

Conflict Detection In Affective Language Content: Affective Incongruency In Semantically Correct Sentences Describing Social Interactions Activates The Anterior Cingulate Cortex

Gesche Schauenburg (✉ gesche.schauenburg@gmail.com)

Freie Universität Berlin

Arash Aryani

Freie Universität Berlin

Chun-Ting Hsu

Freie Universität Berlin

Tobias Schröder

University of Applied Science

Markus Conrad

Universidad de La Laguna

Christian von Scheve

Freie Universität Berlin

Arthur M. Jacobs

Freie Universität Berlin

Research Article

Keywords: neuroscientific, anterior cingulate cortex, replication, Affect Control Theory

Posted Date: June 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-549518/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Neurocognitive studies on the emotion-language relation report a significant influence of affective content on the level of single words. However, it is rather difficult to investigate such influence on the sentence level – partly due to a missing theoretical approach to integrate multiple affective meanings. In a previous EEG study, we used impression formation equations based on *Affect Control Theory* to calculate affective congruency of sentences describing social interactions to successfully predict ERP effects. The current study was crucially motivated by the assumption that the reported early effect of affective incongruency was related to automatic conflict detection associated with activation in anterior cingulate cortex (ACC). To test this hypothesis, we replicated the study as an event-related fMRI design: we visually presented the same sentences to 23 participants (12 f, 11 m) in a silent reading task while measuring differences in the hemodynamic response in two conditions of affective congruency. The ROI analysis results showed expected enhancement of neural activity for affectively incongruent sentences in left ACC, supporting the assumption that affective language content influences meaning making already at basic semantic processing stages. The reported results represent a replication of neuroscientific evidence for ACT's mathematical model of impression formation.

Introduction

Research on the emotion-language relation indicates that affective language content influences meaning processing, perception, and behavior. Concerning words in isolation in visual language processing, studies report such influences of affective information in a variety of tasks and on a range of behavioral and neuropsychological measurements. These results support the assumption of an rapid-automatic influence of emotion word content [1, 2]. However, it seems to be more challenging to tackle such effects at the level of sentences and phrases - not least due to a missing theoretical approach concerning the interplay between affective meanings of multiple words in word context. In a recent event-related potential (ERP) study [3], we approached this issue by employing a mathematical model of impression formation based on *Affect Control Theory* (ACT; [4, 5] to determine affective congruency of complete sentences describing social interactions.

In a design highly controlled for confounding variables for semantic processing, we were able to report ERP effects related to implicit processing of affective congruency. Crucially, there was an early negative waveform for affectively incongruent compared to congruent sentence-final words over anterior electrode sites in the N2/P2 component (130–270 ms after target word onset). That is, incongruency of affective meanings seems to be part of early semantic processing. We considered this finding a correlate of early *affective conflict* detection. Neuroimaging studies ranging from investigating basic perceptual processing to the rather complex processing of moral judgment report that the ACC is related to automatic conflict processing [6–9]. The ACC is well-connected to limbic structures and prefrontal regions [10]. The former is predominantly associated with emotional and automatic processing. The latter is mainly associated with executive control such as decision making, attention allocation, information integration and updating. Therefore, the ACC is very well suited to monitor cognitive and affective information inflow

from limbic structures and to recruit prefrontal areas for more deliberate processing (e.g., dlPFC; 9) such as emotion regulation, social cognition, and decision making when conflict is detected. Due to these peculiar structural and functional features, ACC seems to be potentially relevant for predictions derived from ACT.

Similar to other consistency theories (e.g., [11]), ACT postulates that when incoming representations of social events are affectively congruent to each other they can be easily integrated into our stream of perception, but when they are affectively incongruent, we "mentally stumble" to integrate representations with clashing affective meanings or concepts. In such situations, individuals try to re-interpret respective social events to (re-)establish affective congruency. Hence, the ACC could be vital for a neuropsychological foundation of ACT's theoretical predictions concerning the integration of information flow, the "mental stumbling", and the recruitment of prefrontal areas for re-appraisal and re-interpretation.

In the present study, we aim to complement our findings from the ERP study by Schauenburg et al. (2019) by testing the hypothesis of whether affective incongruency drives ACC activation. Therefore, we converted the previous ERP study's design to an event-related fMRI study design as closely as possible. The same sentence stimuli were visually presented in a word-by-word manner. While we included three levels of affective incongruency (low, medium, high) in the previous ERP study, we only included extreme conditions (low congruent vs. high incongruent) in the present fMRI study limit scanning time. We included the same attention questions to keep participants attentively reading the stimuli (i.e., semantically processing). For further analyses and as a manipulation check, we gathered likelihood ratings on described interactions since it is well established that an event's perceived likelihood correlates with its affective congruency [12]. To verify our results, we contrasted the experimental sentence reading condition with a control condition showing hash signs (#) matched in single sign and thread length with word and sentence length.

