Learning and Instruction 47 (2017) 80–90

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

The effect of anticipated achievement feedback on students' semantic processing as indicated by the N400 cloze effect

Eva M. Janssen*, Sanne H.G. van der Ven, Anne H. van Hoogmoed, Paul P.M. Leseman

Department of Child, Family and Education Studies, Utrecht University, P.O. Box 80140, 3508 TC Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 24 December 2015
Received in revised form
29 September 2016
Accepted 7 October 2016
Available online 2 November 2016

Keywords:
Anticipated achievement feedback
Emotions
Semantic processing
N400
Event-related potentials

ABSTRACT

The present study examined the effects of anticipated achievement feedback on students' semantic processing on the neural level, using event-related brain potentials (ERPs). Participants ($N = 79$) anticipated either self-referential or normative achievement feedback regarding an announced upcoming test. Additionally, their performance expectations (low vs. no expectations) were orthogonally manipulated. Subsequently, students' on-line semantic processing was assessed by measuring the N400 cloze effect, a component in the EEG signal of which the amplitude is associated with semantic processing. Within the low performance-expectation condition, no effect of anticipated feedback on semantic processing was found. Within the no-performance-expectation condition, participants anticipating self-referential feedback showed a more widely distributed N400 cloze effect than participants anticipating normative feedback. The results confirmed the hypothesis that the mere expectation of a particular type of feedback can affect students' semantic processes.

1. Introduction

In the present study we analyse neural correlates of anticipated feedback, by comparing the effects of two types of anticipated feedback on semantic processing, using event-related brain potentials (ERPs) as measured with an electroencephalogram (EEG). In education, norm-referenced tests are widely used for several high-stakes purposes, such as students' access to higher education and future positions in society. However, there are indications that these tests have a negative impact on students' performance. One study showed that the emotional response to achievement feedback after test performance can hamper students' future learning efforts (Mangels, Good, Whiteman, Maniscalco, & Dweck, 2012). Another study demonstrated that the mere expectation of later achievement feedback influenced students' emotions, before any feedback was actually received (Pekrun, Cusack, Murayama, Elliot, & Thomas, 2014): notably, anticipating normative feedback evoked negative emotions in students. It seems clear that feedback and its related emotions affect students and their academic performance, but the neurocognitive mechanisms underlying the effects of feedback are still unknown.

1.1. Achievement goals, emotions, and feedback

According to achievement goal theorists, people form impressions of the reasons why they are performing a particular task (Ames, 1992). These impressions form their achievement goals. Two contrasting goal types that have received most attention are mastery goals and performance goals. A mastery goal emphasises the intrinsic value of learning and is characterised by the belief that ability can be developed (Butler, 1987). The focus is on developing one's competence or "mastery" of a certain skill (Ames, 1992). A performance goal, on the other hand, emphasises the extrinsic value of learning and is characterised by the belief that one's ability reflects a fixed capacity (Dweck, 1986). The focus is on one's performance compared to others (Ames, 1992). Recent achievement-goal research applies a trichotomous framework, in which performance goals are further divided into performance-avoidance and performance-approach goals (Elliot & Harackiewicz, 1996; Pekrun et al., 2014; Pekrun, Elliot, & Maier, 2006; 2009). Students with performance-avoidance goals focus on avoiding the negative effect of failure, whereas students with performance-approach goals focus on reaching the positive value of success.

Both mastery and performance achievement goals evoke particular achievement emotions in students (Dweck, 1986; Pekrun

* Corresponding author.
E-mail addresses: e.m.janssen@uu.nl (E.M. Janssen), s.vanderven@uu.nl (S.H.G. van der Ven), a.h.van.hoogmoed@rug.nl (A.H. van Hoogmoed), p.p.m.leseman@uu.nl (P.P.M. Leseman).

http://dx.doi.org/10.1016/j.learninstruc.2016.10.003

et al., 2006; 2009). Achievement emotions are emotions associated with competence-relevant activities or outcomes such as obtained test grades (Pekrun, Elliot, & Maier, 2009). Based on a meta-analysis of research on achievement goals and emotions, Huang (2011) concluded that mastery goals generally evoke positive achievement emotions, performance-avoidance goals evoke negative achievement emotions, and performance-approach goals can evoke both positive and negative achievement emotions in students.

Numerous studies have shown the profound effect of feedback on achievement goals and emotions (Hattie & Timperley, 2007; Ilgen & Davis, 2000; Ilies & Judge, 2005; Kluger & DeNisi, 1996; Mangels et al., 2012). More importantly, these studies also highlight that feedback-related emotions can have significant impact on students' learning. For example, negative achievement emotions can impair reading comprehension and performance on academic tests (Eysenck, Derakshan, Santos, & Calvo, 2007). In addition, anxiety has been found to impair attention and working memory, together with more subtle changes in learning performance, such as less effective organisation of semantic information (Zeidner, 2014, pp. 265–288).

Students are not only affected by feedback *after* their performance. Pekrun et al. (2014) revealed that merely the expectation of a particular type of achievement feedback can already evoke achievement emotions in students. Their results showed that, without knowing their actual performance, students expecting self-referential feedback (i.e., feedback on their personal improvement between a pre- and post-test) generally reported to have mastery achievement goals and positive achievement emotions regarding an announced upcoming test. In contrast, students expecting normative feedback (i.e., feedback on their test performance relative to their fellow students) generally reported to have performance goals, along with both negative and positive emotions. Students who had formed performance-avoidance goals reported more negative emotions, whereas students who had formed performance-approach goals reported more positive emotions.

In sum, there is strong evidence that (anticipated) feedback affects students' academic performance (for a review, see Valiente, Swanson, & Eisenberg, 2012). However, research on the neurocognitive mechanisms underlying the effects of feedback on cognitive processes is lacking.

1.2. Semantic processing and emotions

Cognitive processes, such as effortful control and depth of processing during task performance, play a large role in the relationship between achievement emotions as induced by (anticipated) academic feedback, and academic performance (Valiente et al., 2012). In the present study we focus on the relationship between anticipated achievement feedback and semantic processing. The semantic processing system is responsible for understanding concepts and giving meaning to incoming stimuli, and can therefore be considered highly important for students' learning.

1.2.1. Semantic processing and N400

Semantic processing has a measurable neural correlate in the form of the N400-event related potential (ERP). It can be measured with an EEG and typically appears at the centro-parietal sites of the scalp (for a review, see Kutas & Federmeier, 2011). In a typical N400 paradigm, also used in this study, sentences are presented that differ in semantic congruency. The magnitude of the N400 amplitude reveals to what extent a person judges a sentence as semantically plausible: the more semantically unexpected the sentence-ending, the more negative the deflection around 400 ms after presentation of the final word (Delong, Urbach, Groppe, & Kutas, 2011). For example, the sentence "*In that library the pupils borrow pillows*" has a semantically unexpected ending: the final word has a very low so-called cloze probability. On the other hand, the sentence "*In that library the pupils borrow books*" semantically makes sense; the final word has a very high cloze probability. Hence, if one processes the previous two sentences, the N400 deflection for the low-cloze (LC) sentence-ending should be more negative than the deflection for the high-cloze (HC) sentence-ending. The relative difference in amplitudes between high-cloze (HC) and low-cloze (LC) sentences is called the N400 cloze effect.

