acceptability judgements were demanded. Instead, readers were more engaged in cognitive processing, as indicated by an anterior negativity for immoral scenarios. In other discourse comprehension studies, such an anterior negativity has been taken to index language-related working memory demands, that is, when alternative but likely text inferences have to be maintained or integrated simultaneously within the situation model (Baggio, Van Lambalgen, & Hagoort, 2008; Xiang & Kuperberg, 2015). Similarly, when explicit moral judgements are required, it appears possible that the integration of either the likely (anticipated) moral action or the unlikely immoral action then differentially taxes working memory (for similar reasoning, see Xiang & Kuperberg, 2015). Thus, it is conceivable that working memory load and the demands on integrating linguistic information into the situation model is higher in the case of immoral than moral items, giving rise to the enlarged anterior negativity. The assumption that the present anterior negativity effect is related to working memory functions implies that cognitive-semantic processing plays a major role when explicit moral judgements are required.

Also behavioural evidence indicates the task-dependent processing of morality materials. Thus, responses were slower for moral than emotional judgements, whereas the speed of emotional judgements was the same for moral and emotion materials. On the one hand, this result accords with the idea that a more complex and therefore time-consuming cognitive decision process underlies moral decision making than it is the case for emotional decision making. On the other hand, it also indicates that the way readers process information about the persons and events described in the text, the information they focus on and evaluate, depends on their specific goals. In addition, faster judgement responses to emotionally negative and immoral scenarios than to neutral and moral scenarios might be attributed to the fact that the former items are more salient.

In sum, the second key ERP finding of the present doctoral thesis concerns the task-

dependent processing of incoming moral information. An emotional judgement and a reading-for-comprehension task triggered affective processing of moral information as indicated by an LPP effect. The moral judgement task shifted the focus to the cognitive-semantic aspects of moral content indicated by the anterior N400. This provides clear evidence for the assumption that the processing of morality scenarios depends on the specific circumstances: cognitive processing when explicit moral judgements are required, and affective processing when the moral content is either implicitly processed during reading for comprehension or when the emotional content of morality scenarios is in focus. This finding accords with other discourse comprehension studies in which the LPP, and hence affective processing of linguistic input, is modulated by various variables, including the specific discourse context and task demands (e.g., Delaney-Busch & Kuperberg, 2013; Lai et al., 2012; Xiang & Kuperberg, 2015). It further accords with fMRI evidence that cognitive processes are more dominant when the task requires explicit moral judgements rather than merely the passive processing of moral content and vice versa (Sevinc & Spreng, 2014).

4.2.3 Does moral content of incoming linguistic information trigger embodied processing?

We examined the embodied processing of incoming moral and emotional content in two studies. ECG was sensitive to emotion materials, but was not reliably influenced by morality conditions. Phasic EDA was not sensitive to the present morality materials either. Given that EDA and ECG are established methods to measure emotional arousal (Bradley & Lang, 2000) it appears that present morality and emotion materials are presumably not very arousing during discourse comprehension.

By contrast, facial EMG yielded interesting results. As mentioned above, in both studies (Study 2 and 3) the control experiment showed that participants were sensitive to negative affect as indicated by increased corrugator and levator activity for negative compared to neutral words and pictures. Whereas, it was expected that immoral scenarios trigger negative emotions, corrugator activity was not or only slightly, though reliably, influenced by our experimental manipulations. This is surprising given that recent fEMG studies observed increased corugator activity when participants read or heard about moral transgressions and performed an emotional rating task ('t Hart, Struiksma, Van Boxtel, & Van Berkum, 2018; Cannon et al., 2011). Instead, levator activity was increased for moral transgressions when participants performed an emotion-related task in Study 2,

which in turn is in line with previous research (Cannon et al., 2011; Krumhuber et al., 2018). Krumhuber and colleagues found that only levator but not corrugator was sensitive to vignettes describing social-cultural norm violations compared to neutral vignettes. The authors took levator activity to indicate disgust when participants read about the described behaviour, suggesting that norm transgressions entail a moral component. Also Cannon and colleagues (2011) found that levator was sensitive to moral transgressions of the purity and fairness foundations, indicating disgust. Critically, according to the authors, they asked participants to make a moral judgement ("How negative or positive was this behaviour?"). In my opinion, this kind of judgement is not a moral one, but an emotional judgement, which is more in line with that of Study 2.