Results

2.1. Behavioral Performance

All 23 participants included in the final data set correctly responded to attention questions above chance ($\geq 66\%$) with a mean accuracy of 86.77 %.

2.2. fMRI Results

Table 1. MNI-Coordinates and statistics for fMRI analyses.						
<i>Anatomical Definition</i>		<i>MNI</i>			<i>CL_size</i>	<i>Peak T</i>
		<i>x</i>	<i>y</i>	<i>z</i>		
<i>AllSentences > AllThreads (whole-brain corrected)</i>						
R	Limbic lobe (hippocampus)	26	-44	6	521	7.22
L	Basal ganglia (caudate tail	-15	-29	20	687	7.03
LR	Dorsolateral and medial prefrontal cortex (BA9)	-5	56	30	1554	7.00
L	Lateral temporal lobe (BA38, 22, 21)	-57	6	-15	728	6.80
R	Sensory association cortex (BA5)	21	-41	65	534	6.71
R	Lateral temporal lobe (BA38, 21)	51	8	-27	447	6.17
L	Angular area (BA 39,19)	-47	-60	23	840	6.02
L	Basal ganglia (caudate body)	-9	18	9	491	5.97
L	Inferior frontal gyrus (BA45, 44)	-56	29	15	335	5.86
L	Ventral posterior cingulate area (BA23)	-3	-57	11	2362	5.75
R	Angular area (BA39)	56	-60	21	152	5.17
L	Basal ganglia (globus pallidus)	-27	-11	-5	145	5.08
<i>AffectivelyIncongruent > AffectivelyCongruent (small volume corrected)</i>						
L	Anterior cingulate cortex (BA32)	-14	-44	8		
					14	4.72
Significant peak voxel for all clusters at $p < .05$ FWE-corrected; name of anatomical definition and x,y,z-coordinates are given for first brain region of relevant cluster, in parentheses involved brain regions are listed corresponding to Brodmann areas : CL_size = Cluster size, L/R = Left/right, BA = Brodmann's area, MNI = Montreal Neurological Institute						

The comparison between *AllSentences > AllThreads* revealed activations in brain areas canonically associated with visual language and semantic processing. Small volume correction with the bilateral ACC mask showed increased neural activity in left ACC (BA 32) for affectively incongruent compared to affective congruent sentence processing. Please see Table 1 for an overview of the results. Figure 1 shows brain activation for *AllSentences > AllThreads*. Figure 2 shows brain activation in ACC for *AffectIncongruent > AffectCongruent* sentences.

2.3. Likelihood Ratings

Participants rated social interactions described in affectively incongruent sentences ($AM = 3.88$, $SD = .91$; $n = 106$) less likely than those described in affectively congruent sentences ($AM = 4.53$, $SD = .91$; $n = 106$; $T(210) = 5.163$; $p < .001$).

Discussion

This study investigates whether ACT-based impression formation equations can predict ACC activation in affectively incongruent yet semantically correct sentences describing social interactions. This investigation was motivated by a previous EEG study where we found an early effect of affective incongruency processing in the P2/N2 latency range [3]. We suggested this early effect to be a neural correlate of conflict detection which is related to ACC activation. This study's main goal was two-folded: a) to replicate previous findings from our EEG study and b) to test the hypothesis that the reported early effect of affective congruency violation is related to conflict processing as reflected by increased ACC activation. To ensure that expected findings tap into meaning making processes, we included attention questions and a control condition of hash signs (non-reading) to compare with brain activation in the experimental condition. As a manipulation check of affective congruency, we gathered likelihood ratings on the interactions illustrated in sentence stimuli. T-Test comparison confirmed well-known deflection-likelihood relations, i.e., events of greater affective deflection are perceived as less likely as compared to those of low affective deflection (12). Due to our highly controlled stimuli in the current design, this relation cannot be explained by other lexico-semantic variables known to influence expectancy related processes (cf. Schauenburg et al., 2019).

Taken together, behavioral and fMRI data strongly support the assumption that participants were indeed semantically engaged during sentence processing: answer accuracy to attention questions was above chance suggesting participants were attentively reading presented sentences and the results of whole level brain contrasts are in line with prevailing literature. Language comprehension and semantic processing are commonly associated with mainly left-lateralized activity in the lateral temporal lobe, including superior temporal gyrus (BA22), medial temporal gyrus (BA21), and temporal pole (BA38), the inferior frontal gyrus (BA44, 45), parietal lobe (BA39), posterior cingulate cortex (BA23) and prefrontal cortex (BA9), such as weaker activity in homologous areas of the right hemisphere [13–16]. We suggest that activity in caudate is linked to task demands depending on intention and attention-related processes necessary to keep presented linguistic representations active and suppress competing cognition [17]. Hippocampal activation has been shown to support flexible cognition related to online and incremental social and language processing crucial to successful behavior and creation of meaning of our world [18–20]. Therefore, we relate reported hippocampal activation to processes implicated in forming mental representations of presented social interactions, i.e., to bind and update prior knowledge of relevant concepts within actual, contextual and relational information.

