



# Elaborative feedback: Engaging reward and task-relevant brain regions promotes learning in pseudoword reading aloud

Samantha R. Mattheiss<sup>1</sup> · Edward J. Alexander<sup>2</sup> · William W. Graves<sup>1</sup>

Published online: 5 December 2017  
© Psychonomic Society, Inc. 2017

## Abstract

Although much is known about the cognitive and neural basis of establishing letter-sound mappings in learning word forms, relatively little is known about what makes for the most effective feedback during this process. We sought to determine the neural basis by which elaborative feedback (EF), which contains both reward-related and content-specific information, may be more helpful than feedback containing only one kind of information (simple positive feedback, PF) or the other (content feedback, CF) in learning orthography-phonology (spelling-sound) mappings for novel letter strings. Compared to CF, EF activated the ventromedial prefrontal cortex, implicated in reward processing. Compared to PF, EF activated the posterior middle temporal, superior temporal, and supramarginal gyri—regions implicated in orthography-phonology conversion. In the same comparison, EF also activated the left fusiform gyrus/visual word form area—implicated in orthographic processing. Also EF, but not CF or PF, modulated activity in the caudate nucleus. In a postscan questionnaire, EF and PF were rated as more pleasant than CF, suggesting that modulation of the caudate for EF may be due to the coupling of reward and skill content. These findings suggest the enhanced effectiveness of EF may be due to concurrent activation of reward-related and task-relevant brain regions.

**Keywords** Orthography · Phonology · Reward · Feedback · Reading

The ability to map between orthography (written word form) and phonology (auditory word form) is an essential literacy skill, and deficits in this skill have been linked to poor reading outcomes. While instruction that focuses on orthography-phonology mapping has been shown to be effective, reading deficits persist for many individuals. Moreover, negative affective reactions to reading difficulties may further impair one's progress. While previous research has found positive, specific feedback to be most effective, no studies to our knowledge have tested the use of such feedback specifically with an orthography-phonology task; moreover, no studies to our knowledge have demonstrated neural activation that may be responsible for the effectiveness of this type of feedback. Specifically, while previous research shows activation of

reading network regions for orthography-phonology learning (McCandliss, Cohen, & Dehaene, 2003; Sandak et al., 2004; Schlaggar & McCandliss, 2007), no studies to our knowledge have demonstrated activation of these same regions during the receipt of feedback.

The purpose of this study was to test the effectiveness of positive, specific feedback for learning the correct pronunciation for novel pronounceable letter strings (pseudowords), and to identify neural mechanisms that may be responsible for the effectiveness of this type of feedback. We hypothesized that feedback containing both positive and content-related components would be more effective than feedback with either component alone. Neurally, we hypothesized that the mechanism for the effectiveness of this feedback would be reflected in activation of neural regions implicated in reward processing as well as those implicated in the target skill—orthography to phonology processing.

---

✉ Samantha R. Mattheiss  
samantha.smolin@rutgers.edu

<sup>1</sup> Department of Psychology, Rutgers University, Smith Hall Room 4-206, 101 Warren Street, Newark, NJ 07102, USA

<sup>2</sup> Department of Psychology, Tufts University, Medford, MA 02155, USA

## Orthographic-phonological conversion

Reading instruction generally focuses on orthographic to phonological conversion. In the field of education, this skill is

called “decoding” and has been defined as the ability to use individual letters and combinations of letters (graphemes) to access phonemes (elementary sound units of a word) to compute the overall sound form of the word (Vellutino, Scanlon, & Spearing, 1995). During the early stages of learning to read, children learn grapheme-phoneme (“b”–/b/) mappings and then learn to blend phonemes to form words (/b/–/a/–/t/). As reading progresses, these associations become automatic; they learn to segment words into syllables (Lesch & Martin, 1998), and blend the syllables together to form whole words. More advanced readers tend to apply this blending strategy (Orsolini, Fanari, Tosi, De Nigris, & Carrieri, 2006) and do so more automatically as reading proficiency increases (Morais, 2003).

Several lines of evidence have converged on the consensus that instruction in grapheme-phoneme correspondences is crucial for learning to read. A meta-analysis found that instruction that systematically teaches grapheme-phoneme correspondences and how to blend them to form words was highly effective compared to interventions that teach students to read whole words (Ehri, Nunes, Stahl, & Willows, 2001). Such systematic instruction was also more effective than whole-language approaches, where grapheme-phoneme correspondence is taught unsystematically, such as when the need arises in the context of a story (Ehri et al., 2001). Grapheme-phoneme conversion and syllabic analysis approaches have been found to be effective across multiple studies. For example, training first graders to read words at the phoneme (“b— a—t”) and at the onset-rime (e.g., “b—at”) level, where the onset is the initial consonant or consonant cluster and the rime is the following vowel and consonants, was more effective than training at the whole-word level (“bat”; Haskell, Foorman, & Swank, 1992), suggesting that segmenting words is an effective strategy. Children who received instruction that taught them to decode words by comparing them to other similar words (“bat”–“hat”) with additional instruction in grapheme-phoneme analysis improved more than children receiving instruction without the additional grapheme-phoneme analysis (Ehri, Satlow, & Gaskins, 2009). Adolescents taught to segment words into syllables, compared to those taught to read whole words without such segmentation, showed greater improvements in reading and spelling novel words (Bhattacharya & Ehri, 2004). In the proposed research, we use an orthography-phonology training paradigm, which teaches and reinforces word segmentation and blending of syllables to form words.

Regarding reading problems, research suggests that impaired orthographic-phonological conversion underlies the reading difficulties experienced by individuals with dyslexia (Harm & Seidenberg, 1999; Lyon et al., 2003; Shaywitz & Shaywitz, 2005) and other reading impairments (Vellutino, Fletcher, Snowling, & Scanlon, 2004; Vellutino et al., 1995). Instruction aimed at teaching orthography-phonology

conversion has been found to be highly effective for struggling readers (Tressoldi, Vio, & Iozzino, 2007; Wentink, Van Bon, & Schreuder, 1997; Williams, 1980). Yet despite improvements, many individuals with dyslexia have persistent reading difficulties (Gabrieli, 2009; Shaywitz et al., 1999; Torgesen, 2006). One possibility is that affective reactions to the experience of difficulty in reading tasks may contribute to the persistence of literacy problems. For example, such difficulty may lead to low confidence in one’s ability (Dweck & Bempechat, 1983) and feelings of helplessness (Dweck, 1975). This could result in low motivation and less time spent reading, further interfering with reading progress. Therefore, incorporating feedback that promotes positive affect as well as relevant cognitive skills may help improve outcomes for these individuals. It has been suggested that researchers as well as educators should seek to improve motivation in addition to reading skill among poor readers, given the correlation between motivation and skill level (Morgan & Fuchs, 2007). In the current study, we pair training in orthography-phonology mapping with relevant feedback, with the aim of enhancing the effectiveness of reading instruction due to recruitment of both affective and cognitive processes.

## The neural correlates of orthography-phonology conversion

Reading novel letter strings aloud is expected to activate neural systems associated with orthography-phonology conversion. On the orthographic side, the left ventral occipitotemporal cortex (vOT), part of the fusiform gyrus, is often referred to as the “visual word-form area” (VWFA; Dehaene & Cohen, 2011). Damage to this region has been shown in numerous studies to be associated with pure alexia (Binder & Mohr, 1992; Damasio & Damasio, 1983; Leff et al., 2001), a selective deficit in processing visual input for words that leaves other functions such as writing intact. In healthy participants, activity in this region has been shown to increase for processing letter sequences with increasing similarity to English (Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006) and French orthography (Vinckier et al., 2007), suggesting that this area responds to letter combinations. Thus, studies of participants with and without brain damage seem to converge to support a role for the vOT in orthographic processing.

Neural regions implicated in phonological processing include the left posterior superior temporal gyrus (pSTG; Graves, Grabowski, Mehta, & Gordon, 2007; Graves, Grabowski, Mehta, & Gupta, 2008; Price, 2012; Rumsey et al., 1997). Evidence includes modulation of activity in the pSTG by exposure to pseudowords, with decreasing activation for pseudowords the more they are heard and repeated (Graves et al., 2008), suggesting that this region is recruited

for processing the phonological form of the word. Phonological processing is also impaired in conduction aphasia, and the pSTG was shown to be the only area of overlap between damage leading to conduction aphasia and activation for healthy participants in phonological tasks (Buchsbaum et al., 2011).

Converting orthography to phonology has been shown to involve the left supramarginal (SMG) and posterior middle temporal (pMTG) gyri. Previous studies have shown that damage to the left SMG impairs mapping between phonology and orthography (Alexander, Fischer, & Friedman, 1992; Roeltgen, Sevush, & Heilman, 1983). A functional neuroimaging study that manipulated word characteristics found that activity in the SMG was associated with decreasing bigram frequency, or the frequency with which two letters co-occur (Graves, Desai, Humphries, Seidenberg, & Binder, 2010). Such findings suggest this region may be recruited for converting orthography to phonology.

Struggling readers demonstrate differences in neural activation during reading tasks, especially in regions implicated in spelling-sound mapping. One study found that second graders with dyslexia, compared to typical readers, failed to show specialization of the fusiform gyrus for letters compared to symbols (Maurer et al., 2007), suggesting lack of specialization of this region in dyslexia. Meta-analyses have also found that struggling readers underactivated the MTG, STG, and SMG compared to typical readers (Maisog, Einbinder, Flowers, Turkeltaub, & Eden, 2008; Richlan, Kronbichler, & Wimmer, 2009).

Since the current study involves learning with feedback, we note that research has demonstrated improvements in reading performance following remediation, with corresponding increases in relevant neural regions. Specifically, increases in activity in the left temporoparietal cortex correlated with improvements in reading (Eden et al., 2004; Temple et al., 2003).

While research demonstrates activation of these left-lateralized regions during reading tasks, as well as the relationship between such activity and successful reading, no studies to our knowledge have demonstrated recruitment of regions implicated in orthography-phonology mapping during the receipt of instructional, content-based feedback. Thus, we aimed to determine whether these areas were recruited during feedback. We predicted that instructional, content-based feedback (referred to here as content feedback, CF), as well as feedback that was both rewarding and content based (referred to here as elaborative feedback, EF), would recruit prominent areas of the reading network, including the left MTG, STG, SMG, and vOT.