In addition, in Study 2, the zygomaticus was sensitive to experimental manipulations contrary to our expectations, showing stronger activation for immoral compared to moral scenarios. Assuming that zygomaticus activity indicates positive affect (Larsen et al., 2003), one may wonder why we found stronger activation of the zygomaticus given that immoral behaviour is assumed to trigger negative emotions. It is important to note, however, that in previous studies presenting visual stimuli, zygomaticus activity was ambiguous in neutral and negative emotional states (Tan et al., 2012). Also in case of complex linguistic stimuli zygomaticus was not a reliable indicator of valence (Larsen et al., 2003). For instance, smiling activity could be also interpreted as wry, sarcastic, and smirking which can be hardley interpreted as a complaisant positive affect ('t Hart et al., 2018). Given that corrugator and levator activity in the control experiment was increased to emotional-negative stimuli, the increased and simultaneous activation of zygomaticus and levator and corrugator for immoral compared to moral scenarios might suggest that the present zygomaticus activity results from cross-talks of nearby muscles that are associated with negative effect (Cannon et al., 2011; Tassinary, Cacioppo, & Vanman, 2007).

Together, the current increased activation of levator and corrugator fit with the assumption that moral transgressions elicit negative emotions which are simulated during reading. Moreover, it appears that an emotional judgement task enhanced the emotional salience of morality scenarios. In line with our ERP findings, embodied processing depended on the task participants had to perform. Extending Niedenthal et al. (2009), not only embodied processing of isolated emotional words but also emotional discourse comprehension depends on whether the task is related to the processing of the emotional content. We did not find evidence for the assumption that facial embodiment has a causal role in emotional language processing, at least as concerns the present materials (but see Havas et al., 2010;

Niedenthal et al., 2009), because embodied processing was absent when having no explicit judgement task. The assumption of a causal role of emotion simulation in language comprehension and its boundary conditions needs to be investigated in future research. To this end, the time course of embodied effects, especially whether the activation of fEMG either precedes or follows the judgement needs to be investigated. If the embodied response follows the judgement, this would also argue against the assumption that emotion simulation has a causal role, but could be taken as an epiphenomenon.

4.2.4 Implications of the present findings at the level of brain and body

The results of the present studies indicate that at the level of brain and body, the way incoming moral information is processed depends on the task participants perform. Moral information is not affectively judged per se, but involves a cognitive component, which dominates if demanded by the task. That is, if incoming moral information needed to be evaluated with regard to moral acceptability, cognitive processing dominates over affective processing.

This conclusion does not fit with SIM's assumption that automatic intuitions always precede cogntive processing in moral judgements (Haidt, 2001). However, it appears to accord with Greene's dual-process theory of moral judgement, which is more flexible than SIM, since it assumes that both affective and cognitive processes are involved in such judgements (Greene et al., 2001, 2004). As outlined earlier, Greene and colleagues examined the brain processes underlying people's moral judgement behaviour with regard to moral dilemma situations that differ in the degree of personal immediacy. Automatic affective processes and controlled cognitive processes work as two dissociable and antagonistic systems in moral judgements. Under some circumstances, emotional processes are faster and precede cognitive processing. However, Greene and colleagues made no assumptions about what exactly triggers a stronger negative emotional reaction. One idea was that the personal immediacy (personal, impersonal) affects whether affective or cognitive processing dominates. By now, evidence supported the involvement of cognitive processing (Greene et al., 2008; Suter & Hertwig, 2011), but explanations for the impact of emotions has been lacking.