Concerning effects of the relevant experimental manipulation on the neural level, we expected reading sentences with incongruent versus congruent affective meanings to correlate more strongly with hemodynamic responses in the ACC, a peculiar brain region monitoring, mediating and integrating

cognitive and emotion-related processes facilitated by its specific connectivity to both limbic brain structures and the prefrontal cortex area [10]. Neural activity in ACC is correlated with error detection, conflict monitoring, expectancy-related processes, and affective experiences [6, 7, 21]. As hypothesized, ACC showed increased neural activity for affectively incongruent sentences. Thus, this study replicates evidence from our previous ERP results and provides a) neuroscientific evidence for the ACT-based mathematical model of impression formation of linguistic representations of social interactions, and b) suggests that violations of affective congruency are reflected by ACC activation even in the absence of explicit task demands, i.e., supporting the assumption of implicit, automatic processing of affective congruency [3].

In the current study, we used ACC activation especially as a dependent variable to assess the influence of affective incongruency on cognitive processing rather than investigating and adding new perspectives to the broad corpus of studies exploring basic mechanisms of ACC activation. However, here we want to give some considerations about how reported ACC activation to affective *Deflection* might be explained in terms of *error signaling to prediction error*. Recently, 22 [22] presented a computational model for mPFC and dlPFC. The Hierarchical Error Representation (HER) model is rooted in the tradition of Reinforcement Learning models which value in surprise as (prediction) error to be a key component in the formalization of association learning as the difference between maximum possible strength and current total strength [23]. By proposing that ACC is not sensible to the affective component of any stimulus, but to surprise (i.e., the discrepancy of the occurrence or non-occurrence of an expected or non-expected event), the HER model provides a considerable and unifying approach to explain ACC activation in various studies. Furthermore, the model assumes a hierarchical organization of frontal lobe processes where error signaling is driven by "bottom-up" and prediction error modeling by "top-down" processes. Within this approach, the reported finding of ACC activation can also be explained as an error signal of predicted affective congruency induced by sentence-final words. Crucially, the ACT-based concept of "affective *Deflection*" is somewhat similar to the concept of "error prediction" because it measures the degree to which an expected affective meaning is violated by a currently perceived affective meaning (which is further mirrored by the inverse correlation of affective *Deflection* and perceived likelihood of an event). Similar to the concept of prediction error, the degree of affective *Deflection* is an experienced-based correlate of the discrepancy between predicted probability and the actually observed social event.

Furthermore, ACT postulates that impressions of affectively incongruent events would initiate processes aiming to regain affective congruency [24]. Within the HER model, corresponding neural processes should be on "higher levels", "top-down" modulating areas in the prefrontal cortex like dlPFC. A more detailed comparison of the HER model and ACT is beyond the scope of the current study. However, we assume that both models could complement each other concerning the computational modeling of precise and testable brain activation and connectivity to corresponding bottom-up and top-down processes throughout the frontal lobe.

Concerning the present results, we found increased activity only in the left ACC. Previous studies on conflict detection provide inconsistent findings concerning each hemisphere's differential contributions

[25]. Crucially, 26 [26] found cognitive control to be localized in the same hemisphere as task execution: the left hemisphere was more involved in a letter-decision, whereas the right hemisphere was more involved in a visuospatial-decision task. Transferred to our study, this finding would explain activation of the left ACC due to the reading task associated with the left hemisphere. Nonetheless, 27 [27] found that differential hemispheric involvement was related to differences between cognitive processes of conflict detection versus error processing rather than to task-induced lateralized processing. 28 [28] reported increased left ACC activation due to conflicting stimuli in a cognitive flanker task, while the right ACC was more involved in an affective flanker task. However, going into more detail concerning the applied tasks and the salience of emotionality (implicit processing of affective congruency in a reading task vs. explicit categorization task of affective words), the comparison of reported findings of lateralized ACC involvement seems to be easier said than done. Future studies should further investigate differential lateralization of ACC activity in different task and stimulus settings to overcome these inconsistencies.