Increasing activation in these task-relevant regions during the receipt of positive feedback has the potential to enhance activity in such regions for struggling readers, thereby bringing neural activity more closely in line with that of typical readers. Evidence of neural activation in task-relevant regions

during the receipt of feedback would provide a basis for the use of such feedback in the classroom, with the aim of benefiting struggling readers.

## Feedback

Recruitment of task-relevant regions during the receipt of feedback should be more effective when feedback indicates a correct response than when it solely provides skill-content information, as feedback in the form of praise has been shown to be rewarding and promote learning (Hattie & Timperley, 2007). Thus, we predicted that feedback that was both positive and content informative (elaborative feedback, EF) would be more effective than feedback that was solely instructional or content based (content feedback, CF).

Feedback that is both rewarding and content informative should also address the adverse affective reactions to experiencing difficulty reading, which have been shown to contribute to the persistence of literacy problems (Dweck, 1975; Dweck & Bempechat, 1983), resulting in low motivation and less time spent reading. Research shows that feedback promotes learning via both cognitive and affective processes, including promoting positive self-evaluation and self-efficacy, as well as confidence to engage in the task (for reviews, see Hattie & Timperley, 2007; Kluger & DeNisi, 1996). Deci, Koestner, and Ryan (1999) also found that positive feedback increases likelihood that students will return to and persist in the task, and that students receiving positive feedback report higher interest in the task. Another study showed that participants receiving performance feedback compared to no feedback showed increased intrinsic motivation (a factor composed of self-rated enjoyment of the experimental task), increased willingness to return for another study, higher perceived task performance, and greater amount of time spent engaging in an activity similar to the task for which feedback was given (Harackiewicz, 1979). Considering such evidence, we predicted that pairing an affective intervention with relevant cognitive skills, specifically with orthography-phonology (spelling-sound) conversion, will improve reading performance.

We also predicted that rewarding, content-informative feedback (EF) would be more effective than feedback that is rewarding but provides no specific information regarding skill content (positive feedback, PF). This prediction is consistent with evidence showing multiple ways by which feedback can be effective. Specifically, feedback has been shown to be most effective when it (1) encourages self-regulation (Hattie & Timperley, 2007), (2) provides specific information about the use of strategies in a task (Brinko, 1993), (3) emphasizes and praises cognitive processes rather than the learner personally (Brinko, 1993; Kamins & Dweck, 1999), and (4) emphasizes learning as compared to performance goals (Heyman &

Dweck, 1992). The EF condition used here seems to align better with these factors than the PF condition.

One example of elaborative feedback being more helpful than simple positive feedback comes from a study that examined the effectiveness of feedback on a thermo-energy learning task. Participants were asked to engage in a computer-based content-learning task about thermodynamics, then answer 12 multiple-choice questions. They received either simple positive or elaborative feedback from an animated agent. Simple positive feedback informed learners about whether or not their answer was correct, while elaborative feedback additionally informed learners about why their answer was correct or incorrect. Compared to positive feedback, elaborative feedback facilitated learning (Lin, Atkinson, Christopherson, Joseph, & Harrison, 2013). The feedback conditions in that previous study were similar to those in the current study, suggesting that elaborative feedback should be more effective than simple positive feedback for the current study as well (hereafter, we will refer to the simple positive feedback condition as “positive feedback”). Still unknown, however, are the neural mechanisms behind this expected difference.

Although previous research suggests that a combination of positive and specific feedback is most effective, no studies to our knowledge have paired such feedback with an orthography-phonology learning task in order to determine its effectiveness for this essential reading skill. This is what we do in the elaborative feedback (EF) condition tested here. We expected elaborative feedback to result in higher accuracy for reading aloud trained items than feedback that is only instructional or reward based.

## Neural activation for feedback processing

The large-scale neural network associated with reward processing has been shown to include the ventromedial prefrontal cortex (vmPFC), dorsal striatum (caudate and putamen), ventral striatum, amygdala, and dopaminergic midbrain (for reviews, see O’Doherty, 2004; Smith & Delgado, 2015). Given that positive feedback is rewarding (Hattie & Timperley, 2007), it is expected to modulate activity in these regions. The dorsal striatum, for example, has been shown to be comparably activated during the receipt of performance feedback as well as during reward and punishment outcomes on a card-guessing task (Tricomi, Delgado, McCandliss, & McClelland, 2006), thus demonstrating its role in processing feedback and extrinsic reinforcers. Further research has shown activation of the striatum for conditions in which an outcome is contingent on an action, suggesting that it may be recruited for processing reinforcement of action due to a reward, rather than the reward itself (O’Doherty, 2004; Tricomi, Delgado, & Fiez, 2004). Likely due to its role in reward-based learning, activity in the striatum during feedback learning may also be related to

subject performance, with activity in the striatum correlating with subject performance and adjustment of their responses due to feedback (Vink, Pas, Bijleveld, Custers, & Gladwin, 2013).

Adding to the novelty of this study is the observation that the literature on the neurocognitive basis of orthography-phonology mapping, and on different aspects of feedback, remain essentially separate. We are aware of no studies that have examined the neural activation associated with the effectiveness of reading-related content plus reinforcement-based feedback, referred to here as elaborative feedback (EF). We expected EF to recruit regions implicated in feedback processing, including the vmPFC and striatum. We also expect it to enhance learning by engaging task-relevant as well as reward-processing regions.

## Age

The main purpose of the study was to test the effectiveness of EF and identify its neural mechanisms. To this end, we sampled from the unimpaired adult population, since these individuals would be able to perform the in-scanner task with minimal complications. Ultimately however, we hope that evidence provided here would motivate future studies to test the generalizability of results to younger populations, especially those actively learning to read. If elaborative feedback is effective for adults, as a next step we feel it would be worth testing for effectiveness in those still learning to read, especially considering previous research demonstrating the effectiveness of positive feedback for children (van Duijvenvoord, Zanolie, Rombouts, Raijmakers, & Crone, 2008). As a first, exploratory step in this direction, we tested whether the effectiveness of the current approach is modulated by age. Specifically, we predicted that relative to the other feedback conditions, the effectiveness of elaborative feedback would be heightened for younger participants.

## The current study

In this study, we compare the ability of different types of feedback to promote learning the pronunciation of novel letter strings (pseudowords). This is done during functional magnetic resonance imaging (fMRI) to determine the neural mechanisms supporting these different types of feedback. Specifically, we tested three different kinds of feedback: simple positive feedback (PF), containing reward-related but not content-related information; content feedback (CF), containing information about the proper blending of letter sounds in the pseudowords; and elaborative feedback (EF), which combines pronunciation-specific and reward-based components. All feedback was given as part of a pseudoword training paradigm, where we expected EF to result in the highest

accuracy. We were then in a position to use fMRI to determine the neural source of the improvements, whether from orthography-phonology based processing, reward-based processing, or both.

## Materials and method

### Participants

Twenty participants were recruited from the Rutgers University–Newark psychology undergraduate student participant pool, the Rutgers graduate student LISTSERV, and Craigslist. One participant reported that he could not hear the feedback and thus was excluded from the study, such that there were 19 remaining participants. All participants provided written informed consent according to Rutgers University Institutional Review Board (IRB) guidelines. Participants were prescreened for no diagnosis of a learning disability, psychiatric, or neurological disorder. The average age of the sample was 24.3 years ( $SD = 3.8$ ), with a range of 19 to 32 years. Fourteen of the participants were female. The average (age-standardized) score on the Wechsler Test of Adult Reading (WTAR), an estimate of verbal intelligence quotient (VIQ; Wechsler, 2001), was 113 ( $SD = 10.49$ ), where the population average is 100. All participants were native English speakers. Four participants were bilingual or multilingual. Participants were compensated for their time at \$30 per hour.

### Stimuli

Stimuli were selected from an existing, normed set of pseudoword stimuli (Gupta et al., 2004), which contains sets of pseudowords ranging from two to seven syllables. A relatively small number of pseudoword stimuli was chosen based on evidence from pilot testing that revealed that 12 to 15 two-syllable pseudowords, each with multiple possible pronunciations, was a reasonable learning task for participants in the given amount of time. The lower end of this range of the 12 to 15 pseudoword range was chosen since the task would be completed while participants were in the scanner, a novel learning environment, with scanner noise.

Three lists of four two-syllable pseudowords, or pronounceable letter strings that are not real words, were selected (see Table 1) so as to control across lists for word characteristics. With word list as the independent variable, three one-way ANOVAs were conducted, yielding no significant differences between word lists in number of letters,  $F(2, 9) = 0.273$ ,  $p > .05$ ; bigram frequency,  $F(2, 9) = 0.462$ ,  $p > .05$  (Miller, Bruner, & Postman, 1954); or biphone frequency,  $F(2, 9) = 0.458$ ,  $p > .05$  (two-phoneme combinations, analogous to bigram frequency; Vitevitch & Luce, 1998). Each list was

**Table 1** The full set of pseudoword stimuli used in the study, selected from Gupta et al. (2004), and surface characteristics controlled across lists. Standard deviations are in parentheses