Greene's theory was based on judgement behaviour in sacrificial moral dilemma situations. As mentioned in Chapter 1 (Section 1.2.2), Greene's materials suffer from several limitations. They were neither validated nor tested in regard to their material-specific linguistic effects. It is an important goal of the present doctoral thesis, that two stan-

dardised and text-based stimulus sets were employed to investigate emotional and moral cognition. For both sets, we offer a comprehensive description of psycholinguistic aspects and rating data (see Method section of Study 1). Emotion materials were employed to examine emotional language processing during discourse comprehension using moderately constraining contexts and identical target words. Together with the similarly constructed emotion materials, cognitive as well as emotional determinants of moral information processing could hence be investigated. Using these well constructed and everyday scenarios, we were able to provide evidence for the idea that affective and cognitive processes work as two separate systems, but the specific interplay is yet an open question.

4.2.5 Does moral information processing differ compared to the processing of emotional information?

Numerous researchers have already asked whether moral evaluations depend on the same general, unspecific mechanisms as other social judgements, or whether moral evaluations build on domain-specific mechanisms (cf. Cushman & Young, 2011; Waldmann et al., 2012). Based on the results of the present doctoral thesis, this question can not be finally answered. First, the major aim was to examine whether moral information undergoes affective processing and whether such processing is task-dependent. To this end, emotional language processing was investigated. Second, morality and emotion materials differ with regard to the wording of the critical sentences and are not matched regarding all potentially relevant word-level or discourse-level dimensions. This means, that it was not appropriate to subject ERP amplitudes for both materials to one common statistical analysis in order to directly compare material-specific ERP effects. Hence, it was decided to indirectly compare the respective material-specific ERP effects providing a first ground on research of emotional language comprehension.

In this respect, it is useful to distinguish between brain and embodied processing. With regard to brain processes, both materials were similarly affectively processed when participants performed an emotion-related task and when they merely read for comprehension. This was also supported by behavioural data that showed faster judgement responses to emotionally negative and immoral scenarios than neutral and moral scenarios, presumably because the former scenarios are more salient. This fits with the assumption that emotional stimuli capture attention and facilitate subsequent processing (Zajonc, 1984). In the case of an explicit moral judgement task, the moral content became more salient and the processing of incoming moral information changed accordingly. Again, behavioural

findings indicated that participants followed task instructions by answering the respective moral judgement questions in the majority of cases as expected. More importantly, the processing was influenced by both the task and the specific moral and emotional content of materials. Immoral items were judged as less acceptable (Experiment 1) and more emotionally moving (Experiment 2) than moral ones and negative items were judged as more emotionally moving than neutral ones. Following the behavioural evidence together with the LPP versus anterior negativity effects, it seems that different aspects of incoming moral information were relevant depending on the task demands. This leads to the assumption that the processing of moral information involves a cognitive component which is only relevant under some circumstances. That is, participants were more engaged in cognitive processing when demanded by the moral judgement task.

With regard to facial emotion simulation, a similar pattern was found. If there were no explicit task demands (i.e. when merely reading for comprehension), both materials were similarly processed and showed no embodied responses. If emotional judgements were demanded, where participants rated how emotionally moved they were, only morality materials provided evidence for the involvement of emotion simulation during language comprehension.

In an implicit (passive reading) task the processing of incoming emotion and morality information is similar, whereas, it differs for an explicit task (emotional compared to moral judgement task). Does the latter finding speak for a domain-specific processing of moral information? To answer the question about domain-specific moral mechanisms, it must be clarified which specific cognitive processes are involved in moral judgements with regard to the anterior negativity. This ERP component may reflect working-memory processes which would speak in favour of general cognitive processing (cf. Baggio et al., 2008; Xiang & Kuperberg, 2015), but it may also reflect a specific cognitive process regarding the moral domain. In my opinion, based on the results of the present studies, moral information processing engaged in general (working memory) mechanisms of social information processing. This is consistent with Greene et al. (2001) who found, increased activation in brain areas that are related to working memory operations during cognitive processing. Thus, what pays off specifically for morality is a stronger cognitive component in a moral judgement task compared to pure emotional scenarios. This anterior negativity and the circumstances under which this it is activated remains to be investigated.