The exact interpretation of the N400 effect is still debated (Lau, Phillips, & Poeppel, 2008). The lexical view relates N400 effects to differences in access to semantic memory representations. In this view, a large N400 cloze effect reflects a facilitated activation of relevant networks of lexical or conceptual representations, which function as a prediction mechanism for the sentence-endings. Alternatively, the integration view states that the N400 effect reflects the process of semantic integration of a target word with the working model of the sentence meaning *after* reading the critical word. Despite these interpretation differences, both views imply that a larger N400 cloze effect can be interpreted as stronger cognitive involvement during sentence reading.

1.2.2. N400 and emotions

Recent electrophysiological evidence has revealed that even mild changes in emotional state can affect the way people semantically process language (e.g., Chwilla, Virgillito, & Vissers, 2011; Egidi & Nusbaum, 2012; Federmeier, Kirson, Moreno, & Kutas, 2001; Pinheiro et al., 2013; Van Berkum, de Goede, van Alphen, Mulder, & Kerstholt, 2013). For instance, Chwilla et al. (2011) manipulated emotional states using either happy or sad video clips. For participants who had watched the happy clips, a large and widespread bilateral N400 cloze effect was present, whereas for the participants who had watched sad clips, the cloze effect was strongly reduced at the left hemisphere and the midline sites. Based on these results, Chwilla et al. (2011) concluded that a positive mood state facilitates subjects' semantic processing. On the other hand, Federmeier et al. (2001) and Pinheiro et al. (2013) used emotion pictures to induce mood states and compared expected sentence-endings with somewhat unexpected sentence-endings. Federmeier et al. (2001) found that a positive mood —when compared to a neutral mood— led to a more positive N400 amplitude for the unexpected sentence-endings. Based on this increase, the authors argued that a positive mood may activate a richer set of semantic properties for upcoming items and therefore a facilitated ability to flexibly accommodate the unexpected sentence-endings. Pinheiro et al. (2013) found that in a positive mood, participants showed equal amplitudes for both expected sentence-endings and somewhat unexpected sentence-endings, both being more positive than the unexpected sentence-endings. Based on these results, the authors also argued that a positive mood may broaden the predictive mechanisms or that words in the semantic memory become more alike. Thus, although the results of Federmeier et al. (2001) and Pinheiro et al. (2013) are not completely similar, both studies conclude that a positive mood broadens predictive mechanisms. These findings are, however, hard to compare with Chwilla et al. (2011) as they did not use the N400 cloze effect as an outcome measure and differed substantially in methods. For example, instead of comparing expected versus highly unexpected (implausible) sentences, Federmeier et al. (2001) and Pinheiro et al. (2013) compared expected sentences with two categories of somewhat more unexpected (but still plausible) sentence-endings. In conclusion, the results on the effect of affective state on the N400 are not fully consistent across studies but they consistently reveal that even a mild change in emotional state already affects the way in which meaning of language is processed.

1.3. The present study and hypotheses

In sum, anticipated achievement feedback can affect students' achievement goals and emotions. Even a mild change in emotional state can already affect semantic processing. Together, these results suggest that anticipated achievement feedback might affect semantic processing. The present study was designed to test this hypothesis. We compared the effects of anticipated self-referential feedback and anticipated normative feedback on students' semantic processing by means of the N400 cloze effect, using the method of Chwilla et al. (2011). In line with their results regarding the effects of watching either a happy or a sad video, we hypothesised that anticipated self-referential feedback would result in a larger, more widely distributed cloze effect than anticipated normative feedback, as prior research has shown that anticipating self-referential feedback was associated with more positive emotions than normative feedback (Pekrun et al., 2014).

In addition to the anticipated feedback manipulation, we also included a performance-expectation manipulation because prior research has shown that anticipated normative feedback can evoke both positive and negative achievement emotions, depending on students' performance goals (Pekrun et al., 2014). We posited that low performance expectations evoke performance-avoidance goals and increase the chance of inducing solely negative achievement emotions as a result of anticipated normative feedback. Therefore, prior to the test on which the feedback was expected, half of the participants (randomly assigned) received a pre-test in which good performance was impossible, which was expected to induce low performance expectations for the ‘actual’ test. The other half received a pre-test in which knowledge of test performance was not possible, which was expected not to induce any specific performance expectations for the ‘actual’ test. We hypothesised that – within the normative feedback group – negative performance expectations would result in a smaller cloze effect compared to neutral performance expectations.

2. Method

2.1. Participants

The sample consisted of 91 recently enrolled first-year students from the Faculty of Social and Behavioural Sciences at Utrecht University. Following Chwilla et al. (2011), only females were included, because there are indications that female and male participants process meaning differently depending on emotional state (e.g., Federmeier et al., 2001). Participation was voluntary and rewarded with a gratification of ten euro. All participants were right-handed native speakers of Dutch without a diagnosed reading disability or a mental, neurological or chronic physical illness. All participants gave written informed consent before the start of the study. Twelve participants were excluded from the analyses: four because of technical problems and eight participants lost too many segments due to to EEG artefacts (see Section 2.6). The final analyses included 79 participants (age: $M = 18.71$ years, $SD = 1.29$). The study was approved by the ethics committee of the faculty.

2.2. Experimental conditions

The participants were randomly assigned to the 2 × 2 experimental conditions (see Table 1).

2.2.1. Anticipated feedback

The participants were informed they would receive performance feedback on a two-part aptitude test concerning statistics. In reality, the test did not measure statistical insight (see Section 2.2.2). We chose for statistical content because many first-year university students in the social and behavioural sciences experience statistics anxiety (Onwuegbuzie & Wilson, 2003) and we expected this test to evoke relatively strong achievement emotions. The test was framed as an indicator of future success in the obligatory course of Methods and Statistics with the intention to increase the value participants would attach to their performance on this test. Depending on feedback condition, the participants anticipated either self-referential feedback ($n = 40$) or normative feedback ($n = 39$). In the self-referential condition, a written test instruction informed the participants they would receive feedback on their improvement (“Once you have completed Part 2, you will receive feedback based on your personal level of progress. You will hear how much you have improved compared to your performance on Part 1”). The participants in the normative feedback condition read that they would receive feedback on their performance relative to other students (“Once you have completed Part 2, your performance will be evaluated in terms of norm scores. You will hear whether you scored within the range of the highest 5%, 20%, or 50%, or the lowest 50%, 20%, or 5% of all students”).

2.2.2. Performance expectations

In order to manipulate participants' performance expectations across the feedback conditions, half of the participants ($n = 40$) received a difficult first part of the aptitude test. The test consisted of five different types of multiple choice problems: figure series (6 ×); number series (5 ×); numerical reasoning (3 ×); syllogisms (6 ×); and analogies (5 ×). The items were presented on a computer screen and had to be completed within a very short time frame. For some problems the correct answer was not among the answering options, making it impossible to perform well. The total duration of the test was 10 min. The participants did not receive feedback on this first part, but a pilot study indicated that the test indeed created a feeling of poor performance. The assumption was that, since they performed poorly on the first part of the test, participants would have a low performance expectation for the second part of the test, on which they expected feedback. This condition will be referred to as the *low-expectation condition*.