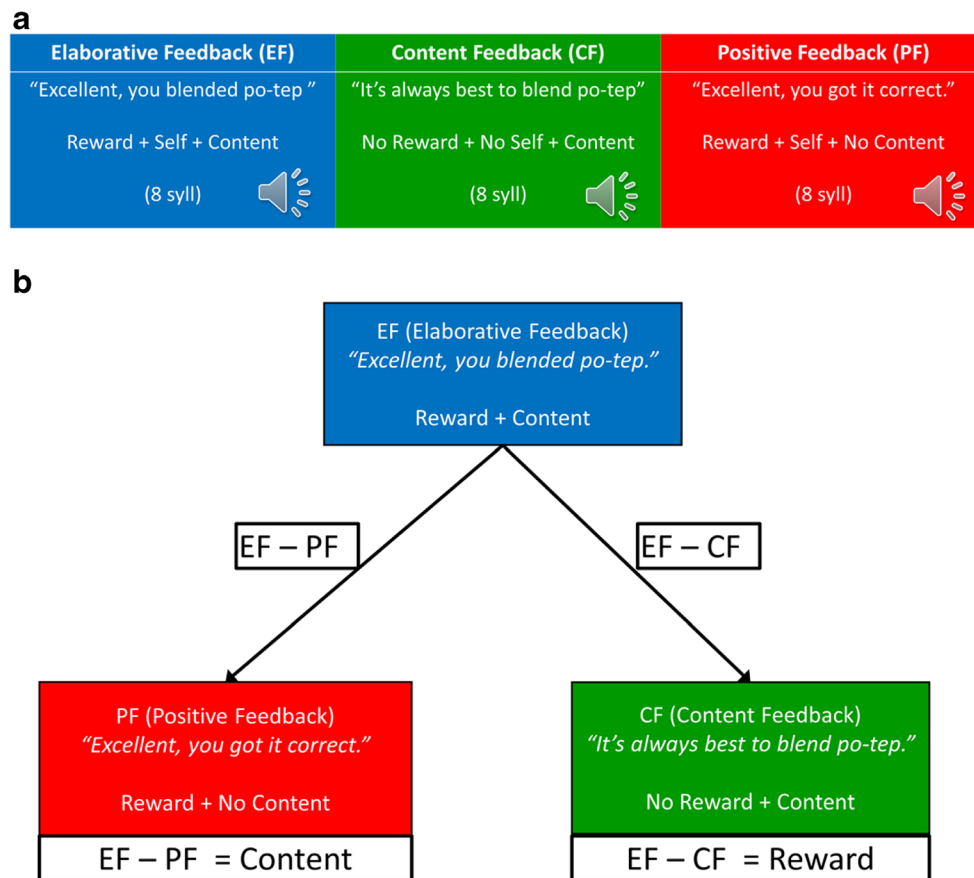
List	A	B	C
<b>Pseudowords</b>	Tesib tinabe desut kegide	Terid botet kopise kedene	gilat bisob butib tosote
<b>Bigram frequency</b>	0.006 (0.002)	0.008 (0.002)	0.007 (0.001)
<b>Biphone frequency</b>	0.002 (0.002)	0.002 (0.001)	0.002 (0.001)

paired with one type of feedback. The pseudoword list/feedback pairs were counterbalanced across participants, so that the type of feedback was paired with each pseudoword list for an equal number of participants. This was done as an additional control to ensure effects of feedback type were not confounded with pseudoword list.

Three types of feedback, provided only on correct trials, were compared in order to identify the influence of content and reward processing. Elaborative feedback (EF) provides the learner with explicit information regarding the skill he/she applied (e.g., “Great! You blended po-*te*p”). Positive feedback (PF; e.g., “Great! You got the answer correct”) provides positive reinforcement but no task-specific information. Content feedback (CF; e.g., “It’s always best to blend po-*te*p”) provides task-specific information without explicitly positive reinforcement. These conditions are illustrated in Fig. 1. Auditory feedback in all three conditions contained exactly eight syllables, and the duration was approximately 4 seconds. Auditory stimuli were recordings of a female voice with an American-English accent. They were recorded using an Audio-Technica ATR-1200 Cardioid Dynamic Vocal/Instrument Microphone, plugged into a laptop computer using Audacity software. The feedback stimuli were matched across feedback conditions, as verified with feedback condition as the independent variable in each ANOVA analysis. They were matched for duration (measured in milliseconds),  $F(2, 33) = 0.106$ ,  $p > .05$ , with EF  $M = 3,197$  ( $SD = 152.8$ ), CF  $M = 3,197$  ( $SD = 8.2$ ), PF  $M = 3,199$  ( $SD = 0$ ); average pitch,  $F(2, 33) = 0.748$ ,  $p > .05$ , with values of EF  $M = 226.88$  Hertz (Hz;  $SD = 3.8$ ), CF  $M = 229.42$  Hz ( $SD = 7.95$ ), PF  $M = 228.09$  Hz ( $SD = 0$ ); and intensity,  $F(2, 33) = 0.531$ ,  $p > .05$ , with values of EF  $M = 76.47$  decibels (dB;  $SD = 1.45$ ), CF  $M = 75.96$  dB ( $SD = 1.32$ ), and PF  $M = 76.14$  dB ( $SD = 0$ ).

### Procedure

There were three phases in the study, all of which were conducted during scanning. During the first phase (the learning phase), each pseudoword appeared once on the screen, in random order, for 1 second. Participants were asked to read the pseudoword aloud while speaking clearly into the



**Fig. 1** **a** The three types of feedback are displayed. Each type was matched on number of syllables, duration (s), pitch (Hz), and intensity (dB). **b** Planned contrasts between the three feedback types

Advanced Noise Cancelling Fiber Optic Microphone for fMRI. After the pseudoword was read aloud, the correct pronunciation was played through the speakers. The participant was instructed to listen carefully to the pronunciation of each pseudoword.

In the second phase (feedback phase), each pseudoword appeared on the screen, and the participant was asked to read it aloud into the microphone (see Fig. 2a). The experimenter sat in the control room and pressed a “y” or “n” for correct or incorrect answers, in order for participants to receive appropriate feedback based on their response. Participants saw an “X” after an incorrect response. After a correct response, participants received one of the three types of feedback, as described above. Each word appeared six times successively, regardless of whether or not the answer was correct.

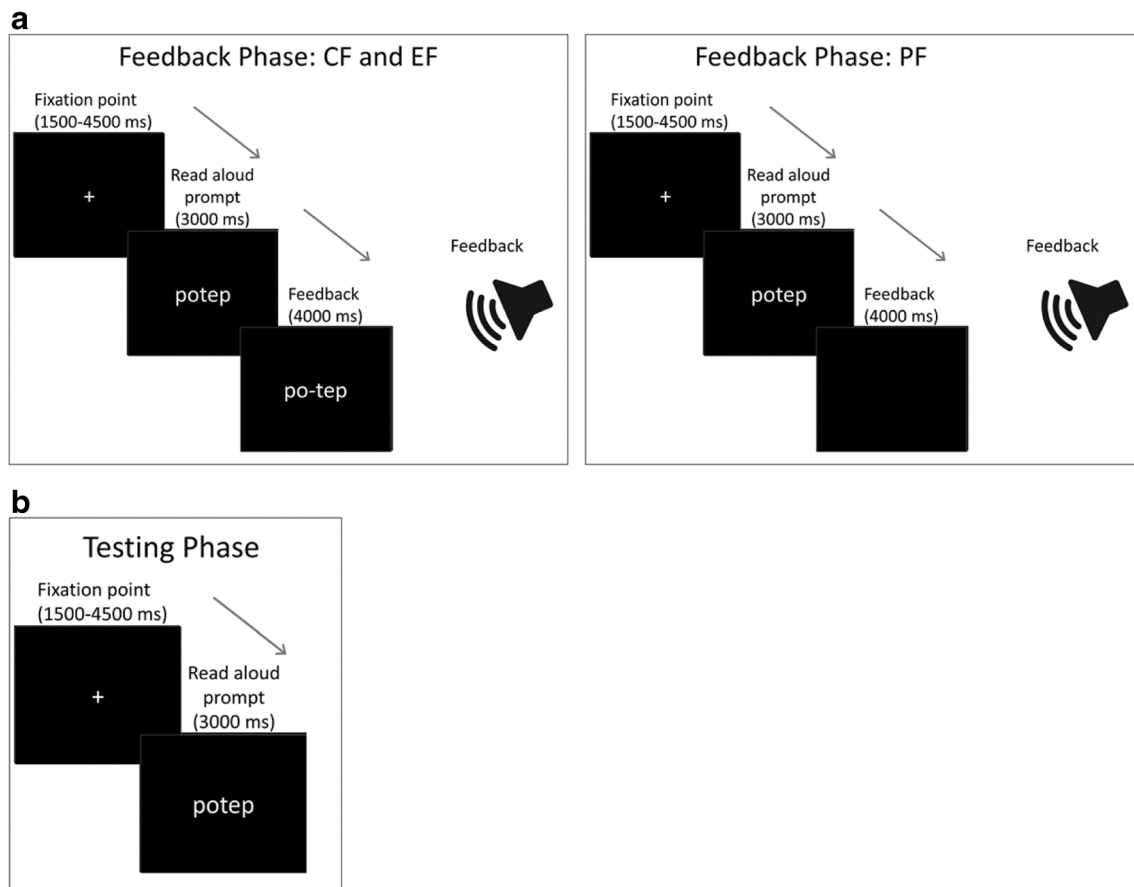
In the EF and CF conditions, the hyphenated pseudoword was presented visually on the screen starting at 2 seconds from the start of the auditory feedback (e.g., “You blended po-tep”). Visual presentation of the pseudoword lasted 2 seconds and ended simultaneously with the completion of the auditory feedback (see Fig. 2a). The experiment was presented using PsychoPy (Peirce, 2007), a standardized, open-source software platform for conducting computerized psychology experiments.

During the last phase (testing phase), each pseudoword appeared once on the screen in randomized order, and the participant was asked to read it aloud (see Fig. 2b). During each of the three phases, a fixation cross appeared on the screen between trials, with randomized interstimulus intervals following a fully event-related design.

To familiarize participants with the procedure prior to the start of the study, participants also engaged in a practice run, which included a shortened version of each of the three study phases.

### Postscan measures

To facilitate comparison of measured performance for each feedback type with participants’ subjective experience, participants were given an 8-item Likert scale questionnaire after the experiment. The scale ranged from 1 to 5, assessing their ratings of pleasantness and helpfulness of each type of feedback, and self-rating their level of reading difficulty as a child and as an adult. An example of an item assessing the helpfulness of a feedback type is “To what degree was the following feedback helpful: ‘It’s best to blend —,’” with responses ranging from 1 (*not helpful at all*) to 5 (*very helpful*). An example of an item assessing the pleasantness of a feedback



**Fig. 2** **a** Feedback phase for content feedback (CF) and elaborative feedback (EF), where audio feedback was paired with visual presentation of the hyphenated pseudoword; Feedback phase for positive feedback (PF), where audio feedback was paired with no visual presentation of the

pseudoword. The duration of each stimulus presentation is shown in milliseconds (ms). **b** Testing phase, where pseudowords were presented on the screen and participants asked to read each pseudoword aloud

type is “To what degree was the following feedback pleasant: ‘It’s best to blend — —’,” with responses ranging from 1 (*very unpleasant*) to 5 (*very pleasant*). Level of reading difficulty as a child and adult were assessed with two separate questions: “To what degree did you experience difficulty reading, as a child (or adult), in learning to read?” Responses ranged from 1 (*no difficulty at all*) to 5 (*severe difficulty*).