We consider the current study to be a fruitful and integrating approach of employing a social psychological model for an affective neuroscientific research question concerning the influence of affective language content on sentence processing. The applied concept of affective congruency as provided by ACT captures features of social information influencing linguistic representation of social events, which go beyond common variables known to influence semantic processing. However, some limitations need to be mentioned. First, we see that hash signs were probably not an ideal cognitive baseline in this design. They served as visually comparative stimuli, but what cognitive processes participants might be engaged in when perceiving hash signs is not strictly controlled. This also questions the comparability of processing demands in terms of task difficulty, see 14. Second, we presume that a combined EEG-fMRI study with the same group of participants would probably provide better-coupled information about the temporal evolution of cognitive processes and their neural correlates. Third, individual differences in social cognition (e.g., empathy trait, theory of mind or mentalizing abilities), personality traits, or language processing competency could be investigated in future studies.

Our study is the first attempt to explore differential brain activity to manipulations of affective congruency as operationalized by ACT-based mathematical model of impression formation. Consistent with our hypothesis rooted in the reported findings of a previous EEG study, affective incongruency was related to increased ACC activation – even in the absence of an explicit processing task in semantically correct sentences. This study's observations emphasize the basic influence of affective language content during meaning making and are in line with the previous EEG study results.

Methods

4.1. Participants

Twenty-seven German native speakers gave written consent and participated in this study, which was conducted in accordance with the Declaration of Helsinki and had been approved by the Ethics

Committee of Freie Universität Berlin, Germany. All participants were right-handed (German adaption of [29]) with normal or corrected-to-normal vision and no history of psychological or neurological diseases. Participation was monetarily rewarded (8 €/h). The data of four participants were excluded due to too much movement during data acquisition (more than 5 mm in any orientation) or due to too many wrong answers to attention questions (more than 16, i.e., one-third wrong answers). The final set included data from 23 subjects (mean age 24.70 years, range 19–34; 12 women, 11 men).

4.2. Stimuli

In our ACT-based approach, affective congruency describes affective meaning maintenance. This defines the degree to which a concept's affective meaning is maintained during the (linguistic) representation of a current event. ACT operationalizes affective congruency reversely by the *Deflection* between a concept's fundamental (without context) and its transient (in-context) affective meaning. ACT-based impression formation regression equations provide such *Deflection* measurements for linguistically represented social situations based on substantial rating data of affective meanings (for more detail, see [3, 30]). We built German sentence stimuli with words for which affective meaning information was available [30–34] and applied impression formation equations to determine affective *Deflection* for each sentence. A total of 212 sentences, i.e., 106 sentences for each condition of low and high affective *Deflection*, were selected.

Figure 3 displays some examples of sentence stimuli and illustrates allocation to both experimental conditions. All sentences describe social interactions in a basic subject-verb-object structure. Both experimental conditions included same sentence contexts (= subject + verb) and final words (= objects), they only differed in their combination resulting in different *Deflection* value which was calculated according to 30 [30]. Sentence stimuli were balanced across conditions concerning variables known to influence language processing, such as cloze probability, contextual constraints, and co-occurrences. See table 2 for an overview of study design and stimuli characteristics. For an elaborated description of sentence stimuli generation, please see 3 [3].

Table 2. Overview of study design and stimuli example.				
	<u>Experimental sentences</u>		<u>Control threads</u>	
	Affectively Congruent	Affectively Incongruent	Without @	With @
Stimulus example	The hussy chokes the beast	The hussy chokes the fairy	### ##### ### #####	### ##### #@### ### #####
N per run	26 (27)	27 (26)	10 (11)	11 (10)
<i>N total</i>	<i>106</i>	<i>106</i>	<i>42</i>	<i>42</i>
Question example	Did the hussy strangle the beast? (yes)		Was there an @? (no)	
N yes, per run	6	6	6	6
N no, per run	6	6	6	6
N total, per run	12	12	12	12

To allow for experimental manipulation verification by contrasting reading versus non-reading, we included a control condition presenting non-verbal signs (hash signs): a total of 84 hash-sign threads were matched with the experimental condition, i.e., the length of single hash-sign strings was matched to word length, and hash-sign threads were matched to sentence length, i.e., 5–6 words. There were two control conditions: hash-sign threads with and without @signs.

4.3. Procedure

Sentences and hash-sign threads were presented visually in the same manner: A trial began with the presentation of a 250 ms fixation cross followed by a 100 ms blank screen. Sentences and hash-sign threads were then presented in a word-by-word fashion: each word or hash-sign string with less than nine letters (hash signs) was presented for 250 ms, words with more than nine letters (hash signs) were presented for 300 ms respectively to ensure complete perception (controlled across runs). Since both conditions included the same word stimuli, presentation time of single words was perfectly balanced. There was a 100 ms blank between words (hash signs), which improved smooth reading perception as tested before the final experiment started. Each stimulus was presented in the center of the screen. The Inter-Stimulus-Interval (ISI) between sentences and hash-sign threads, respectively, was optimized with optseq2 (<http://surfer.nmr.mgh.harvard.edu/optseq/optseq2.help.txt>) jittering a mean 2000 ms interval. To keep participants attentive following the trials, yes-no-questions (balanced for type of correct answer)

were randomly assigned and had to be answered by a button press (index finger = yes, middle finger = no). There were twelve attention questions for each condition of affective deflection, which focused on the preceding sentence using synonyms to make answering sufficiently challenging. In control trials, twelve attention questions were randomly presented, asking if an @-sign was present in the preceding hashmark thread. Jittering order and stimulus presentation order over different conditions (low affective deflection, high affective deflection, control condition, attention question) were optimized to ensure a maximal signal-to-noise ratio.