The other half of the participants ($n = 39$) completed a first part of a test in which they could not be aware of their test performance. The test was comparable to the Digit Symbol Substitution Test from the Wechsler Intelligence Scales (Wechsler, 1981). The participants had to match as many statistical symbols to digits as possible within a time frame of 5 min. Moreover, one adjustment was made to make the test a little bit more complicated: if a digit was in boldface, the participants had to subtract the boldfaced digit from the number 12 and match the resulting number to the

Table 1
The 2 × 2 experimental conditions and sample sizes.

	Self-referential feedback	Normative feedback	Total
Low performance expectation	21	19	40
No performance expectation	19	20	39
Total	40	39	79

corresponding statistical symbol. This condition served as a control condition that would not evoke specific expectations about performance on the second part of the test. This condition will be referred to as the *no-expectation condition.*

Prior to the experiment, all participants completed an online questionnaire at home, in which they were asked about their achievement goals with regard to the course Methods and Statistics (Achievement Goal Questionnaire, Elliot & Murayama, 2008). There were no significant differences between the four experimental conditions in prior mastery goals, $F(3,75) = 0.73$, $p = .536$, $\eta_p^2 = 0.028$, or performance goals, $F(3, 75) = 0.64$, $p = .594$, $\eta_p^2 = 0.025$.

2.3. Semantic processing

In order to measure the effect of anticipated feedback on semantic processing, an N400 ERP task was administered at two moments: a baseline before feedback information had been provided (T1); and immediately after the moment participants were informed about the type of feedback they would receive (T2). During both tasks, the participants were presented with 65 declarative sentences in Dutch. The two sets of sentences were counterbalanced across time points. The final word of each sentence varied in cloze probability. Half of these sentences ended with highly expected words: the high-cloze (HC) condition. The other half ended with highly unexpected words: the low-cloze (LC) condition. The sentences were adapted from the study of Nieuwland and Van Berkum (2006).

2.3.1. Procedure

The sentences were presented in serial visual presentation mode at the centre of the computer screen. They were presented in random order and each sentence was presented only once. Words were presented in black capital letters on a white background. Each sentence started with a fixation cross (2600 ms), followed by a blank screen (300 ms). Duration of each word was 345 ms; only the final words were presented longer (600 ms). Blank screens (300 ms) were presented in between words. The list of sentences was split up into two sub-blocks; there was a pause of 10 s during which a text informed the participants they were halfway through the task. The duration of both ERP tasks was 15 min.

2.3.2. Memory test

To increase engagement, participants were asked to carefully attend to all sentences (for a similar procedure, see Federmeier et al., 2001) as a memory test would be administered after each ERP measurement. The memory test consisted of a paper sheet containing fifteen sentences drawn from the ERP task. These sentences were presented without the final word. Participants were asked to complete as many sentences as they could within 2 min. The number of correct answers was counted.

2.4. Achievement emotions

To check whether the achievement emotions had been successfully induced, participants' prospective emotions regarding Part 2 of the aptitude test were assessed with an adapted and translated version of the Achievement Emotions Questionnaire (AEQ; Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011). Respondents were instructed to report their emotions before taking Part 2 of the test on a five-point Likert scale (1 = strongly disagree – 5 = strongly agree). The adaptation of the questionnaire included the removal of nine questions that concerned participants' preparation for the test (e.g., “I'm so proud of my preparation that I want to start the exam now”), as preparation was not required for the test, and the replacement of the word *exam* by *statistics test*. The questionnaire consisted of fourteen questions: five addressed the positive achievement emotions *hope* (3 items; e.g., “I have great hope that my abilities will be sufficient in order to complete the statistics test”; Cronbach's α in the current study = .89) and *enjoyment* (2 items; e.g., “I look forward to the statistics test”; Cronbach's α = .61), and nine questions addressed the negative achievement emotions *anxiety* (5 items; e.g., “I worry whether the statistics test will be too difficult”; Cronbach's α = .80) and *hopelessness* (4 items; e.g., “I have lost all hope that I have the ability to do well on the statistics test”; Cronbach's α = .83).

2.5. Procedure

Participants were seated in a separate room in the laboratory. During electrode placement the participants were informed about the study procedures. They were informed that the aim of this study was to improve the course Methods and Statistics. To increase the credibility of the ERP task in this context, we explained that sentence comprehension is also thought to be related to statistical insight. Immediately after the EEG-preparation, the baseline measurement (T1) of the ERP task started.

After the first ERP task, participants were informed that they now had to complete Part 1 of a two-part aptitude test which would predict their future success in the course Methods and Statistics (performance-expectation manipulation); written instruction informed them that they would only receive feedback on Part 2 of the test and included the type of feedback they would receive (feedback manipulation). Directly after the first part of the aptitude test, the second measurement (T2) of the ERP task started. The computer screen provided participants with the same instruction as during the baseline measurement but they were now also reminded about the feedback they would receive on Part 2 of the statistics test, which was said to be administered after the second ERP measurement. During the pause between the two blocks of sentences, the text on the computer screen repeated the type of feedback participants would receive after completing the second part of the aptitude test. After finishing the second ERP task, participants filled out the adapted AEQ that served as a manipulation check. Finally, the participants were informed that the study had a different research goal from what was explained earlier and that they did not have to complete the second part of the aptitude test. All participants were asked what they thought the actual research aim could be. Although three participants mentioned the word “feedback” in their answer, no participant was able to answer this question correctly. Participants were then debriefed on the real purpose of the study.

2.6. EEG recording, preprocessing, and analysis

The ERP data were recorded with 32 active Ag-AgCl electrodes mounted on an elastic electrode cap (Biosemi Active 2 system). Electrode impedance was kept below 50 kΩ. The signals were amplified and digitized on-line at 2048 Hz. To measure eye movements, additional electrodes were placed below the left eye and at the left and right temple. After recording, the EEG signals were imported into the Matlab-based Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). First, the data were down-sampled to 512 Hz. Subsequently, signals were detrended and filtered with a 0.5–40 Hz band-pass filter and referenced to the mean of the left and right mastoids. The signals were corrected for horizontal and vertical eye movements by employing the Independent Component Analysis method (Jung et al., 2000). Next, all channels were screened for artefacts. When a channel deviated substantially from the other channels, while the signal in other