## Data analysis

Response times (RT) for participants’ spoken responses during the testing phase were calculated as the duration from the onset of the visual pseudoword display to the onset of participant speech. A rater who was blinded to feedback type analyzed each pseudoword response for timing by listening to and viewing the waveform of the audio file corresponding to each pseudoword response. Accuracy during feedback and testing phases was also recorded. If both the first and second syllable of the pseudoword were pronounced correctly, including correct pronunciation of the vowel sound (short or long), the pronunciation was marked as correct and received a score of

1. If either or both of the syllables were read incorrectly, the item received a score of 0, signifying incorrect. Accuracy was scored by independent raters who were blind to the feedback type with which each pseudoword was trained. A subset of items from the testing phase ( $n = 72$ ) were scored by both raters to ensure interrater reliability, yielding a Cohen’s kappa (Cohen, 1960) of 0.65, a value established by previous studies to be in the acceptable range (Landis & Koch, 1977). Accuracy during the feedback phase was recorded by PsychoPy based on the experimenter’s keypress (“y” or “n”) after listening to the participant’s response.

On two out of 228 total trials (0.9% of trials), the experimenter mistakenly coded accurate participant responses as being inaccurate. The experimenter informed the participant immediately that an experimenter error was made. This occurred during the feedback phase of the experiment. In the behavioral analysis of testing-phase accuracy and reaction time (RT), the data for these two pseudowords were simply excluded from the analysis. For the neuroimaging analysis (discussed further, below), the feedback trials for these two pseudowords were coded separately in the regression analysis

as errors and thus were not included in the feedback condition regressors.

To examine the effectiveness of the different feedback types, a one-way analysis of covariance (ANCOVA) was performed, using accuracy during the testing phase as the dependent variable, accuracy during the feedback phase as a covariate, and feedback category as the independent variable of interest with three levels (EF, CF, and PF). Testing accuracy was calculated by averaging each participant's accuracy for reading aloud the four pseudowords in each feedback condition. Feedback accuracy was calculated by averaging each participant's accuracy for each of the six repetitions of the four pseudowords during the feedback phase. Accuracy during the feedback phase was included as a covariate because feedback was given only for correct answers during this phase, such that the number of correct responses determined the number of times the feedback was received. If, for example, an individual was not successful at pronouncing the pseudoword correctly during the feedback phase, he or she would not have heard the feedback specified for the given target pseudoword. This, in turn, would presumably affect the influence of feedback on performance during the testing phase.

An additional mixed-effects linear-regression model was used to determine the contrast coefficients for each feedback condition, adjusted for the covariate (accuracy during the feedback phase) and for random subject-level variance, following the general approach outlined by Baayen, Davidson, and Bates (2008). In this model, EF and CF were entered as regressors, with PF serving as an implicit baseline for comparison with the other two conditions. As such, the resulting coefficient for EF represents the effect of EF compared to PF; and the coefficient for CF represents the effect of CF compared to PF. This resulted in coefficients representing the effect of each feedback type on testing-phase accuracy, adjusted for feedback-phase accuracy and for random subject variance.

### fMRI data acquisition and analysis

Magnetic resonance imaging data were collected on a 3-T Siemens Magnetom Trio Tim Scanner with a 12-channel head coil. A T1 high-resolution anatomical brain scan was collected for each participant, using a three-dimensional magnetization-prepared rapid gradient-echo (MPRAGE) sequence, with a TR of 1,900 milliseconds (ms) and a TE of 2.52 ms (matrix =  $256 \times 256$  voxels, 176 contiguous 1-mm axial slices, field of view, FOV = 256 mm, flip angle = 9 degrees). Four runs of blood-oxygen-level-dependent (BOLD) data were collected using a gradient-echo echoplanar imaging (EPI) sequence (TR = 2,000 ms, TE = 25 ms, FOV = 208 mm, matrix =  $64 \times 64$ , flip angle = 60 degrees). Each volume consisted of 35 axial slices ( $3.25 \times 3.25 \times 3$ -mm voxels).

All images were preprocessed using the AFNI software suite (<http://afni.nimh.nih.gov/afni>; Cox, 1996). For each

participant, the first six images in each run were ignored due to initial saturation. Slice timing and motion correction were applied to the time series images, and the high-resolution structural scan was then aligned to these images (Saad et al., 2009). Additional noise covariates from motion correction parameters and signal in the ventricles were entered as regressors of no interest. Regressors of interest were included for each feedback condition paired with pseudoword presentation for each feedback type, during the feedback phase, and analyzed using the AFNI program for statistical regression, 3dDeconvolve. Pseudoword stimuli were paired with trial number using the amplitude modulated option in 3dDeconvolve, in order to analyze modulation of neural activity by trial number (Presentations 1–6). Only correct trials were included in the main analysis. Trials with erroneous responses (e.g., mispronounced pseudoword) were modeled separately. Each subject's anatomical scan was aligned to the Talairach atlas (Talairach & Tournoux, 1988), and this alignment solution was applied to align each subject's image regression results to the same atlas. The group analysis was subsequently conducted on these registered images for each condition. A brain mask excluding most white matter and cerebrospinal fluid was applied to all contrast images as well as a smoothing kernel of 6-mm FWHM. A two-tailed  $t$  test with a voxelwise threshold of  $p < .005$  and a cluster correction of  $617 \mu\text{l}$  (mapwise corrected  $p < .05$ ) was used. Note that this voxelwise threshold is more strict than the  $p < .01$  threshold identified by Eklund, Nichols, and Knutsson (2016) as being problematic for controlling false positives in fMRI. The AFNI program 3dttest++ was used to obtain group contrast images. Each individual's WTAR score was entered as a covariate in order to account for any effects of differences in VIQ. An additional analysis examined the correlation between accuracy during the testing round for each particular feedback type and neural activation during pseudoword presentation as well as receipt of feedback.

## Results

### Performance

Performance was analyzed in terms of RT and accuracy. During the testing phase, analysis of RT showed that pseudowords did not reliably differ by feedback training condition (see the rightmost column in Table 2). In terms of accuracy during the testing phase, there was a significant main effect of feedback type,  $F(1, 15) = 3.77$ ,  $p < .05$ . Pseudowords trained with elaborative feedback (EF; e.g., “Excellent, you blended po-tep”) were read aloud during the testing phase with higher accuracy than those trained with content feedback (CF; e.g., “It's best to blend po-tep”) and positive feedback (PF; “Excellent, you got it correct”), as shown in the Testing



**Table 2** Mean accuracy during the feedback phase (Feedback Acc), accuracy during the testing phase (Testing Acc), and response time during the testing phase (Testing RT) for each condition. Standard deviations are in parentheses

	Feedback Acc	Testing Acc	Testing RT
<b>EF</b>	.86 (.12)	.65 (.20)	1.33 (.85)
<b>CF</b>	.81 (.15)	.55 (.21)	1.33 (.76)
<b>PF</b>	.88 (.11)	.49 (.25)	1.25 (.69)

Acc column in Table 2 and Fig. 3a. Pairwise comparisons revealed a significant difference in testing accuracy (unadjusted for accuracy during the feedback phase) between EF and PF,  $t(18) = 2.19, p < .05$ ; but no significant difference between EF and CF,  $t(18) = 1.24, p > .05$ . The linear regression model provided an estimate of the effect size of EF and CF relative to PF, after taking into account feedback accuracy. This model showed an increase of 0.18 points, or about an 18% increase, in accuracy ( $p < .01$ ) for EF, and an increase of 0.12 points ( $p > .05$ ), or a 12% increase in accuracy for CF, in comparison to

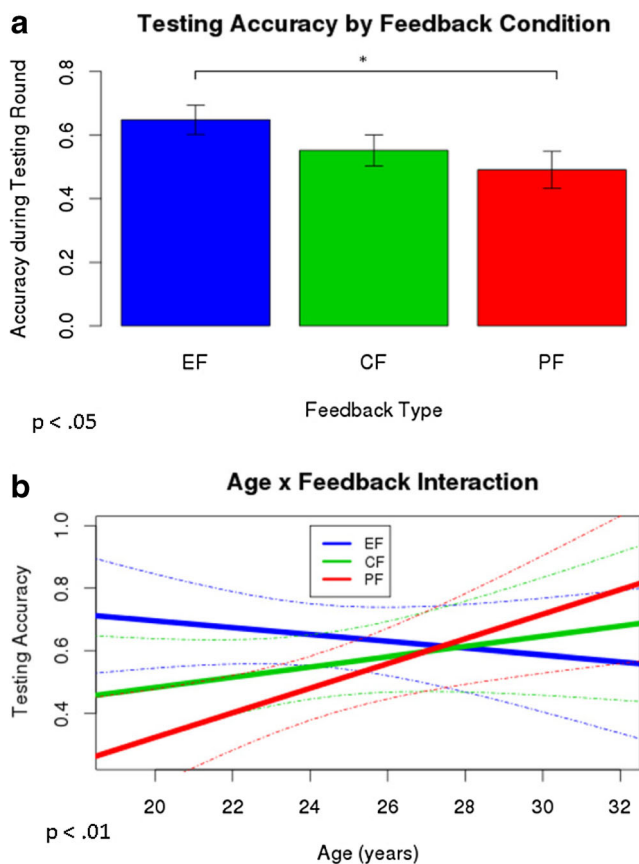
reference variable, PF. As expected, accuracy during the feedback phase, during which the participant was actively learning the pseudoword pronunciations, did not differ by condition:  $F(1, 15) = 2.41, p > .05$  (see the Feedback Acc column in Table 2). Accordingly, pairwise comparisons revealed no significant difference between EF and CF on feedback accuracy,  $t(34) = 1.094, p > .05$ ; EF and PF,  $t(34) = -0.667, p > .05$ ; or EF and PF,  $t(34) = -1.73, p > .05$ .