4.3 Outlook: Open questions and further research directions

Although the present ERP results of Study 1 indicated the task-dependent processing of moral content, the cognitive component of moral information and its specific nature remains to be investigated. It is a major limitation of the present studies that they focused mainly on the affective aspects of moral information processing. As mentioned above, the anterior negativity might reflect cognitive processing but this is an isolated results which needs to be replicated. It should be an aim of further research to examine the functional significance of this ERP component and whether it reflects the domain-specific processing of moral content. This could be achieved by manipulating the emotional salience of morality materials (e.g., varying arousal) or by varying between different cognitive tasks (e.g., "Do you think this behaviour is conforming to the law?").

Another aspect of cognitive processing in moral information processing was concerned in a recent study of 't Hart et al. (2018). Participants read scenarios describing a moral or immoral protagonist, to whom happens a good or bad thing while recording fEMG corrugator activity. 't Hart and colleagues differentiated between two critical events. For example, if a narrative involves a protagonist acting immorally (first critical event), increased corrugator activity reflected negative affect. Whereas the subsequent critical event, for example something good happens to the immoral protagonist, leads to conflicting affect between language-driven simulation (positive) and fairness-based moral evaluation (negative). According to the authors, the pattern of corrugator activity depends on the moral status of the protagonist, which influences subsequent information processing. The results of 't Hart et al. (2018) are in line with the affective language comprehension model (ALC) most recently developed by Van Berkum (in press), which approaches the interfaces between language comprehension and emotion. It is based on the assumption that an emotional evaluation does not only take place in the mere comprehension of lexical meaning, but also, when readers evaluate this linguistic input with regard to their own beliefs. Van Berkum differentiates between meaning comprehension and evaluation, in which evaluation is the emotional reaction to representations that become available after understanding what is described in the narrative. The evaluation process implies controlled, cognitive processing of incoming information. According to the ALC, we could differentiate between two phases within the moral judgement process. First, the perception of an agent acting in a certain way, and second, the evaluation of the agent's action on the basis of the context-based situation model. As proposed by 't Hart et al. (2018), the results of Study 3 would then suggest that, only morality scenarios describe anything which is

worthy to cognitively evaluate. However, both Van Berkum (in press) and 't Hart et al. (2018) make no assumptions about how these two processes exactly interact and remains to be investigated.

Another study concerning the relevance of cognitive processing of incoming moral information is given by Yang et al. (2013). By using ERPs, they identified that physical compared to moral disgust occurred at different phases. The evaluation of moral information was processed prior to that of physical disgust. Therefore, Yang et al. (2013) argued that the processing of moral content precedes to those of emotional content. Also, fMRI evidence showed that the neural networks activated by basic disgust feelings compared to disgust elicited by moral transgressions covered distinct brain regions (Schaich Borg, Lieberman, & Kiehl, 2008; Moll et al., 2002, 2005).

A further limitation concerns the type of judgement task which was chosen for Study 2. As Niedenthal et al. (2009) proposed, embodied processing depends on the task which should be related to emotional content. Their fEMG effects were absent when participants judged emotional words in an lexical decision task. This has not yet been tested for fEMG, EDA and ECG measures for morality materials. Study 2 provided evidence, that an explicit task enhances the emotional salience of morality scenarios with regard to embodied responses, but it is unclear, whether they are still present if participants perform a moral judgement task as in Experiment 1 of Study 1. For emotion materials, an emotional judgement task did not affect the emotional salience of materials with regard to embodied responses anyway.

Based on the present three studies, I can only speculate about the exact relation between brain processes and physiological embodied responses. In the first case, the simulation of emotions during the processing of incoming information starts with the reactivation of experienced-based representations in the brain (Damasio, 1994). The brain sends a signal to innervate the concerned muscle and embodied responses would follow brain processing. Alternatively, both brain and embodied processes appear simultaneously. The time course of ERPs and fEMG responses needs to be clarified to substantiate assumptions about the causal role of embodied processing of incoming information. If responses in facial EMG follow the moral judgement, emotion simulation could be taken as an epiphenomenon. Thus, it should be an aim of future studies to combine all psychophysiological methods in a single experiment to investigate the various signals with regard to the exact time course of brain and embodied responses.