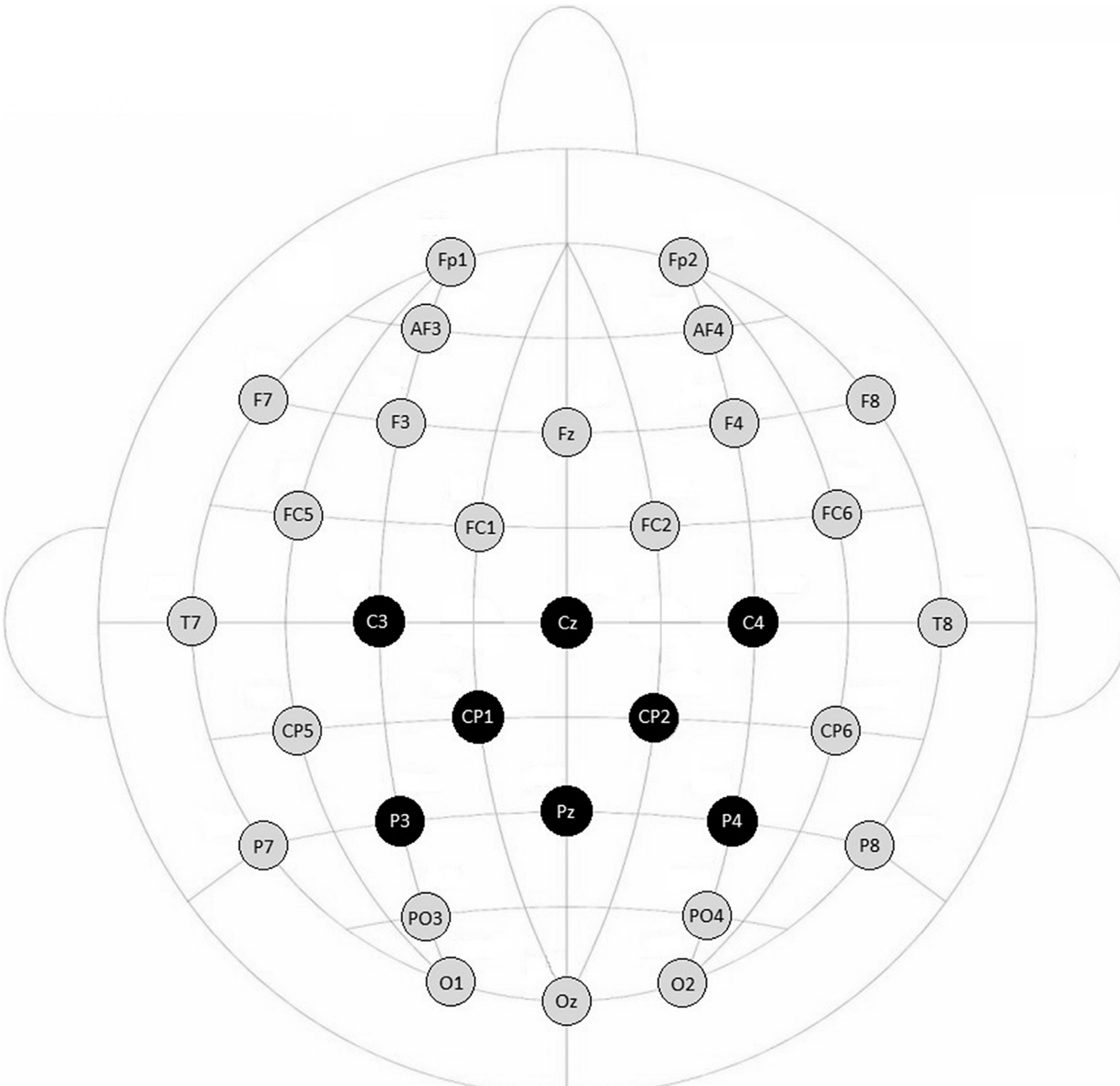

Fig. 1. Birds-eye view of the electrode layout on an idealised head. The centro-parietal sites (used for analyses), are displayed in black.

channels did not contain artefacts in a substantial amount of the trials, this channel was marked as a bad channel. Bad channels were reconstructed based on a linear combination of surrounding channels (bad channels were never neighbouring channels). Data were segmented per condition (HC vs. LC) from 100 ms before to 1000 ms after the onset of the critical word. Segments were baseline corrected and then manually screened for artefacts. For all participants, a minimum of 20 trials in each condition was required. As reported before (Section 2.1.), eight participants did not meet this criterion and were excluded from the analyses. For T1, a mean of 24.90 ($SD = 3.11$) and 24.69 ($SD = 2.77$) trials were available for computation for the HC and LC condition respectively. For T2, a mean of 25.51 ($SD = 3.02$) and 24.43 ($SD = 2.63$) trials per condition were available for computation.

Mean amplitudes in the 300–500 ms epoch after the critical word onset were used to measure the N400 component in the analyses. Because the N400 is largest over the centro-parietal sites (Kutas & Federmeier, 2011), analyses were restricted to the cloze effects at these electrodes (see Fig. 1). The cloze effects were computed by subtracting LC amplitudes from the HC amplitudes. A high value (i.e., large difference) reflected a large cloze effect. A four-way mixed ANOVA was carried out, with Time (Tl vs. T2) and Site (C3, Cz, C4, CP1, CP2, P3, Pz, and P4) as within-subjects factors and Anticipated Feedback (self-referential vs. normative) and Performance Expectation (low vs. no) as between-subjects factors.[1] Assumptions of normality and homogeneity of variance were checked and met. No multivariate outliers were detected. Whenever sphericity could not be assumed, a Greenhouse-Geisser correction was applied (Greenhouse & Geisser, 1959). Corrected *p*-values are reported along with the original degrees of freedom. All ANOVAs were conducted in SPSS version 22.0. Post hoc analyses (see Sections 3.2.2. and 3.2.3.) were conducted using the JASP software (JASP Team, 2016) as the assumption of equal variances is dropped in SPSS when conducting post hoc analyses, which was considered inadequate.

[1] As an alternative to using difference scores, we conducted the analyses also with high vs. low cloze probability as an extra within-subjects factor, with highly similar results. For the sake of readability, these results are not included here, but can be found in the Supplemental Materials.

Table 2
Overview of the means (M) and standard deviations (SD) on the achievement emotions questionnaire (AEQ) for the experimental conditions.

Emotions[a]	Low performance expectation				No performance expectation			
	Self-referential feedback		Normative feedback		Self-referential feedback		Normative feedback	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Hope	2.84	0.85	2.84	0.76	3.49	0.56	3.52	0.60
Enjoyment	2.69	0.81	2.71	0.63	3.11	0.64	3.10	0.58
Anxiety	2.21	0.68	2.20	0.75	1.79	0.62	1.78	0.81
Hopelessness	2.01	0.77	1.97	0.67	1.41	0.45	1.39	0.50

Note. [a] Range 1.00–5.00.

3. Results

3.1. Manipulation check

The effects of anticipated feedback (Feedback) and performance expectation (Expectation) on participants' self-reported emotions for Part 2 of the aptitude test were analysed using a two-way MANOVA with the four types of achievement emotions as dependent variables: hope, enjoyment, anxiety and hopelessness. The reported effect size is partial eta-squared (η_p^2), for which 0.01 is considered small, 0.06 medium, and 0.14 large (Cohen, 1988). The descriptive statistics for the experimental conditions are reported in Table 2.

Using Pillai's trace, the effect of Feedback was not significant for any of the reported achievement emotions, $V = 0.00$, $F(4, 72) = 0.02$, $p = .999$, η_p^2 s $= 0.001$. Nevertheless, there was a significant effect of Expectation on the achievement emotions, $V = 0.24$, $F(4, 72) = 5.56$, $p = .001$, η_p^2 s $= 0.236$. Participants in the low-expectation condition, reported less hope and enjoyment, and more anxiety and hopelessness regarding Part 2 of the test than participants in the no-expectation condition, hope: $F(1, 75) = 17.39$, $p < .001$, $\eta_p^2 = 0.188$; enjoyment: $F(1,75) = 7.04$, $p = .010$, $\eta_p^2 = 0.086$; anxiety: $F(1, 75) = 6.77$, $p = .011$, $\eta_p^2 = 0.083$; hopelessness: $F(1,75) = 18.53$ $p < .001$, $\eta_p^2 = 0.198$. No significant interaction effects between Feedback and Expectation were found, $ps \geq .934$, η_p^2 s $< .001$.