Because four of the final 19 participants were bilingual or multilingual, we tested for any influence of number of languages spoken on task performance. An ANCOVA analysis, with testing accuracy as the dependent variable, feedback accuracy as the covariate, and the interaction of feedback condition with languages spoken (bilingual or multilingual vs. monolingual) revealed no main effects of languages spoken on accuracy,  $F(1, 16) = 0.55, p > .05$ , nor any interaction effects of languages spoken with feedback condition,  $F(1, 16) = 0.61, p > .05$ .

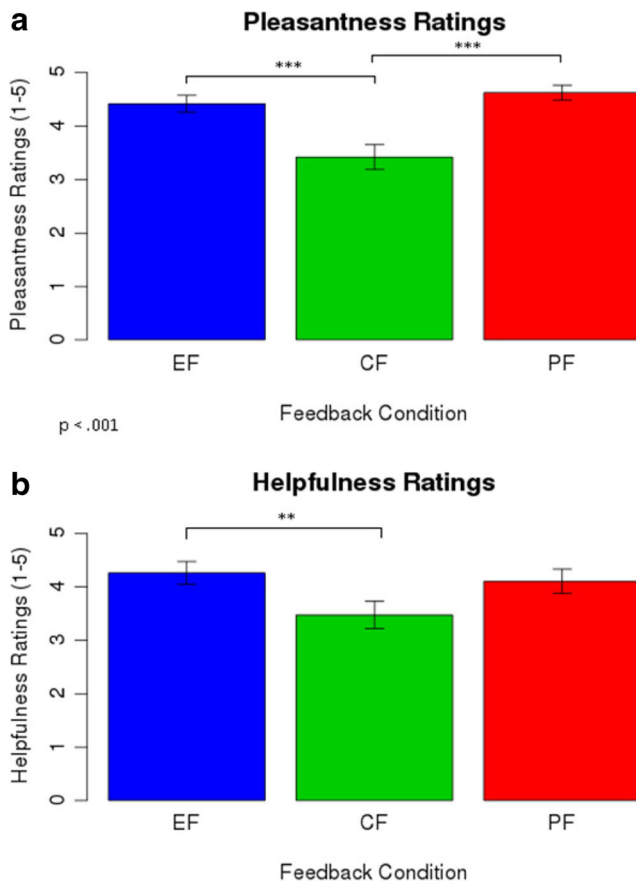
Self-report ratings of reading difficulty experienced as a child and as an adult were collected to ensure that effects of feedback were stable across reading levels, and that there were not any major differences according to self-reported reading levels. An ANCOVA analysis revealed no main effect of self-reported reading difficulty as a child,  $F(1, 16) = 0.347, p > .05$ , or as an adult,  $F(1, 16) = 0.127, p > .05$ , and no significant interaction between feedback condition and reading difficulty as a child,  $F(1, 16) = 0.339, p > .05$ , or adult,  $F(1, 16) = 0.436, p > .05$ .

We also performed an exploratory analysis to determine if there is any initial evidence of greater benefit for EF relative to the other feedback types for participants of decreasing age. To maximize power within the limited age range of our participants, age was treated as a continuous variable. Using an ANCOVA analysis on accuracy data in the testing phase, we tested for the multiplicative interaction of feedback conditions with age. This revealed a reliable interaction, with a greater effect of feedback condition for younger participants,  $F(1, 16) = 6.53, p < .01$  (see Fig. 3b). This interaction held even when the effects of VIQ and the interaction of VIQ and feedback condition were entered into the model,  $F(1, 15) = 6.35, p < .01$ .

To determine whether or not the most effective types of feedback were also experienced as the most pleasant, we performed additional post hoc analyses of subjective ratings provided by participants. Positive feedback was rated as the most pleasant type of feedback on a scale from 1 to 5, with 1 being *very unpleasant* and 5 being *very pleasant*. EF was rated the second most pleasant (EF:  $M = 4.42, SD = 0.69$ ; CF:  $M = 3.42, SD = 1.12$ ; PF:  $M = 4.63, SD = 0.59$ ),  $F(1, 18) = 26.56, p < .001$  (see Fig. 4a). EF and PF were rated as equally pleasant. When considered together with the fact that EF resulted in higher accuracy than PF, this pattern suggests that the effectiveness of EF may be due to pairing of skill-content and



**Fig. 3** **a** Elaborative feedback (EF) resulted in the highest accuracy, compared to positive feedback (PF) and to content feedback (CF); ANCOVA:  $p < .05$ . **b** Interaction of age and feedback condition (elaborative feedback, EF; content feedback, CF; positive feedback, PF). Younger participants showed a stronger effect of feedback condition. Dotted lines represent 95 percentile confidence intervals. (Color figure online)



**Fig. 4** **a** Positive feedback (PF) and elaborative feedback (EF) were rated as more pleasant than content feedback (CF). **b** EF was rated as more helpful than CF. Error bars represent standard error. Two asterisks (\*\*) indicates  $p < .01$ . Three asterisks (\*\*\*) indicates  $p < .001$

reward information. Finally, congruent with effectiveness of feedback, EF was rated as the most helpful on a scale from 1 to 5 (EF:  $M = 4.26$ ,  $SD = 0.93$ ; CF:  $M = 3.47$ ,  $SD = 1.12$ ; PF:  $M = 4.11$ ,  $SD = 0.99$ ),  $F(1, 18) = 3.97$ ,  $p < .05$  (see Fig. 4b).

### Neuroimaging results: Comparisons by feedback type

Analyses of neural activation by feedback type during the feedback phase yielded results that were generally in line with our hypotheses. Directly contrasting EF and CF, EF significantly activated the vmPFC (see Table 3; Fig. 5a), while CF activated a large portion of the bilateral frontal cortex, including the bilateral dlPFC and IFG as well as the bilateral AG and MTG (see Table 3; Fig. 5a). Directly contrasting EF and PF, EF activated the left vOT/FG, left pSTG extending to pMTG, and SMG (see Table 3; Fig. 5b), while PF activated the posterior cingulate cortex (pCing; see Table 3; Fig. 5b). Directly contrasting CF and PF, CF activated the bilateral vOT/FG, left pSTG, SMG, and IFG, while PF activated the vmPFC (see Table 3; Fig. 5c).

Considering that the vmPFC is in an area sensitive to fMRI signal dropout, the mean signal in the vmPFC ROI (the

resulting cluster of the EF > CF contrast) was plotted against the beta values for each participant's EF–CF contrast (see Fig. S2 in the Supplementary Materials). No significant correlation was found between the mean signal and beta value for EF–CF,  $r(19) = -0.14$ ,  $p > .05$ , thus verifying that the effect was not driven by low signal levels in this region.

Distinct from the feedback phase just described, this study was not designed to test neural activation during the testing phase, given the small number of words per condition in that phase. Accordingly, no activation differences were found among feedback conditions during this phase.

### Neuroimaging results: Pseudoword comparisons

The fact that the feedback directly followed pseudoword presentation allowed us to examine how the BOLD signal during pseudoword presentation was modulated by feedback type. We also included trial number as a covariate to see how neural activation over the course of the six learning trials may have differentially changed by feedback type. Note that it is this whole-brain analysis that established the location of significant interaction effects. Follow-up analyses to describe the pattern of the interactions were then performed. Activation in the left IFG was modulated by trial number differently for EF compared to CF (see Fig. 6a) and for EF compared to PF (see Fig. 6b). Since the resulting clusters in the EF–CF and EF–PF analyses covered both the IFG and STG, these regions were plotted separately to visualize the activation pattern in each region. Two ROIs were generated with a radius of 3 mm surrounding a maximum point in the IFG ( $x = -50$ ,  $y = 17$ ,  $z = 2$ ; see Table 4) and a maximum point in the STG ( $x = -45$ ,  $y = -5$ ,  $z = 1$ ; see Table 4). Parameter estimates were plotted for each trial for both the STG and IFG separately for EF versus CF. Since the overall pattern was the same for pseudowords in these two ROIs, and the cluster covered a greater portion of the IFG, only the IFG plot is shown. To better characterize the interaction effect, a linear regression was conducted on the parameter estimates for the IFG sphere, with trial number as the independent variable. This analysis showed no effect of trial number on IFG activation for pseudowords trained with EF,  $F(1, 18) = -1.69$ ,  $\beta = -0.23$ ,  $p > .05$ , but a significant positive effect of trial number on IFG activation for pseudowords trained in the CF condition,  $F(1, 18) = 5.39$ ,  $\beta = 0.77$ ,  $p < .0001$ .

Similarly, in the EF versus PF contrast, one ROI was generated in the left IFG at the local maximum ( $x = -44$ ,  $y = 2$ ,  $z = 11$ ; see Table 4) and another in the left STG ( $x = -56$ ,  $y = -4$ ,  $z = 2$ ; see Table 4). Parameter estimates were plotted separately, and again, the overall pattern was the same for both the left IFG and left STG. Only the IFG parameter estimates are shown in Fig. 6b. A linear regression on the parameter estimates for the IFG ROI revealed no effect of trial number for EF,  $F(1, 18) = -0.36$ ,  $\beta = -0.04$ ,  $p > .05$ , but a significant main

**Table 3** Talairach coordinates and peak *t* statistic for each cluster significantly activated in the main feedback comparisons