To ensure that participants were familiar with the experimental setting and task, they began with a practice run. The practice run consisted of eight trials, five sentence trials, and three hash-sign trials, each followed by an attention question. Emoticons gave feedback to participants' responses: frowning emoticons indicated a wrong answer, smiling ones indicated a right answer. Feedback was given only during practice trials. We used Presentation® software (Neurobehavioral Systems) to present stimuli and record participants' answers to attention questions. Practice runs were followed by four experimental blocks in between which participants could take a break.

4.4. Design

The experiment consisted of four runs, each containing 74 trials (26 resp. 27 sentences of low affective deflection; 26 resp. 27 sentences of high affective deflection, 21 control trials) and nine attention questions; assigned trials were randomized within each run, and the running order was randomized across subjects. For an overview, please see Table 1.

4.5. fMRI Data Acquisition

Functional data were acquired on a Siemens Tim Trio 3T MR scanner. Four runs of 282 volumes were measured using a T_2^* -weighted echo-planar sequence [slice thickness: 3 mm, no gap, 37 slices, repetition time (TR): 2 s, echo time (TE): 30ms, flip angle: 70°, matrix: 64 × 64, field of view (FOV): 192 mm, voxel size: 3.0 mm × 3.0 mm × 3.0 mm] and individual high-resolution T1- weighted anatomical data (MPRAGE sequence) were acquired (TR: 1.9, TE: 2.52, FOV: 256, matrix: 256 × 256, sagittal plane, slice thickness: 1 mm, 176 slices, resolution: 1.0 mm × 1.0 mm × 1.0 mm).

4.6. fMRI Preprocessing

The fMRI data were preprocessed and analyzed with software package SPM12 (www.fil.ion.ucl.ac.uk/spm). Preprocessing covered slice-timing correction, realignment for motion correction, and sequential co-registration. Structural images were segmented into grey matter, white matter, cerebrospinal fluid, bone, soft tissue, and air/background with the 'New Segment' module [35]. A group anatomical template was created with the DARTEL (Diffeomorphic Anatomical Registration using Exponentiated Lie algebra; 36) toolbox from the segmented grey and white matter images. Transformation parameters for structural images were then applied to functional images to normalize them to the brain template of the Montreal Neurological Institute (MNI) supplied with SPM. Functional

images were resampled to a resolution of $1.5 \times 1.5 \times 1.5$ mm and spatially smoothed with a kernel of 6 mm full-width-at-half-maximum during normalization.

4.7. fMRI GLM Analysis

We realized statistical parametric maps by multiple regressions of the data onto a model of the hemodynamic response. In the subject-level analysis, this model contained regressors for both experimental and control conditions. Regressors were convolved with canonical hemodynamic response function in SPM12. Beta-images of each of the four conditions in the factorial design for each participant were used at the group level to a random-effect paired t-test analysis (Holmes & Friston, 1998) to explore potential differential correlations between the two experimental conditions (*AffectIncongruent* > *AffectCongruent*) and between reading and non-reading conditions (*AllSentences* > *AllThreads*). Thus, we contrasted experimental conditions vs. control conditions ($[AffectIncongruent + AffectCongruent] > [HashSignWith@ + HashSignWithout]$, in short: *AllSentences* > *AllThreads*) at the whole-brain level to test if the experiment successfully tapped into the language processing network. Whole-brain fMRI analyses included a cluster forming threshold for uncorrected $p < .001$, then a cluster-level threshold of family-wise error (FWE) corrected $p < .05$ for the entire image volume. For our experimental manipulation (*AffectIncongruent* > *AffectCongruent*), we also performed a small volume correction (SVC) with a bilateral ACC mask (BA32 mask defined by WFU Pickatlas Tool; 37). In the ROI analyses, we applied an initial voxel-level threshold of uncorrected $p < .001$ for the entire image, followed by the threshold of voxel-level family-wise error (FWE) corrected $p < .05$ after applying SVC with the bilateral ACC mask.