Conclusion

Immoral actions of others that clash with our moral beliefs are very likely to be experienced as emotionally aversive. This thesis provides evidence that we are able to rapidly evaluate or judge behaviour as good or bad, as positive or aversive, or as right and wrong. The processing of morally laden language has emerged as a new domain in emotional language comprehension. To make a judgement about the valence of someone's action incoming information is affectively processed. To consider a moral judgement, cognitive processes are needed. How incoming linguistic moral information is processed depends on a number of factors like the contextual information, the circumstances that contribute to a moral decision, and the task which participants have to perform. To conclude, emotional and cognitive processing both contribute to moral information processing.

The main achievement of the present doctoral thesis is that it extends previous studies investigating discourse-based effects in emotional and moral language comprehension. The present work provides a new set of text-based scenarios to study emotional and moral cognition. I assume that these text materials recruit the affective system and are motivationally relevant. In the present thesis, I have provided evidence for discourse-based meaning construction of emotional content as well as for discourse-based meaning construction of moral content. The meaning construction of moral information is task-dependent. Future research would need to take into account the potential impact of task demands when elucidating the nature of the potential cognitive and affective processes contributing to moral evaluations and decisions.

References

Amodio, D. M., Bartholow, B. D., & Ito, T. A. (2013). Tracking the dynamics of the social brain: ERP approaches for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *9*, 385–393.

Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, *10*, 229–240.

Avramova, Y. R., & Inbar, Y. (2013). Emotion and moral judgment. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*, 169–178.

Baggio, G., Van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language*, *59*, 36–53.

Barsalou, L. W. (1999). Language comprehension: Archival memory or preparation for situated action? *Discourse Processes*, *28*, 61-80.

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, *2*, 716–724.

Bartholow, B. D., & Amodio, D. M. (2009). Brain potentials in social psychological research: A brief review and tutorial. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in social neuroscience* (pp. 198–232). New York: Guilford Press.

Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, *8*, 536–554.

Baumrind, D. (1986). Sex differences in moral reasoning: Response to Walker's (1984) conclusion that there are none. *Child Development*, *57*, 511–521.

Bradley, M. M., & Lang, P. (2007). Emotion and motivation. In I. J. Cacioppo, L. Tassinary, & G. B. (Eds.) (Eds.), *Handbook of psychophysiology* (pp. 581–607). New York: Cambridge University Press.

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. *Cognitive Neuroscience of Emotion*, *25*, 49–59.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*, 34–44.

Cacioppo, J. T., Crites, S. L., Berntson, G. G., & Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should they not affect the event-related brain potential? *Psychological Science*, *4*, 108–112.

Cannon, P. R., Schnall, S., & White, M. (2011). Transgressions and expressions: Affective facial muscle activity predicts moral judgments. *Social Psychological and Personality*

Science, *2*, 325–331.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*, 215–224.

Chen, P., Qiu, J., Li, H., & Zhang, Q. (2009). Spatiotemporal cortical activation underlying dilemma decision-making: an event-related potential study. *Biological Psychology*, *82*, 111–115.

Choe, S. Y., & Min, K.-H. (2011). Who makes utilitarian judgments? the influences of emotions on utilitarian judgments. *Judgment and Decision Making*, *6*, 580–592.

Christensen, J. F., Flexas, A., Calabrese, M., Gut, N. K., & Gomila, A. (2014). Moral judgment reloaded: a moral dilemma validation study. *Frontiers in Psychology*, *5*, 607.

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and Language*, *122*, 211–226.

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, *11*, 393–399.

Clore, G. L., & Huntsinger, J. R. (2009). How the object of affect guides its impact. *Emotion Review*, *1*, 39–54.

Clore, G. L., Wyer Jr, R. S., Dienes, B., Gasper, K., Gohm, C., & Isbell, L. (2001). *Affective feelings as feedback: Some cognitive consequences.* New York: Lawrence Erlbaum Associates Publishers.

Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: a process dissociation approach. *Journal of Personality and Social Psychology*, *104*, 216–235.