Location of extreme point	Cluster size ( $\mu$ l)	<i>x</i>	<i>y</i>	<i>z</i>	Peak <i>t</i> statistic
<i>EF &gt; CF</i>					
Ventromedial prefrontal cortex	1,250	0	31	−3	3.91
<i>CF &gt; EF</i>					
Right lateral frontal cortex	14,993	51	21	27	4.45
R middle frontal gyrus		44	18	41	4.31
R inferior frontal gyrus (triangularis)		43	24	12	4.22
R inferior frontal gyrus (opercularis)		35	2	35	4.01
R inferior frontal gyrus (orbitalis)		45	31	−8	3.93
R inferior frontal gyrus (orbitalis)		33	22	−2	3.95
Dorsal frontal cortex	4,488				
R medial superior frontal gyrus		8	29	42	3.64
L medial superior frontal gyrus		−8	3	56	3.58
L inferior/ middle frontal gyrus	3,993	−38	23	20	4.04
L angular gyrus	2,877	−35	−56	35	4.23
R superior temporal sulcus	2,784	49	−26	−2	4.49
L lateral middle frontal gyrus	2,528	−51	10	47	3.75
Right parietal lobe	2,144				
R angular gyrus		36	−61	41	3.48
R inferior parietal lobule		52	−52	36	3.31
Right fronto-medial cortex	2,012				
R middle frontal gyrus		25	50	11	4.79
R superior frontal gyrus		12	47	30	3.36
L middle temporal gyrus	1,173	−54	−37	4	3.73
L claustrum	946	−29	21	0	4.44
R precuneus	625	11	−52	35	3.95
<i>EF &gt; PF</i>					
Left occipital cortex	2,408				
Left fusiform gyrus (posterior)		−41	−53	−9	4.33
Left fusiform gyrus (anterior)		−38	−39	−17	3.99
Left temporal / parietal cortex	904				
L superior temporal gyrus (posterior)		−43	−43	14	3.47
L supramarginal gyrus		−49	−27	23	3.45
L middle occipital gyrus	624	−27	−85	−1	3.48
<i>PF &gt; EF</i>					
R posterior cingulate	2,171	15	−50	35	3.70
<i>CF &gt; PF</i>					
Left frontal cortex	13,453				
L precentral sulcus		−40	1	33	4.62
L inferior frontal gyrus (triangularis)		−48	19	25	4.28
L inferior frontal gyrus (orbitalis)		−51	16	6	4.21
L precentral gyrus		−47	−1	49	3.60
L inferior frontal sulcus		−29	24	7	3.32
L inferior frontal gyrus (triangularis)		−53	39	12	4.23
Right frontal cortex	9,687				
R precentral sulcus		37	7	29	4.87
R inferior frontal gyrus (triangularis)		40	27	15	4.27
R middle frontal gyrus		51	16	27	3.72
Left temporal cortex	5,227				
L middle temporal gyrus		−48	−45	7	5.30

**Table 3** (continued)

Location of extreme point	Cluster size ( $\mu$ l)	$x$	$y$	$z$	Peak $t$ statistic
L superior temporal gyrus (posterior)		−61	−29	7	4.02
L superior temporal gyrus		−43	−24	2	3.50
L superior temporal gyrus (anterior)		−58	7	−5	3.38
Left occipital cortex	2,303				
L fusiform gyrus		−39	−50	−9	4.07
L fusiform gyrus (posterior)		−36	−65	−11	4.02
R inferior parietal lobule	1,834	36	−51	31	4.70
L middle occipital gyrus	1,324	−28	−84	1	4.19
L superior frontal gyrus (posterior)	1,169	−5	16	51	4.02
R fusiform gyrus	1,106	42	−50	−8	3.94
Right occipital cortex	1,073				
R lingual gyrus		23	−85	−10	3.67
R middle occipital gyrus		34	−75	8	3.53
L subcallosal gyrus	886	−2	12	−11	4.37
L inferior temporal gyrus	735	−41	−35	−19	3.92
PF > CF					
Medial frontal cortex	776				
R mid-orbital gyrus		1	45	3	3.48
L mid-orbital gyrus		0	65	7	3.24
R precentral gyrus	586	19	−20	57	3.57

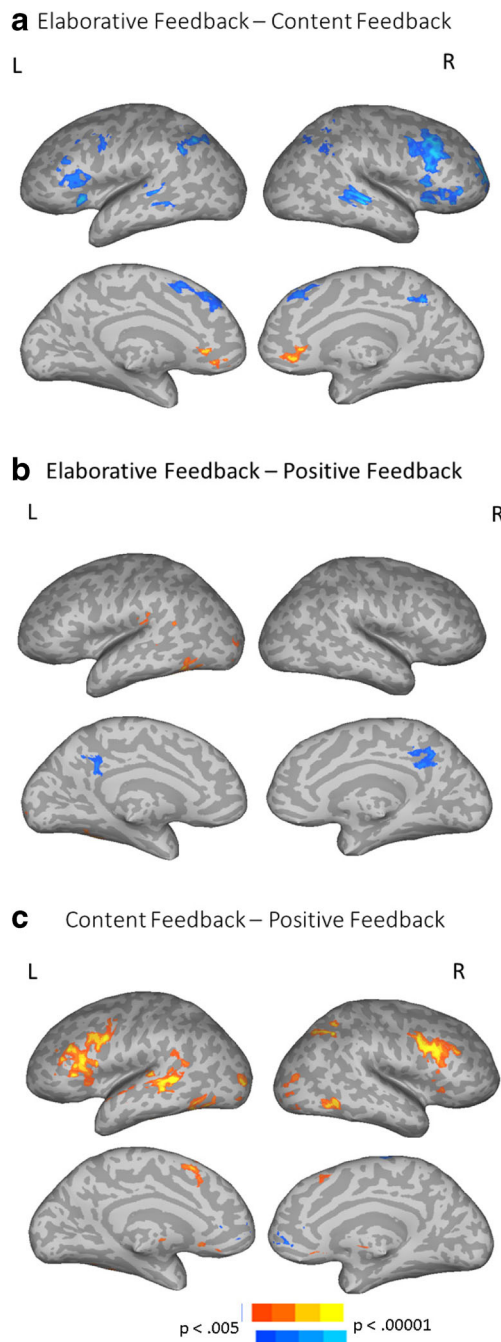
effect of trial number on IFG activation for pseudowords trained in the PF condition,  $F(1, 18) = -4.208$ ,  $\beta = 0.33$ ,  $p < .01$ .

Trial number during the feedback phase also modulated activity in the caudate differently for EF compared to CF (see Fig. 6a) and PF (see Fig. 6b). Since the cluster included a small portion of the cingulate as well, a 3-mm sphere surrounding the maximum points in the caudate for both the EF–CF contrast ( $x = -13$ ,  $y = 15$ ,  $z = 15$ ; see Table 4) and the EF–PF contrast ( $x = -13$ ,  $y = 7$ ,  $z = 20$ ; see Table 4) were created. Again, the pattern was the same in the caudate as in the overall cluster, but only the parameter estimates for the caudate were plotted so as to exclude activation in surrounding brain regions. The overall pattern in both the left and right caudate was similar as well for both the EF–CF and EF–PF contrasts; only the left caudate plots are shown in Fig. 6b. To better characterize the modulation by trial number in the EF–CF and EF–PF contrasts, we conducted a linear regression on the caudate parameter estimates for each condition. For EF compared to CF, trial number significantly negatively predicted neural activation in the caudate ROI,  $F(1, 18) = -5.92$ ,  $\beta = -0.55$ ,  $p < .0001$ ; but there was no significant effect of trial number for CF,  $F(1, 18) = -.47$ ,  $\beta = -0.05$ ,  $p > .05$ . In the EF–PF contrast, trial number significantly negatively predicted neural activation in the caudate ROI,  $F(1, 18) = -4.2$ ,  $\beta = -0.55$ ,  $p < .0001$ ; but there was no effect of trial number for PF,  $F(1, 18) = -0.16$ ,  $\beta = -0.02$ ,  $p > .05$ .

## Discussion

### Interpretation of behavioral results

Feedback that praises the learner for correctly applying the target skill (EF) was predicted to be more helpful than that which only praises the learner (PF) or only provides skill-content information (CF). In line with these predictions, participants learned to pronounce novel letter strings with higher accuracy when the pseudowords were trained with EF than with PF or CF, suggesting that EF promotes learning the orthography-phonology conversion for novel letter strings. This is consistent with previous literature demonstrating feedback to be most effective when it provides specific information about the use of strategies (Brinko, 1993) and more elaborative compared to simple information (Lin et al., 2013). The novel contribution, however, is that this study provides specific evidence for the use of EF to support an essential literacy skill—orthography to phonology conversion. Given that difficulties in orthography-phonology conversion typically underlie reading difficulties (Blomert, 2011; McNorgan, Randazzo-Wagner, & Booth, 2013), improving upon this skill is of high importance for struggling readers. While systematic instruction in orthography-phonology conversion has previously been shown to be effective (Ehri et al., 2001), our study has demonstrated an additional advantage through the use of a particular kind of feedback that explicitly combines



**Fig. 5** **a** Contrasting elaborative feedback (EF) with content feedback (CF), EF (warm colors) activated the vmPFC. CF (cool colors) bilaterally activated the dlPFC, IFG, AG, and MTG. **b** Contrasting EF with positive feedback (PF), EF (warm colors) activated the left SMG, pSTG extending to pMTG, and vOT/FG. PF (cool colors) activated the precuneus/pCing. **c** Contrasting CF with PF, CF (warm colors) activated the left STG, SMG, and bilateral IFG and vOT/FG. PF (cool colors) activated the vmPFC. (Color figure online)

information about orthography-phonology mappings with positive verbal reinforcement for correct answers.