3.2. ERP results

The waveforms for T1 and T2 are displayed in Fig. 2 and Fig. 3, respectively. The grand average ERPs are presented separately for the experimental conditions. The waveforms revealed that the critical words elicited a broad negative-going deflection around 400 ms, the N400 component. The difference in amplitude between the HC and the LC condition in the N400 windows suggested the presence of an N400 cloze effect, which was confirmed by two dependent *t*-tests for both time points, conducted on the mean HC and LC amplitudes (average of the eight electrodes). Both at T1 and T2, LC sentences elicited significantly stronger negative deflections than HC sentences, T1: $t(78) = 9.42$, $p < .001$, $d = 1.06$ T2: $t(78) = 7.77$, $p < .001$, $d = 0.87$.

3.2.1. Comprehensive analysis

A four-way mixed ANOVA was conducted (Table 3, Comprehensive analysis). As hypothesised, the results revealed a significant four-way interaction with a small to medium effect size between Site, Time, Feedback, and Expectation. This suggested a difference in distribution of the cloze effect over the electrode sites, between the two time points, the anticipated feedback conditions, and the

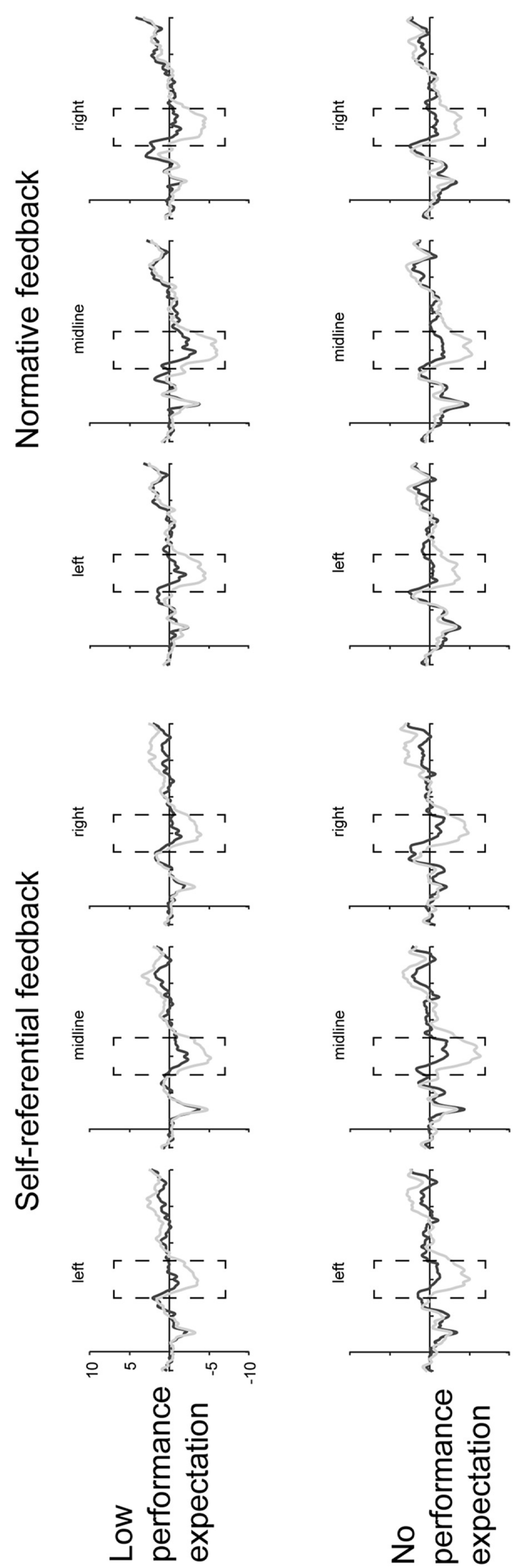

Fig. 2. Grand ERP averages at the centro-parietal sites for the anticipated feedback and performance-expectation conditions at T1, time locked to the onset of the critical word of both sentence conditions (high-cloze and low-cloze). The dashed rectangles indicate the time window (300–500 ms) in which the N400 cloze effect was measured.

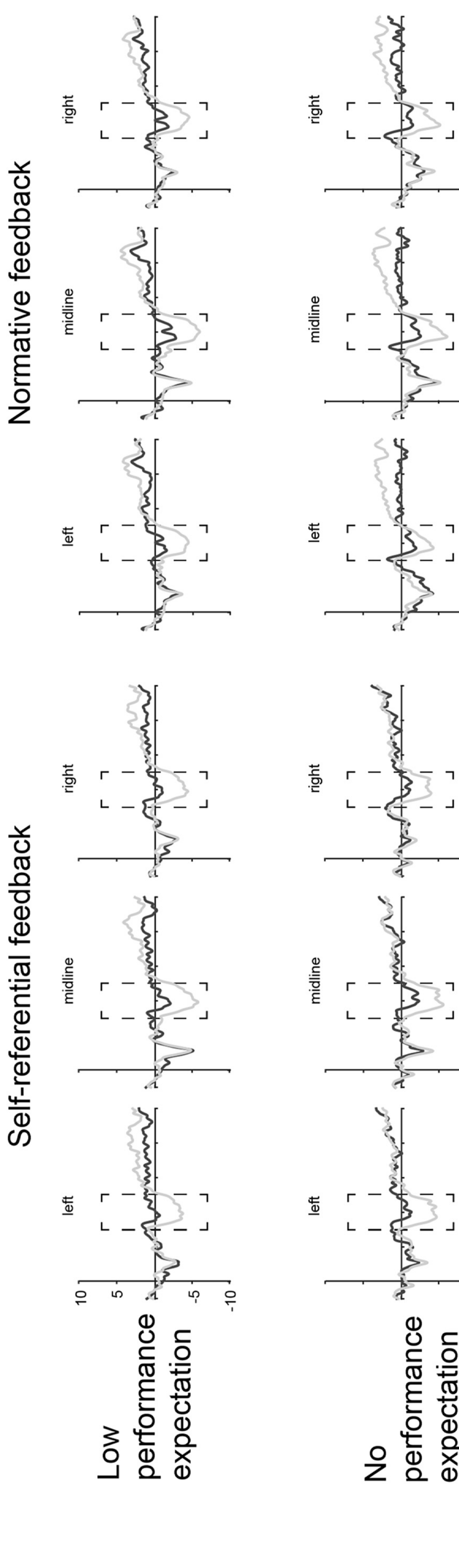

Fig. 3. Grand ERP averages at the centro-parietal sites for the anticipated feedback and performance-expectation conditions at T2, time locked to the onset of the critical word of both sentence conditions (high-cloze and low-cloze). The dashed rectangles indicate the time window (300–500 ms) in which the N400 cloze effect was measured.

performance-expectation conditions. A separate three-way mixed ANOVA for T1 with the factors Site, Feedback, and Expectation (Table 3, T1) showed no significant main or interaction effects, indicating that, at the baseline measurement, the size of the cloze effect was equal across the sites, and did not differ between the feedback conditions and the performance-expectation conditions ($M = 2.71$ μV, $SD = 2.56$).