Crites, S. L., Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1995). Bioelectrical echoes from evaluative categorization: II. A late positive brain potential that varies as a function of attitude registration rather than attitude report. *Journal of Personality and Social Psychology*, *68*, 997–1013.

Cushman, F., & Young, L. (2011). Patterns of moral judgment derive from nonmoral psychological representations. *Cognitive Science*, *35*, 1052–1075.

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment testing three principles of harm. *Psychological Science*, *17*, 1082–1089.

Damasio, A. R. (1994). Descartes' Error. Emotion, Reason and the Human Brain. New York (Grosset/Putnam).

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. *Handbook of Psychophysiology*, *2*, 200–223.

Delaney-Busch, N., & Kuperberg, G. (2013). Friendly drug-dealers and terrifying puppies: Affective primacy can attenuate the N400 effect in emotional discourse contexts. *Cognitive, Affective, & Behavioral Neuroscience*, *13*, 473–490.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–374.

Duska, R., & Whelan, M. (1975). Kohlberg's Theory of Moral Development. In *Moral Development: A Guide to Piaget and Kohlberg* (pp. 42–79). New York: Paulist Press.

Fields, E. C., & Kuperberg, G. R. (2012). It's all about you: An ERP study of emotion and self-relevance in discourse. *NeuroImage*, *62*, 562–574.

Fields, E. C., & Kuperberg, G. R. (2016). Dynamic effects of self-relevance and task on the neural processing of emotional words in context. *Frontiers in Psychology*, *6*, 2003.

Filik, R., Hunter, C. M., & Leuthold, H. (2015). When language gets emotional: Irony and the embodiment of affect in discourse. *Acta Psychologica*, *156*, 114–125.

Filik, R., & Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology*, *45*, 554–558.

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in Brain Research*, *156*, 185–203.

Forgas, J. P. (2012). *Affect in Social Thinking and Behavior*. Psychology Press.

Gable, P. A., Adams, D. L., & Proudfit, G. H. (2015). Transient tasks and enduring emotions: the impacts of affective content, task relevance, and picture duration on the sustained late positive potential. *Cognitive, Affective, & Behavioral Neuroscience*, *15*, 45–54.

Gernsbacher, M. A. (2013). *Language comprehension as structure building*. Psychology Press.

Gibbs, R. W. (2003). Embodied experience and linguistic meaning. *Brain and Language*, *84*, 1–15.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*, 1–26.

Glenberg, A. M., Webster, B. J., Mouilso, E., Havas, D., & Lindeman, L. M. (2009).

Gender, emotion, and the embodiment of language comprehension. *Emotion Review*, *1*, 151–161.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55–130.

Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to mcguire, langdon, coltheart, and mackenzie. *Journal of Experimental Social Psychology*, *45*, 581–584.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144–1154.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*, 389–400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*, 883–899.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In *The Neurocognition of Language* (pp. 273–317). Oxford University Press UK.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*, 438–441.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*, 814–834.

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*, 998–1002.

Haidt, J., & Hersh, M. A. (2001). Sexual morality: The cultures and emotions of conservatives and liberals 1. *Journal of Applied Social Psychology*, *31*, 191–221.

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*, 55–66.

Haidt, J., & Joseph, C. (2007). The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules. *The*

Innate Mind, *3*, 367–391.

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*, 613–628.

Haidt, J., et al. (2003). The moral emotions. *Handbook of Affective Sciences*, *11*, 852–870.

Hajcak, G., MacNamara, A., Foti, D., Ferri, J., & Keil, A. (2013). The dynamic allocation of attention to emotion: simultaneous and independent evidence from the late positive potential and steady state visual evoked potentials. *Biological Psychology*, *92*, 447–455.

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Developmental Neuropsychology*, *35*, 129–155.

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2012). ERPs and the study of emotion. *The Oxford Handbook of Event-related Potential Components*, *441*, 440–472.

Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J., & Davidson, R. J. (2010). Cosmetic use of botulinum toxin-a affects processing of emotional language. *Psychological Science*, *21*, 895–900.

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review*, *14*, 436–441.