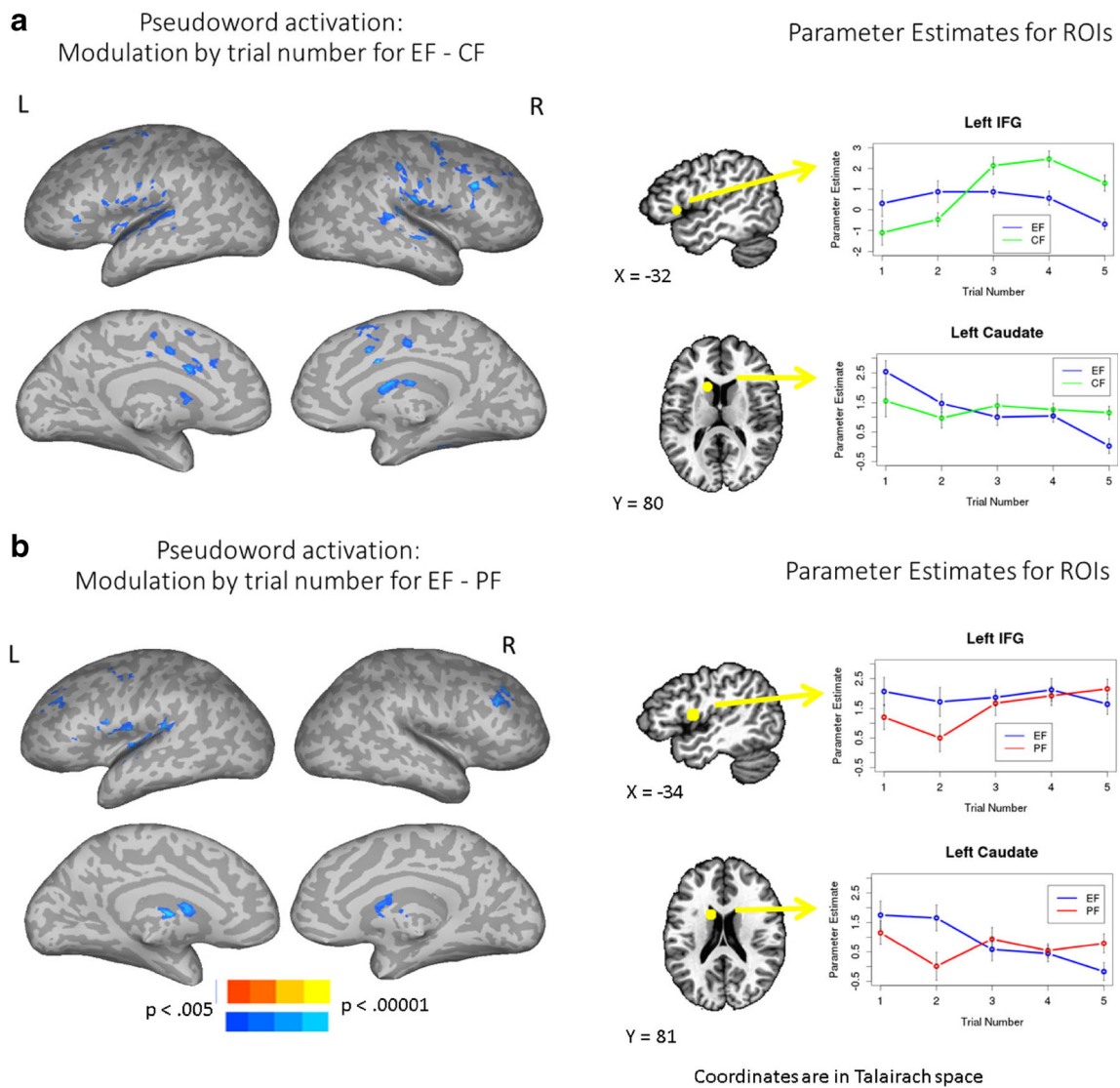
The post-hoc analysis examining the interaction of age with feedback condition shows that in our sample, younger individuals benefited more from EF, which included both

reward and explicit skill content, compared to feedback that simply indicated that their answer was correct (PF). This is consistent with previous research showing that adolescents may be more sensitive than adults to positive social feedback (Cauffman et al., 2010; Jones et al., 2014) and less able to learn from negative feedback (Cauffman et al., 2010). Although, in those two studies, adolescents up to age 17 years were compared to adults ages 18 to 25 years, recent research demonstrates that during emerging adulthood, 18 to 25 years, cognitive and affective patterns may resemble those of adolescence (Arnett & Jensen, 2000). This is relevant to the current study, where participants ranged in age from 19 to 32 years, offering evidence that responsiveness to feedback type may also differ between periods of emerging and later adulthood. While we take caution in interpreting these results due to the limited age range of the participants, we do suggest that our results are promising enough to warrant further testing of positive, specific feedback for younger individuals.

The analysis of subjective ratings suggests that feedback in the form of praise elicits positive affect and a sense of efficacy, yet may not be effective in promoting learning without additional informational (skill) content. EF was rated as more pleasant and helpful than CF, and resulted in the highest mean accuracy score. These results suggest that by combining praise with skill content, EF elicits reward value associated with the target skill, thus improving learning. If future studies demonstrated that these results generalized to struggling readers, this would be particularly important for these individuals, whose adverse affective reactions to their difficulties may further deter their progress (Dweck, 1975; Dweck & Bempechat, 1983). Feedback that is both rewarding and informative may counter these negative affective reactions, thereby potentially improving learning for struggling readers.

### Interpretation of neuroimaging results: Feedback processing

As predicted, EF recruited neural regions implicated in both reward and orthography-phonology content. Compared to CF, EF recruited the vmPFC, implicated in both reward (Blair et al., 2006; for review, see Smith & Delgado, 2015) and self-processing (D'Argembeau, 2013). Given that EF was rated as more rewarding than CF, activation of this region for EF compared to CF is likely due to reward processing. The vmPFC has been implicated in encoding reward value, including expected monetary outcomes (Knutson, Fong, Bennett, Adams, & Hommer, 2003) and social rewards (Somerville, Kelley, & Heatherton, 2010). In one study, participants with low self-esteem showed increased activity in this region for positive compared to negative social feedback (Somerville et al., 2010), suggesting a role for the vmPFC in evaluating the salience of a social reward. Since in EF the learner's actions are being perceived by another, it may be that the



**Fig. 6** **a** Left: Interaction of EF–CF pseudoword presentation during the feedback phase with modulation by trial number. Right: Graphs represent parameter estimates for each trial number for an ROI in the left IFG sphere (shown on the sagittal slice,  $x = -32$ ) and left caudate sphere (shown on the axial slice,  $y = 80$ ). **b** Left: Interaction of EF–PF

pseudoword presentation during the feedback phase with modulation by trial number. Right: Graphs represent parameter estimates for the ROI in the left IFG sphere (shown on the sagittal slice,  $x = -34$ ) and left caudate sphere (shown on the axial slice,  $y = 81$ ). Coordinates are in Talairach space. (Color figure online)

vmPFC is recruited due to this social aspect. However, interpretation of the results of the EF–CF contrast is limited due to the presence of two differences between the EF and CF conditions. EF involves both reward and self-processing, as compared to CF, which involves only task-relevant information. Thus, activation of the vmPFC for EF compared to CF cannot be attributed specifically to reward or self-processing. However, research suggests that activation of vmPFC for self-processing may be due to assigning personal value for self-related concepts (D’Argembeau, 2013). For example, this region was recruited for processing evaluative feedback (e.g., “You’re a great girl!”; Pan, Hu, Li, & Li, 2009). Therefore, even if activation of vmPFC was due to self-processing in the

EF compared to CF condition, it is likely that such recruitment is also reward related.

On the flip side of the same contrast, CF showed greater activation than EF in a number of regions previously shown to be involved in reading, such as the MTG, IFG, and AG (Binder, Medler, Desai, Conant, & Liebenthal, 2005; Fiez, Balota, Raichle, & Petersen, 1999; Graves et al., 2010; Turkeltaub, Eden, Jones, & Zeffiro, 2002), as well as the dlPFC (Barbey, Koenigs, & Grafman, 2013; Owen, McMillan, Laird, & Bullmore, 2005), implicated in working memory. It is somewhat surprising that CF resulted in greater activation of reading network regions, compared to EF, given that skill-content was equated across these two conditions.

**Table 4** Talairach coordinates and peak *t* statistic for each pseudoword contrast significantly modulated by trial number

Location of extreme point	Cluster size ( $\mu$ l)	<i>x</i>	<i>y</i>	<i>z</i>	Peak <i>t</i> statistic
EF pseudowords (modulation by trial number) > CF pseudowords (modulation by trial number)					
No significant activations					
CF pseudowords (modulation by trial number) > EF pseudowords (modulation by trial number)					
Left frontal / temporal / parietal cortex	5,416				
L superior temporal gyrus		−67	−9	6	4.59
L inferior frontal gyrus (triangularis)		−50	17	2	3.89
L superior temporal gyrus		−45	−5	1	3.67
L postcentral gyrus		−53	−7	17	3.59
L superior temporal gyrus		−61	−25	15	3.42
Right lateral parietal / temporal cortex	2,979				
R postcentral gyrus		60	−23	17	4.14
R superior temporal gyrus (anterior)		59	−9	−5	3.67
R superior temporal gyrus (posterior)	2,165	51	−31	8	4.55
R middle frontal gyrus	1,784	43	20	27	4.61
R middle frontal gyrus	1,534				
Precentral sulcus		30	−10	43	3.89
Precentral sulcus		51	9	38	3.23
Right caudate / cingulate cortex	1,499				
R caudate (body)		14	7	21	3.76
R caudate (body)		16	−10	24	3.72
R cingulate gyrus		10	−10	39	3.47
L cingulate gyrus	1,486	−16	−3	46	4.25
Left caudate / cingulate cortex	1,113				
L cingulate gyrus		0	10	29	4.29
L caudate (body)		−13	15	15	3.52
R postcentral gyrus	1,106	54	−23	41	3.75
Left insular / temporal cortex	1,038				
L insula		−37	−15	23	3.97
L superior temporal gyrus (posterior)		−43	−30	8	3.91
R superior frontal gyrus (posterior)	735	9	12	52	3.44
R culmen	625	24	−40	−25	4.12
EF pseudowords (modulation by trial number) > PF pseudowords (modulation by trial number)					
No significant activations					
PF pseudowords (modulation by trial number) > EF pseudowords (modulation by trial number)					
Left frontal / temporal cortex	4,751				
L insula		−44	2	11	4.26
L transverse temporal gyrus		−33	−30	9	4.17
L superior temporal gyrus		−56	−4	2	3.18
R middle frontal gyrus	1,471	43	28	29	4.12
L thalamus	988	−7	−7	16	4.75
Right caudate / lentiform nucleus	879				
R caudate		21	2	19	4.14
R lentiform nucleus		15	8	6	3.27
L caudate	651	−13	7	20	3.99
L superior frontal gyrus	587	−28	33	31	3.64
L middle frontal gyrus	576	−28	0	43	3.61

Such activation may be due to greater attention to the skill-content in CF, when an individual is told what skill should be applied in order to achieve the correct answer. Interestingly, although CF recruited reading network regions to a greater degree than EF, EF resulted in higher accuracy. Thus, it is likely that pairing reward and skill-content is responsible for the effectiveness of EF. Further, activation of the dlPFC for CF is consistent with research showing activation of this region for negative feedback (Zanolie, Teng, Donohue, van Duijvenvoorde, Band, Rombouts & Crone, 2008), especially when feedback is informative for achieving a correct answer on the subsequent trial (Zanolie, van Leijenhorst, Rombouts & Crone, 2008). Although CF was provided on correct and not incorrect trials, its form was similar to negative feedback, in that it provided the learner with information regarding the skill which should be applied in order to achieve the correct answer. Thus, dlPFC activation for CF is consistent with this literature.