In contrast to T1, the three-way mixed ANOVA for T2 (Table 3, T2) showed a significant three-way interaction with a medium effect size between Site, Feedback, and Expectation, indicating that, after the experimental manipulations, the distribution of the cloze effect over the sites differed between the feedback conditions and the performance-expectation conditions. To break down this three-way interaction, we conducted two separate two-way mixed ANOVAs for both performance-expectation conditions with the factors Site and Feedback.

3.2.2. Low-performance expectation

The ANOVA for the low-expectation condition showed no significant interaction or main effects of Feedback (Table 4, Low-expectation condition), suggesting that –in contrast to our hypothesis– the N400 results did not differ between the two feedback groups within the low-expectation condition. Furthermore, a significant medium-sized main effect of Site was present, indicating a difference in distribution of the cloze effect over the sites. Bonferroni post hoc tests revealed one significant difference: the cloze effect at Cz (midline) was larger than at P4 (right hemisphere), $p < .001$. The remaining sites did not differ significantly in their size, $ps \geq .146$, indicating that size of the cloze effect was more or less equal across the sites ($M = 2.92$ μV, $SD = 3.11$).

3.2.3. No-performance expectation

The ANOVA for the no-performance expectation condition showed a large-sized interaction effect between Feedback and Site (Table 4, No-expectation condition), suggesting that –in contrast to the low-performance expectation group– the distribution of the N400 effect over the sites differed between the feedback conditions. To break down this two-way interaction, we conducted two separate repeated measures ANOVAs for both feedback conditions within the no-expectation condition with Site as within-subjects factor.

3.2.3.1. Self-referential feedback. The ANOVA for the self-referential feedback condition revealed a large-sized main effect of Site, indicating a difference in distribution of the cloze effect over the sites, $F(7, 126) = 3.39$, $p = .018$, $\eta_p^2 = 0.159$. Bonferroni corrected post hoc tests revealed two significant, and one marginally significant difference. The cloze effects at C3 and CP1 (left hemisphere) were both larger than the cloze effect at P4 (right hemisphere), $ps \leq .040$. Furthermore, the cloze effect at CP1 was larger than at C4, $p = .066$. The remaining sites did not differ significantly in their size, $ps \geq .132$. In sum, the cloze effect in the self-referential condition was not equally distributed, with a stronger effect at the left hemisphere ($M_{\text{left hemisphere}} = 3.43$ μV, $SD = 3.55$; $M_{\text{midline}} = 3.09$ μV, $SD = 3.84$; and $M_{\text{right hemisphere}} = 2.36$ μV, $SD = 3.47$).

3.2.3.2. Normative feedback. The ANOVA for the normative feedback condition also revealed a large-sized main effect of Site, indicating a difference in distribution of the cloze effect over the sites, $F(7, 133) = 4.51$, $p = .002$, $\eta_p^2 = 0.192$. Bonferroni corrected post hoc tests revealed seven significant, and one marginally significant difference: the cloze effects at Cz, Pz (midline), CP2, C4, and P4 (right hemisphere) were all larger than the cloze effect at P3 (left hemisphere), $ps \leq .040$. Furthermore, the cloze effects at Cz and P4 were both larger than at C3 (left hemisphere), $ps \leq .046$, and the cloze effect at C4 was larger than

Table 3
Results of the four-way mixed ANOVA on time, site, feedback, and expectation (comprehensive analysis) and the results of the two separate three-way mixed ANOVAs on site, feedback and expectation for T1 and T2.

Source	Comprehensive analysis				Split by measurement occasion							
					T1				T2			
	F	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2
Main effects												
Time	0.00	1, 75	.992	0.000								
Site	3.64	7, 525	.006	0.046	1.96	7, 525	.097	0.025	3.84	7, 525	.003	0.049
Feedback	0.38	1, 75	.541	0.005	0.01	1, 75	.939	0.000	0.68	1, 75	.414	0.009
Expectation	0.00	1, 75	.957	0.000	0.31	1, 75	.579	0.004	0.27	1, 75	.602	0.004
Interaction effects												
Site × Feedback	2.44	7, 525	.044	0.032	1.14	7, 525	.340	0.015	2.58	7, 525	.031	0.033
Site × Expectation	0.65	7, 525	.632	0.009	0.98	7, 525	.423	0.013	0.76	7, 525	.570	0.010
Feedback × Expectation	1.02	1, 75	.315	0.013	1.28	1, 75	.261	0.017	0.10	1, 75	.748	0.001
Site × Feedback × Expectation	2.50	7, 525	.040	0.032	0.45	7, 525	.785	0.006	5.27	7, 525	< .001	0.066
Time × Site	1.68	7, 525	.142	0.022								
Time × Feedback	0.43	1, 75	.513	0.006								
Time × Expectation	0.53	1, 75	.469	0.003								
Time × Site × Feedback	0.91	7, 525	.472	0.012								
Time × Site × Expectation	1.17	7, 525	.316	0.016								
Time × Feedback × Expectation	0.20	1, 75	.655	0.003								
Time × Site × Feedback × Expectation	3.77	7, 525	.015	0.038								

at C3, $p = .060$. The remaining sites did not differ significantly in their size, $ps \geq .100$. In sum, the cloze effect in the normative condition was not equally distributed, with a stronger effect at the right hemisphere ($M_{\text{left hemisphere}} = 1.52$ μV, $SD = 2.84$; $M_{\text{midline}} = 2.50$ μV, $SD = 3.12$; and $M_{\text{right hemisphere}} = 2.49$ μV, $SD = 2.82$).

3.2.3.3. Direct comparison of the feedback conditions. In order to directly compare the size of the cloze effect between the two feedback conditions (within the no-expectation condition), we conducted eight independent *t*-tests (one for each electrode), with feedback condition as independent variable and the size of cloze effect as dependent variable. The analyses revealed one significant and one marginally significant difference, both with a medium effect size. At C3 and P3 (left hemisphere), the mean cloze effects were larger for participants in the self-referential condition than for participants in the normative condition, C3: $t(37) = -2.05$, $p = .048$, $d = 0.66$, P3: $t(37) = -1.94$, $p = .060$, $d = 0.62$. The size of the cloze effect at the remaining sites did not differ significantly between the two feedback conditions, $ps \geq .166$, $ds \leq 0.45$. In sum, the size of cloze effect at the midline and right hemisphere was equal in both feedback conditions, and the size of the cloze effect at the left hemisphere was stronger in the self-referential feedback condition compared to the normative feedback condition.

3.3. Correlation analyses

To gain more insight into the meaning of the cloze effect, exploratory correlation analyses were conducted between the size of cloze effect at the eight electrode**s** and the mean AEQ ratings, and the post-task memory test scores. Not one of the sites correlated significantly with the self-reported emotions on the AEQ, $rs \leq -.13$, $ps \geq .239$. The performance on the memory test, on the other hand, correlated positively with half of the sites (CP1, CP2, P3, Pz), indicating that a larger cloze effect was related to better performance on the memory test, *r*s ranged between .25 and .31, *p*s ranged between .005 and .028.

Table 4
Results of the two-way mixed ANOVAs on site and feedback for the low-expectation and the no-expectation condition.