4.8. Post-Scan Sentence Ratings

For further analyses and as a manipulation check of affective deflection, we gathered likelihood ratings on the described interactions of presented sentences after scanning had been completed. Participants were asked to rate each sentence concerning the question "How likely do you rate the described interaction" from "not at all likely" to "extremely likely" on a 7-point-Likert scale. Mean likelihoods were computed across all 106 sentences separated for each condition of affectively congruent vs. incongruent interactions. Obtained likelihood means were compared by t-test.

Declarations

6. Acknowledgements

This research was realized by a scholarship to G.S. by German Research Foundation (DFG) within Cluster of Excellence "Languages of Emotion" (Freie Universität, Berlin, Germany). Partial funding for this study was provided by A. J. and by a grant to M.C. and C.v.S. through the Cluster of Excellence "Languages of Emotion" of the DFG.

7. Author contributions

G.S. conceived, planned, carried out the experiment, processed and analysed the data, and wrote the manuscript. C.-T.H. assisted developing the event-related fMRI design and wrote the code for data acquisition. A.A. assisted with data processing and analyses. T.S., A.J., C.v.S., and M.C. were involved in planning and supervised the work. All authors reviewed the manuscript.

8. Additional Information

Conflict of interest statement: None.

References

1. Bernat, E., Bunce, S. & Shevrin, H. Event-related brain potentials differentiate positive and negative mood adjectives during both supraliminal and subliminal visual processing. *International Journal of Psychophysiology* **42**, 11–34 (2001).
2. Citron, F. M. M. Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and language* **122**, 211–226 (2012).
3. Schauenburg, G. *et al.* Making sense of social interaction: Emotional coherence drives semantic integration as assessed by event-related potentials. *Neuropsychologia* **125**, 1–13 (2019).
4. Heise, D. R. *Expressive Order. Confirming Sentiments in Social Actions* (Springer Science + Business Media LLC, 2007).
5. Heise, D. R. *Understanding events. Affect and the construction of social action*. 1st ed. (1979).
6. Botvinick, M. M., Cohen, J. D. & Carter, C. S. Conflict monitoring and anterior cingulate cortex: an update. *Trends in cognitive sciences* **8**, 539–546 (2004).
7. Braem, S. *et al.* The Role of Anterior Cingulate Cortex in the Affective Evaluation of Conflict. *Journal of cognitive neuroscience* **29**, 137–149 (2017).
8. Carter, C. S. *et al.* Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science (New York, N.Y.)* **280**, 747–749 (1998).
9. Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. The neural bases of cognitive conflict and control in moral judgment. *Neuron* **44**, 389–400 (2004).
10. Stevens, F. L., Hurley, R. A. & Taber, K. H. Anterior cingulate cortex: unique role in cognition and emotion. *The Journal of neuropsychiatry and clinical neurosciences* **23**, 121–125 (2011).
11. Heider, F. Attitudes and cognitive organization. *The Journal of psychology* **21**, 107–112 (1946).
12. Heise, D. R. *Surveying cultures. Discovering shared conceptions and sentiments* (Wiley, 2010).
13. Binder, J. R. in *fMRI Techniques and Protocols*, edited by M. Filippi (Springer New York, 2016), pp. 355–385.
14. Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral cortex (New York, N.Y.: 1991)* **19**, 2767–2796 (2009).

15. Federmeier, K. D., Wlotko, E. W. & Meyer, A. M. What's "right" in language comprehension: ERPs reveal right hemisphere language capabilities. *Language and linguistics compass* **2**, 1–17 (2008).
16. Ferstl, E. C., Neumann, J., Bogler, C. & Cramon, D. Y. von. The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Hum. Brain Mapp.* **29**, 581–593 (2008).
17. Crosson, B., Benjamin, M. & Levy, I. in *Neural basis of semantic memory*, edited by M. A. Kraut & J. Hart (Cambridge University Press, 2007), pp. 219–244.
18. Rubin, R. D., Watson, P. D., Duff, M. C. & Cohen, N. J. The role of the hippocampus in flexible cognition and social behavior. *Frontiers in human neuroscience* **8**, 742 (2014).
19. Montagrin, A., Saiote, C. & Schiller, D. The social hippocampus. *Hippocampus* **28**, 672–679 (2018).
20. Duff, M. C. & Brown-Schmidt, S. The hippocampus and the flexible use and processing of language. *Frontiers in human neuroscience* **6**, 69 (2012).
21. Kanske, P. & Kotz, S. A. Emotion triggers executive attention: anterior cingulate cortex and amygdala responses to emotional words in a conflict task. *Human brain mapping* **32**, 198–208 (2011).
22. Alexander, W. H. & Brown, J. W. The Role of the Anterior Cingulate Cortex in Prediction Error and Signaling Surprise. *Topics in cognitive science* **11**, 119–135 (2019).
23. Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current research and theory*, edited by A. H. Black & W. F. Prokasy (Appleton Century Crofts, 1972), pp. 64–99.
24. Robinson, D. T., Smith-Lovin, L. & Wisecup, A. K. in *Handbook of the Sociology of Emotions*, edited by J. E. Stets & J. H. Turner (Springer US, 2006), pp. 179–202.
25. Lütcke, H. & Frahm, J. Lateralized anterior cingulate function during error processing and conflict monitoring as revealed by high-resolution fMRI. *Cerebral cortex (New York, N.Y.: 1991)* **18**, 508–515 (2008).
26. Stephan, K. E. *et al.* Lateralized cognitive processes and lateralized task control in the human brain. *Science (New York, N.Y.)* **301**, 384–386 (2003).
27. Lütcke, H. & Frahm, J. Lateralized anterior cingulate function during error processing and conflict monitoring as revealed by high-resolution fMRI. *Cerebral cortex (New York, N.Y.: 1991)* **18**, 508–515 (2008).
28. Ochsner, K. N., Hughes, B., Robertson, E. R., Cooper, J. C. & Gabrieli, J. D. E. Neural systems supporting the control of affective and cognitive conflicts. *Journal of cognitive neuroscience* **21**, 1842–1855 (2009).
29. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
30. Schröder, T. A Model of Language-Based Impression Formation and Attribution Among Germans. *Journal of Language and Social Psychology* **30**, 82–102 (2011).
31. Ambrasat, J., Scheve, C. von, Conrad, M., Schauenburg, G. & Schröder, T. Consensus and stratification in the affective meaning of human sociality. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 8001–8006 (2014).