Herring, D. R., Taylor, J. H., White, K. R., & Crites Jr, S. L. (2011). Electrophysiological responses to evaluative priming: the LPP is sensitive to incongruity. *Emotion*, *11*, 794–806.

Holt, D. J., Lynn, S. K., & Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *Journal of Cognitive Neuroscience*, *21*, 2245–2262.

Huebner, B., Dwyer, S., & Hauser, M. (2009). The role of emotion in moral psychology. *Trends in Cognitive Sciences*, *13*, 1–6.

Ito, T. A., & Cacioppo, J. T. (2000). Electrophysiological evidence of implicit and explicit categorization processes. *Journal of Experimental Social Psychology*, *36*, 660–676.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*, 887–900.

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PloS One*, *7*,

e42366.

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: two faces of moral regulation. *Journal of Personality and Social Psychology*, *96*, 521–537.

Kahane, G., Everett, J. A., Earp, B. D., Farias, M., & Savulescu, J. (2015). 'Utilitarian' judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193–209.

Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., & Tracey, I. (2012). The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience*, *7*, 393–402.

Kintsch, W. (1974). *The Representation of Meaning in Memory.* Lawrence Erlbaum Associates.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*, 163–182.

Kohlberg, L. (1981). The philosophy of moral development: Moral stages and the idea of justice (Essays on Moral Development, Volume 1). *Harper & Row*.

Krumhuber, E. G., Tsankova, E., & Kappas, A. (2018). Examining subjective and physiological responses to norm violation using text-based vignettes. *International Journal of Psychology*, *53*, 23–30.

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.

Lai, V. T., Hagoort, P., & Casasanto, D. (2012). Affective primacy vs. cognitive primacy: Dissolving the debate. *Frontiers in Psychology*, *3*, 243.

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*, 776–785.

León, I., Díaz, J. M., de Vega, M., & Hernández, J. A. (2010). Discourse-based emotional consistency modulates early and middle components of event-related potentials. *Emotion*, *10*, 863–873.

Leuthold, H., Filik, R., Murphy, K., & Mackenzie, I. G. (2012). The on-line processing of socio-emotional information in prototypical scenarios: Inferences from brain

potentials. *Social Cognitive and Affective Neuroscience*, *7*, 457–466.

Leuthold, H., Kunkel, A., Mackenzie, I. G., & Filik, R. (2015). Online processing of moral transgressions: ERP evidence for spontaneous evaluation. *Social Cognitive and Affective Neuroscience*, *10*, 1021–1029.

Lotto, L., Manfrinati, A., & Sarlo, M. (2014). A new set of moral dilemmas: Norms for moral acceptability, decision times, and emotional salience. *Journal of Behavioral Decision Making*, *27*, 57–65.

Luck, S. J. (2014). *An Introduction to the Event-related Potential Technique*. MIT Press.

McGuire, J., Langdon, R., Coltheart, M., & Mackenzie, C. (2009). A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology*, *45*, 577–580.

Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., & Pessoa, L. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *Journal of Neuroscience*, *22*, 2730–2736.

Moll, J., de Oliveira-Souza, R., Moll, F. T., Ignácio, F. A., Bramati, I. E., Caparelli-Dáquer, E. M., & Eslinger, P. J. (2005). The moral affiliations of disgust: A functional MRI study. *Cognitive and Behavioral Neurology*, *18*, 68–78.

Niedenthal, P. M. (2007). Embodying emotion. *Science*, *316*, 1002–1005.

Niedenthal, P. M., Winkielman, P., Mondillon, L., & Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology*, *96*, 1120–1136.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin*, *131*, 510–532.

Osterhout, L. (1997). On the brain response to syntactic anomalies: Manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*, 494–522.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Pastötter, B., Gleixner, S., Neuhauser, T., & Bäuml, K.-H. T. (2013). To push or not to push? Affective influences on moral judgment depend on decision frame. *Cognition*, *126*, 373–377.

Piaget, J. (1932).

The Moral Judgment of the Child.

Polman, E., & Ruttan, R. L. (2012). Effects of anger, guilt, and envy on moral hypocrisy.