Source	Low-expectation condition				No-expectation condition			
	F	*df*	*p*	η_p^2	*F*	*df*	*p*	η_p^2
Site	3.31	7, 266	.008	0.080	1.37	7, 259	.246	0.036
Feedback	0.72	1, 38	.724	0.003	0.64	1, 37	.428	0.017
Site × Feedback	1.15	7, 266	.336	0.029	6.44	7, 259	<.001	0.148

3.4. Summary

The AEQ revealed no differences in reported achievement emotions between the feedback conditions, but there was an effect of performance expectations: participants in the low-expectation condition reported a higher level of negative and a lower level of positive emotions than participants in the no-expectation condition. The ERP results revealed a different pattern. Within the low-expectation condition, no effect of anticipated feedback was found and the cloze effect was more or less equally distributed over the sites with a mean size of 2.92 μV. In contrast, within the no-expectation condition a large interaction effect between Feedback and Site was found, indicating a different N400 distribution between the two anticipated feedback conditions. For the self-referential group, the cloze effect was stronger at the left hemisphere with mean cloze effects of 3.43 μV and 2.36 μV for the left and right hemisphere respectively. In contrast, for the normative group, the cloze effect was stronger at the right hemisphere with mean cloze effects of 1.52 μV and 2.49 for the left and right hemisphere respectively. Direct comparison of the feedback groups within the no-expectation condition showed that the cloze effect at the left hemisphere was significantly stronger in the self-referential group. Finally, exploratory correlation analyses did not reveal a relationship between N400 size and the self-reported achievement emotions but did reveal a positive correlation between N400 size and performance on the memory test.

4. Discussion

The aim of this study was to investigate how students' on-line semantic processing was affected by the mere anticipation of a particular type of achievement feedback and whether students' performance expectations played a role in this relationship. ERP results indicated that anticipated feedback affected the way students processed sentence meaning, but only in the condition without induction of low-performance expectations. In line with our hypothesis, participants anticipating self-referential feedback showed a stronger, more widely distributed N400 cloze effect than participants anticipating normative feedback. No effect of anticipated feedback was found when students' performance expectations were experimentally manipulated to be negative. When discussing the effect of anticipated feedback on semantic processing in Section 4.1, we refer to the findings within the no-expectation condition. In Section 4.2 we provide a possible explanation for the non-significant effect of anticipated feedback on semantic processing within in the low-expectation condition.

4.1. Anticipated feedback and semantic processing

In the current study, the differences in semantic processing between the feedback conditions were mostly found in the distribution of the N400 cloze effect over the scalp. For the self-referential group, the cloze effect was stronger at the left hemisphere and in the normative group, the cloze effect was stronger at the right hemisphere. On the basis of visual half field studies, Kutas and Federmeier (2000) concluded that the left and right hemisphere contribute differently to the semantic processing of language. They argued that the left hemisphere deals with the part of semantic memory that makes a prediction about the forthcoming words in a sentence, which enables subjects to process language fast and efficiently. The right hemisphere, on the other hand, applies a plausibility-based integration strategy, that is, a subjective judgement of how much an item “makes sense” in a given context *after* reading the sentence. A combination of both processing strategies seems most efficient. The difference in the distribution of the cloze effect between feedback conditions in the current study suggests that students anticipating self-referential feedback applied both processing strategies but with a stronger focus on the prediction of the forthcoming words (in line with the semantic facilitation hypothesis), as their cloze effects were located at both hemispheres but more strongly at the left hemisphere. In contrast, students anticipating normative feedback seemed more focused on sentence processing after reading the sentence (in line with the integration hypothesis), as their N400 effect was located more strongly at the right hemisphere. Direct comparison of the feedback groups revealed that the cloze effect at the left hemisphere was slightly stronger in the self-referential group; there was no difference between the feedback groups at the right hemisphere. This suggests that the feedback manipulation in the current study mainly modulated left-hemispheric semantic processing, not, or less so, the right hemisphere processing. This result is in agreement with Chwilla et al. (2011), who found a reduced cloze effect at the left hemisphere for participants in the sad-mood condition.

What the N400 effect exactly indicates in relation to emotion and mood, is still debated. Based on a *reduced* negativity for relatively unexpected sentence-endings, some studies conclude that a positive mood facilitates semantic processing by broadening the range of semantic expectations (Federmeier et al., 2001; Pinheiro et al., 2013). However, other studies suggest that an *increased* negativity for unexpected sentence-endings, especially in case of highly unexpected sentence-endings, reflects enhanced controlled processing or searching in memory to find an appropriate interpretation (Nieuwland & Van Berkum, 2006). Differences in findings may also depend on the used sentence designs. There may be a difference in this regard between slightly unexpected (but still possible) endings and very unexpected (and very difficult to integrate) endings, as in the present study. In the latter case, N400 is likely to point to controlled processing including memory search. In line with Chwilla et al. (2011), we conclude that the larger cloze effect for participants in the self-referential feedback condition should be interpreted as a facilitated semantic processing compared to participants in the normative feedback condition who showed a less robust cloze effect. This was also reflected in correlation-analyses which revealed that a larger cloze effect was related to better performance on the memory task, suggesting better encoding as a result of facilitated or more effortful processing.

4.2. Achievement emotions and semantic processing

The two performance-expectation conditions differed significantly on the achievement emotion questionnaire (AEQ). As intended, participants in the low-expectation condition reported a higher level of negative and a lower level of positive emotions than participants in the no-expectation condition. These differences support predictions from the control-value theory that lack of control and negative expectations reduce positive emotions and increase negative emotions. Furthermore, these findings provide validity evidence for the AEQ. However, when related to the feedback conditions and the N400 results, AEQ findings were remarkable for two reasons. First, the two feedback groups within the no-expectation condition clearly differed in distribution of the cloze effect, while the reported emotions with the AEQ did not. Moreover, in contrast with Pekrun et al. (2014), we did not find any effects of anticipated feedback on the AEQ. These findings indicate that achievement emotions as measured with the AEQ may not underlie the observed differences in the N400 effect. This was also reflected in the correlation analyses, which revealed that not one of the achievement emotions correlated significantly with the size of the cloze effect. In N400 research focusing on the relationship between emotion and semantic processing, emotional state is generally operationalised as mood state. However, mood is different from achievement emotions; it has a lower intensity and lacks a specific referent while achievement emotions specifically refer to an achievement situation. In previous N400 research, subjects' mood states were continuously induced using pictures or video clips, and emotion-manipulation checks were administered at several moments during the mood-induction procedure. Contrastingly, in our study, we induced one relatively strong negative achievement emotions at the beginning of the procedure and the manipulation was checked with the AEQ only once at the end of the procedure, to avoid subjects' awareness of particular interest in emotions. We presuppose that achievement emotions are not directly related to the N400 effect. Instead, the difference in N400 effect between the feedback groups in the no-expectation condition might be caused by a difference in participants' mood state that was not captured with the AEQ. Also note that Pekrun et al. (2014), who found effects of anticipated feedback on the AEQ, administered this test three times, with more questions. Finally, an explanation for the fact that we did not find effects of anticipated feedback on semantic processing in the low-expectation condition could be that the induced achievement emotion, as result of the low-expectation test, overruled possible effects of anticipated feedback on subjects' mood state.

To conclude, it seems that the low-expectation test indeed induced more negative achievement emotions but that these did not affect semantic processing, and that the anticipated feedback —if not overruled by a negative achievement emotion— induced a difference in mood state that influenced semantic processing. However, as our manipulation-check did not measure mood state, further research is needed to confirm this explanation.