32. Schmidtke, D. S., Schröder, T., Jacobs, A. M. & Conrad, M. ANGST: affective norms for German sentiment terms, derived from the affective norms for English words. *Behavior research methods* **46**, 1108–1118 (2014).
33. Schauenburg, G., Ambrasat, J., Schröder, T., Scheve, C. von & Conrad, M. Emotional connotations of words related to authority and community. *Behavior research methods* **47**, 720–735 (2015).
34. Võ, M. L. H. *et al.* The Berlin Affective Word List Reloaded (BAWL-R). *Behavior research methods* **41**, 534–538 (2009).
35. Ashburner, J. & Friston, K. J. Unified segmentation. *NeuroImage* **26**, 839–851 (2005).
36. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113 (2007).
37. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* **19**, 1233–1239 (2003).

Figures

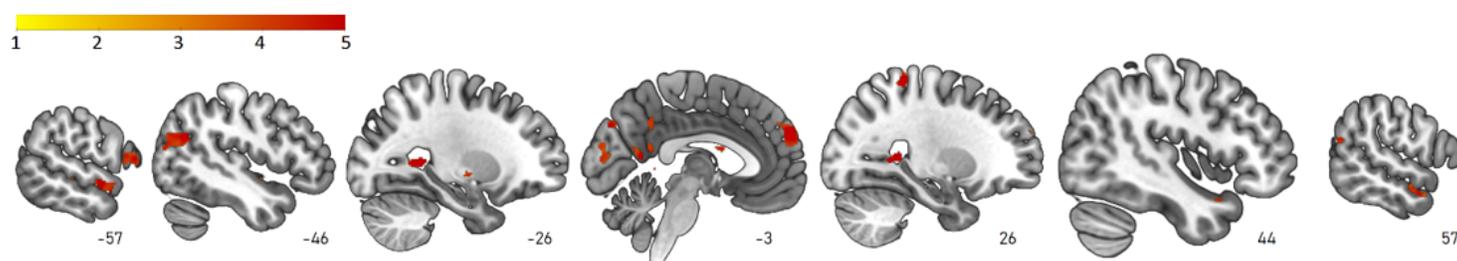


Figure 1

Brain activation for contrast of experimental vs. control condition (AllSentences > AllThreads).

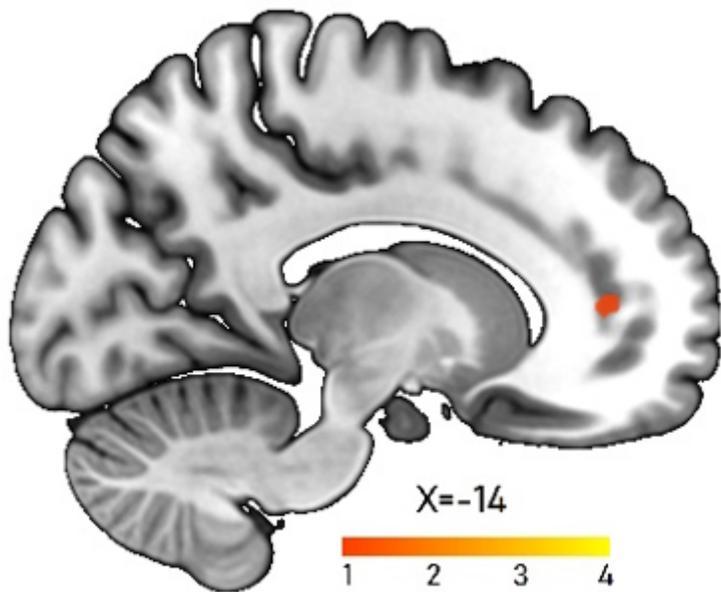


Figure 2

Brain activation for ROI analysis in ACC for contrast of affectively incongruent vs. congruent sentences (AffectivelyIncongruent > AffectivelyCongruent)