Personality and Social Psychology Bulletin, *38*, 129–139.

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, *9*, 29–43.

Rozin, P. (2008). Hedonic" adaptation": Specific habituation to disgust/death elicitors as a result of dissecting a cadaver. *Judgment and Decision Making*, *3*, 191–194.

Sarlo, M., Lotto, L., Manfrinati, A., Rumiati, R., Gallicchio, G., & Palomba, D. (2012). Temporal dynamics of cognitive–emotional interplay in moral decision-making. *Journal of Cognitive Neuroscience*, *24*, 1018–1029.

Schaich Borg, J., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience*, *18*, 803–817.

Schaich Borg, J., Lieberman, D., & Kiehl, K. A. (2008). Infection, incest, and iniquity: Investigating the neural correlates of disgust and morality. *Journal of Cognitive Neuroscience*, *20*, 1529–1546.

Schmader, T., & Lickel, B. (2006). The approach and avoidance function of guilt and shame emotions: Comparing reactions to self-caused and other-caused wrongdoing. *Motivation and Emotion*, *30*, 42–55.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, *34*, 1096–1109.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Sevinc, G., & Spreng, R. N. (2014). Contextual and perceptual brain processes underlying moral cognition: a quantitative meta-analysis of moral reasoning and moral emotions. *PloS One*, *9*, e87427.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*, 387–401.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*, 768–777.

Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition*, *119*, 454–458.

Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., & Traue, H. C. (2012). Repeatability of facial electromyography (EMG) activity over corrugator

supercilii and zygomaticus major on differentiating various emotions. *Journal of Ambient Intelligence and Humanized Computing*, *3*, 3–10.

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2007). The skeletomotor system: Surface electromyography. In J. Cacioppo, L. Tassinary, & G. Berntson (Eds.), *Handbook of Psychophysiology* (p. 267-299). New York: Cambridge University Press.

't Hart, B. H., Struiksma, M. E., Van Boxtel, A. H., & Van Berkum, J. J. (2018). Emotion in Stories: Facial EMG Evidence for Both Mental Simulation and Moral Evaluation. *Frontiers in psychology*, *9*, 613.

Thompson, D., Mackenzie, I. G., Leuthold, H., & Filik, R. (2016). Emotional responses to irony and emoticons in written language: evidence from EDA and facial EMG. *Psychophysiology*, *53*, 1054–1062.

Ugazio, G., Lamm, C., & Singer, T. (2012). The role of emotions for moral judgments depends on the type of emotion and moral scenario. *Emotion*, *12*, 579–590.

Van Berkum, J. J. (in press). Language comprehension, emotion, and sociality: aren't we missing something? *To appear in Rueschemeyer & Gaskell (Ed.). Oxford Handbook of Psycholinguistics.*.

Van Berkum, J. J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, *11*, 657–671.

Van Berkum, J. J., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong? the brain's fast response to morally objectionable statements. *Psychological Science*, *20*, 1092–1099.

Van der Cruyssen, L., Van Duynslaeger, M., Cortoos, A., & Van Overwalle, F. (2009). Erp time course and brain areas of spontaneous and intentional goal inferences. *Social Neuroscience*, *4*, 165–184.

Van Dijk, T. A., Kintsch, W., & Van Dijk, T. A. (1983). *Strategies of discourse comprehension*. Academic Press.

Waldmann, M. R., Nagel, J., & Wiegmann, A. (2012). Moral Judgment. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (p. 274-299). Oxford University Press.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*, 780–784.

Xiang, M., & Kuperberg, G. (2015). Reversing expectations during discourse comprehension. *Language, Cognition and Neuroscience*, *30*, 648–672.

Yang, Q., Yan, L., Luo, J., Li, A., Zhang, Y., Tian, X., & Zhang, D. (2013). Temporal dynamics of disgust and morality: an event-related potential study. *PloS One*, *8*, e65094.

Zajonc, R. B. (1984). *On the primacy of affect.* American Psychological Association.

Zwaan, R. A. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. *Psychology of Learning and Motivation*, *44*, 35–62.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.