4.3. Limitations and future directions

One drawback of the current study regards the sample size. For this reason, replication of our findings is necessary before strong conclusions can be drawn. Furthermore, the generalisability of the findings should be examined. A first step would be to include males in future research. Second, future research should also assess to what extent other types of feedback, performance expectations, and task contents evoke mood states and emotions that affect cognitive processes.

A second drawback of the study concerns the ecological validity. Because of the ERP measurements, it was not possible to conduct the study in a real educational setting. Participants were not actually doing an exam and they were aware that their performance would not affect their actual study results. We suspect that the effects of anticipated feedback and achievement expectations might be stronger in a real educational setting. A challenge for future research is to increase the ecological validity of the experimental paradigm while still allowing for on-line processing measures.

Finally, the manipulation check procedure can be improved. As mentioned before, both in achievement-emotion research and in N400 research emotion questionnaires were administered at several times during the procedure, whereas we only administered the AEQ at the end of our procedure. We did this to prevent participants from becoming aware of our particular interest in their emotions. However, this may have lowered the validity and reliability of our manipulation check, making it impossible to check whether we replicated the findings of Pekrun et al. (2014). Moreover, it rendered a mediation analysis between anticipated feedback, achievement emotions, and semantic processing impossible.

4.4. Conclusion

This study is the first to relate findings regarding the emotional consequences of anticipated feedback directly to students' on-line semantic processing. Moreover, this is the first study that investigated the N400 effect in an educational context. In line with Chwilla et al. (2011), and consistent with current explanations of the N400 effect (Lau et al., 2008), we conclude that the expectation of self-referential feedback evokes a stronger cognitive involvement through a facilitated activation of relevant semantic networks that function as a prediction mechanism. In view of this, the widely used system of normative evaluation in educational practice deserves critical reflection. Feedback on individual progress instead of norm-based feedback in learning may facilitate semantic processing and, thereby, improve learning in students.

Acknowledgments

The authors would like to thank Esther van Geffen and the students, Lotte van Steen and Rachelle de Vries, for their assistance during the data collection, and the reviewers for their helpful comments on the manuscript which helped to improve the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.learninstruc.2016.10.003.

References

Ames, C. (1992). Classrooms: Goals, structures and student motivation. *Journal of Educational Psychology, 84*, 261–271. http://dx.doi.org/10.1037/0022-0663.84.3.261.

Van Berkum, J. J., De Goede, D., Van Alphen, P. M., Mulder, E. R., & Kerstholt, J. H. (2013). How robust is the language architecture? The case of mood. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00505.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology, 79*, 474–482. http://dx.doi.org/10.1037/0022-0663.79.4.474.

Chwilla, D. J., Virgillito, D., & Vissers, C. T. (2011). The relationship of language and emotion: N400 support for an embodied view of language comprehension. *Journal of Cognitive Neuroscience, 23*, 1–15. http://dx.doi.org/10.1162/jocn.2010.21578.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

Delong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology, 48*, 1203–1207. http://dx.doi.org/10.1111/j.1469-8986.2011.01199.x.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040–1048. http://dx.doi.org/10.1037/0003-066X.41.10.1040.

Egidi, G., & Nusbaum, H. C. (2012). Emotional language processing: How mood affects integration processes during discourse comprehension. *Brain and Language, 122*, 199–210. http://dx.doi.org/10.1016/j.bandl.2011.12.008.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70*, 461–475. http://dx.doi.org/10.1037/0022-3514.70.3.461.

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology, 100*, 613.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336. http://dx.doi.org/10.1037/1528-3542.7.2.336.

Federmeier, K. D., Kirson, D. A., Moreno, E. M., & Kutas, M. (2001). Effects of transient, mild mood states on semantic memory organization and use: An event-related potential investigation in humans. *Neuroscience Letters, 305*, 149–152. http://dx.doi.org/10.1146/annurev. psych.093008.131123.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95–112. http://dx.doi.org/10.1007/BF02289823.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*, 81–112. http://dx.doi.org/10.3102/003465430298487.

Huang, C. (2011). Achievement goals and achievement emotions: A meta-analysis. *Educational Psychology Review, 23*, 359–388. http://dx.doi.org/10.1007/s10648-011-9155-x.

Ilgen, D. R., & Davis, C. A. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology, 49*, 550–565. http://dx.doi.org/10.1111/1464-0597.00031.

Ilies, R., & Judge, T. A. (2005). Goal regulation across time: The effects of feedback and affect. *Journal of Applied Psychology, 90*, 453. http://dx.doi.org/10.1037/0021-9010.90.3.453.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*, 163–178.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284. http://dx.doi.org/10.1037/0033-2909.119.2.254.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences, 4*, 463–470. http://dx.doi.org/10.1016/S1364-6613(00)01560-6.

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. http://dx.doi.org/10.1146/annurev.psych.093008.131123.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De) constructing the N400. *Nature Reviews Neuroscience, 9*, 920–933. http://dx.doi.org/10.1038/nrn2532.

Mangels, J. A., Good, C., Whiteman, R. C., Maniscalco, B., & Dweck, C. S. (2012). Emotion blocks the path to learning under stereotype threat. *Social Cognitive and Affective Neuroscience, 7*, 230–241. http://dx.doi.org/10.1093/scan/nsq100.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research, 1118*, 155–167. http://dx.doi.org/10.1016/j.brainres.2006.08.022.

Onwuegbuzie, A. J., & Wilson, V. A. (2003). Statistics Anxiety: Nature, etiology, antecedents, effects, and treatments: A comprehensive review of the literature. *Teaching in Higher Education, 8*, 195–209. http://dx.doi.org/10.1080/1356251032000052447.

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*. http://dx.doi.org/10.1155/2011/156869.

Pekrun, R., Cusack, A., Murayama, K., Elliot, A. J., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction, 29*, 115–124. http://dx.doi.org/10.1016/j.learninstruc.2013.09.002.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2006). Achievement goals and discrete achievement emotions: A theoretical model and prospective test. *Journal of Educational Psychology, 98*, 583. http://dx.doi.org/10.1037/0022-0663.98.3.583.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology, 101*, 115. http://dx.doi.org/10.1037/a0013383.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions

Questionnaire (AEQ). *Contemporary Educational Psychology, 36*, 36–48. http://dx.doi.org/10.1016/j.cedpsych.2010.10.002.

Pinheiro, A. P., Del Re, E., Nestor, P. G., McCarley, R. W., Gonçalves, Ó. F., & Niznikiewicz, M. (2013). Interactions between mood and the structure of semantic memory: Event-related potentials evidence. *Social Cognitive and Affective Neuroscience, 8*, 579–594. http://dx.doi.org/10.1093/scan/nss035.

JASP. Team. (2016). *JASP (version 0.7.5.5)*.

Valiente, C., Swanson, J., & Eisenberg, N. (2012). Linking students' emotions and academic achievement: When and why emotions matter. *Child Development Perspectives, 6*, 129–135. http://dx.doi.org/10.1111/j.1750-8606.2011.00192.x.

Wechsler, D. (1981). *Manual for the wechsler adult intelligence scale-revised*. New York: Psychological Corporation.

Zeidner, M. (2014). Anxiety in education. In R. Pekrun, & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education*. New York: Routledge.