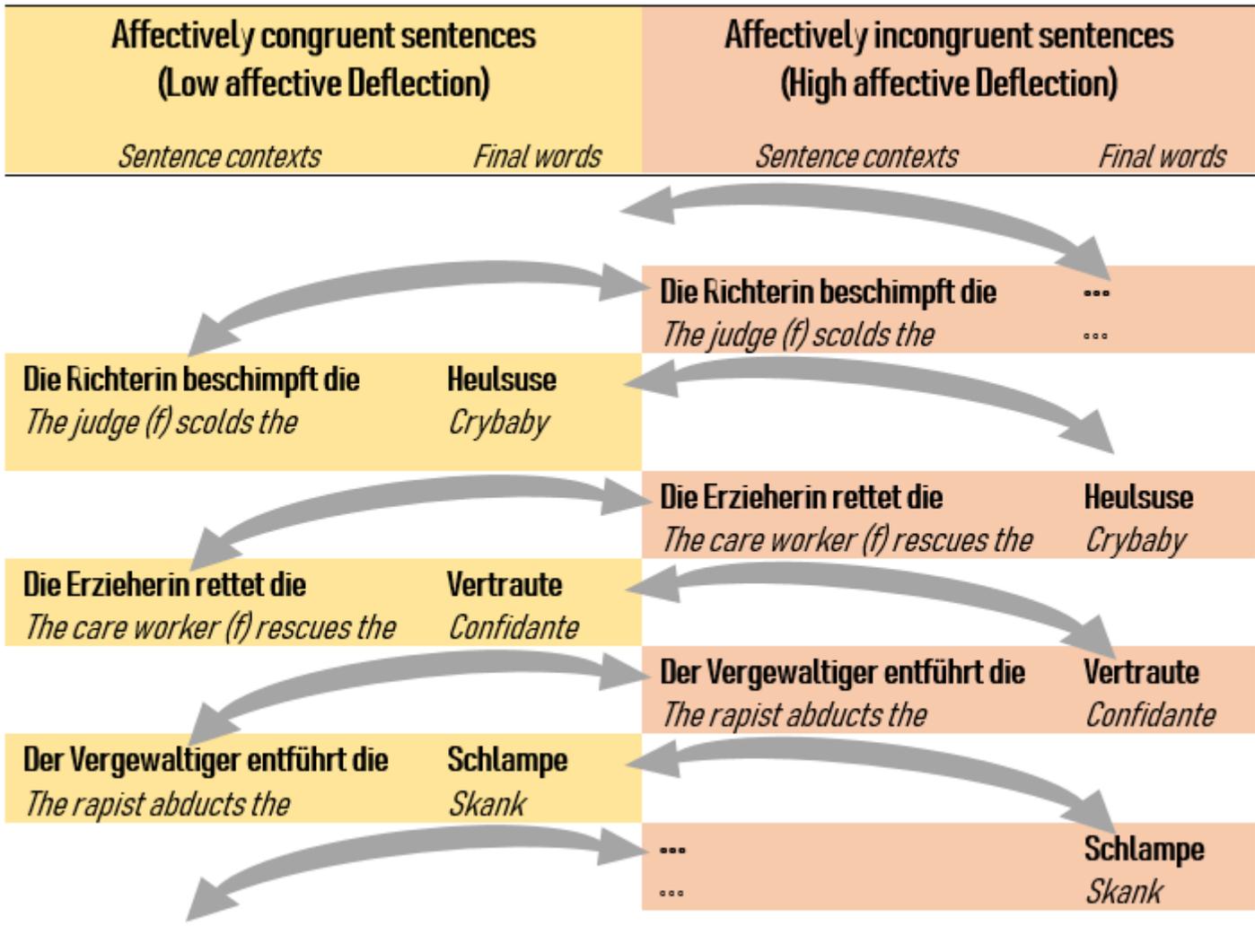


Figure 3

Schematic illustration of stimuli design with a selection of experimental sentences and proxies. The arrows illustrate how every sentence context (= actor + behavior) and every sentence-final word (= object) enters all three conditions of affective deflection. Original sentence stimuli are presented in bold print, translations in cursive below.