Consistent with our predictions, EF compared to PF recruited regions involved in orthography-phonology processing including the fusiform gyrus, MTG, SMG, and STG. While previous studies have shown recruitment of these regions for engaging in orthography-phonology tasks (Binder et al., 2005; Fiez et al., 1999; Graves et al., 2010; Turkeltaub et al., 2002), this is the first study to our knowledge that has demonstrated such activation during the receipt of feedback.

The opposite contrast, PF compared to EF, resulted in activation of the pCing, a region considered part of the putative default-mode network. This is a set of mutually correlated regions typically found to be active during conditions in which participants are not engaged in any particular task (Buckner, Andrews-Hanna, & Schacter, 2008; Gusnard, Raichle, & Raichle, 2001). Given that PF also resulted in lower accuracy than EF, the pCing activation may correspond to the learners being less engaged in the task during the receipt of PF compared to EF.

Lastly, PF was also compared to CF, with the prediction that PF would recruit reward-related regions, and CF would recruit task-related regions. These predictions were also supported. PF recruited the vmPFC, while CF recruited the left MTG, STG, SMG, and vOT/FG. This provides additional confirmation that PF recruited reward-related but not task-related regions, and CF recruited task-related but not reward-related regions, thus validating the results of the EF–PF and EF–CF contrasts.

### Interpretation of neuroimaging results: Pseudowords

Higher accuracy for pseudowords trained with EF, along with a relative lack of activation increase in reading-related neural regions while reading aloud these pseudowords as trial number increased, suggests greater habituation to pseudowords trained with EF. Previous literature has found similar habituation effects for trained items. For example, Graves et al.

(2008) found decreased neural activity in the left pSTG over the course of six trials on which participants repeated auditory pseudowords. In our study, the fact that activation in the left IFG and STG remained steady for pseudowords trained in the EF condition, but increased for the other feedback conditions, combined with the greater accuracy for EF than the other conditions, suggests that these areas may be maintaining information for accurately pronouncing pseudowords. Another possibility is that activation in these areas may reflect increased recruitment of resources in the service of learning to pronounce pseudowords under conditions of less beneficial feedback. This interpretation is consistent with literature finding increased activation of some reading-related regions, including the IFG, under conditions that were more difficult, due to either word characteristics (Graves, Boukrina, Mattheiss, Alexander, & Baillet, 2017; Graves et al., 2010) or domain-general demands (Fox & Raichle, 2007; Fox et al., 2005).

Modulation of the caudate by trial number for EF, but not PF or CF, supports the role of the caudate in feedback processing. However, interestingly, EF and PF were both rated as more pleasant than CF, yet the caudate was modulated by trial number only for EF and not PF, suggesting that recruitment of the caudate for EF may be due to the coupling of rewarding and skill content. Specifically, the caudate has previously been shown to facilitate action–outcome associations (Knutson & Cooper, 2005), and here may have facilitated the binding of reward and reading-related information, thereby leading to higher accuracy in this condition.

Together, behavioral and neuroimaging main results suggest the enhanced effectiveness of EF may be due to concurrent activation of reward-related and reading-relevant regions. Thus, receiving elaborative feedback upon correctly applying a skill may not only promote learning, but may also enhance the reward value of a target skill.

### Limitations

While this study provides evidence of the effectiveness of EF in an orthography-phonology learning task, the question of generalizability of results from adults to children and adolescents remains. The current study has provided proof of concept in the form of evidence that the feedback manipulation is effective among adults. Follow-up analyses also showed EF to be particularly effective in this sample for the younger end (18–23 years) of our adult age range. This suggests it may at least be worth exploring the usefulness of EF with younger developing readers in future studies.

Additionally, since the CF condition provided instructive feedback on spelling-sound content, its status as reinforcing a skill they had just demonstrated may have been a bit ambiguous. Mitigating this concern, however, is the fact that participants did change their answers on subsequent trials when an



“X” was provided to indicate it was incorrect but repeated their answers when CF was provided, demonstrating that they at least understood that CF indicated that their answer was correct.

Regarding the stimuli, although average pitch was matched across feedback types, it is possible that feedback types differed in natural pitch fluctuations due to differences in social and emotional content. We chose not to control pitch fluctuations, as this may have diminished the natural social and emotional cues present in the feedback. While potential differences in pitch fluctuations across feedback types may be a confound in the study, we feel our results are unlikely to be due to differences in basic auditory characteristics of the feedback. Instead, pitch fluctuations have been shown to modulate neural regions in the auditory cortex, while our results demonstrate activation in areas outside this region, including the supramarginal and fusiform gyri. Moreover, if activation in the temporal cortex, for example, were due to pitch intonation differences, one would expect similarity between EF and PF, which were more similar in emotional content (participants rated these two types of feedback as more rewarding than CF). On the contrary, activation in auditory cortex was greater for EF than for PF and greater for CF than for PF. Since EF and CF differed in reward ratings but included the same content (orthography-phonology mapping), the resulting activation in the auditory cortex is likely due to the skill content (present in both EF and CF) rather than reward ratings (different for EF and CF).

Lastly, both CF and EF, but not PF, presented the target pseudoword both visually and auditorily. This was done in order to highlight neural regions implicated in orthography (visual pseudoword presentation) and phonology (auditory pseudoword presentation). The contrast of neural activation during the receipt of feedback for EF–PF and CF–PF were designed such that EF and CF included skill content (pseudoword visual and auditory presentation), but PF did not, and thus was used as a baseline comparison for skill content in these contrasts. However, a limitation remains in that neural activation for EF versus PF and CF versus PF could be due to mere exposure to the word form in the non-PF condition rather than differences due to orthography-phonology mapping, per se. Mitigating this concern, however, is the fact that the neural activation for EF versus PF and CF versus PF recruited the predicted neural regions involved in orthography-phonology mapping rather than solely regions implicated in basic, early visual processing. This limitation is also relevant to the pseudoword comparisons. It is true that for EF, but not PF, the pseudoword was presented on the screen on each trial during the receipt of feedback, and thus one would expect habituation for pseudowords in the EF, but not the PF, condition. However, interestingly, the EF–PF contrast revealed neural activation in much of the same neural regions as the EF–CF contrast, where both EF and CF presented the pseudoword on each trial during the receipt of feedback. Thus, it unlikely that

these results are due to repeated stimulus presentation in the EF and CF but not PF conditions.

## Implications

Elaborative feedback focuses on both skill-related content and praising instances of correct application of the skill. This contrasts with feedback that focuses primarily on skill-content alone. Implementing elaborative feedback requires that educators seek positive behaviors and demonstrations of skill among learners, and subsequently communicate their observations to learners in a positive manner, thus promoting positive affect as well as future applications of the behavior. Concurrent activation of task-relevant and reward regions may, over time, increase an individual's positive affective experience of the target academic skill, thus addressing both cognitive and affective impediments to reading. In terms of neural consequences, activation of task-relevant regions during the receipt of feedback has the potential to enhance function in these regions over time, which is associated with improvements in reading performance.

Since the elaborative feedback used here helped readers who were already skilled, this raises the possibility that it may also help struggling readers. If future studies were to show that elaborative feedback similarly recruited reading-related neural regions among struggling readers, this could be potentially helpful, given that these individuals have been shown to have reduced activation of these brain regions during reading tasks (Maisog et al., 2008; Richlan et al., 2009). Conversely, improving reading skills has been shown to correspond to increasing activation in the same reading-related brain regions (Eden et al., 2004; Temple et al., 2003). Therefore, one implication of this study would seem to be that enhancing activation of these neural regions, as shown here during elaborative feedback, could potentially serve to improve reading skills among struggling readers. The rewarding aspect of elaborative feedback may be particularly beneficial in this process, as struggling readers often experience further detriments due to adverse affective reactions to low performance.

**Acknowledgements** We thank Elizabeth Tricomi and Kent Harber for very helpful discussions and feedback on the neural and psychological bases of feedback.

**Funding** This work was supported by a grant from the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant K99/R00 HD065839) to W.W.G.

## References

- Alexander, M. P., Fischer, R. S., & Friedman, R. (1992). Lesion localization in apractic agraphia. *Archives of Neurology*, 49(3), 246–251. <https://doi.org/10.1001/archneur.1992.00530270060019>

- Amett, J. J., & Jensen, J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, *55*(5), 469–480. <https://doi.org/10.1037/0003-066X.55.5.469>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*. <https://doi.org/10.1016/j.jml.2007.12.005>
- Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205. <https://doi.org/10.1016/j.cortex.2012.05.022>
- Bhattacharya, A., & Ehri, L. C. (2004). Read and spell words. *Journal of Learning Disabilities*, *37*(4), 331–348. <https://doi.org/10.1177/00222194040370040501>
- Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., & Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. *NeuroImage*, *27*(3), 677–693. <https://doi.org/10.1016/j.neuroimage.2005.04.029>
- Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *NeuroImage*, *33*(2), 739–48. <https://doi.org/10.1016/j.neuroimage.2006.06.053>
- Binder, J. R., & Mohr, J. P. (1992). The topography of callosal reading pathways. *Brain*, *115*(6), 1807–1826. <https://doi.org/10.1093/brain/115.6.1807>
- Blair, K., Marsh, A. A., Morton, J., Vythilingam, M., Jones, M., Mondillo, K., ... Blair, J. R. (2006). Choosing the lesser of two evils, the better of two goods: Specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(44), 11379–11386. <https://doi.org/10.1523/JNEUROSCI.1640-06.2006>
- Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *NeuroImage*, *57*(3), 695–703. <https://doi.org/10.1016/j.neuroimage.2010.11.003>
- Brinko, T. (1993). The practice of giving feedback to improve teaching. *The Journal of Higher Education*, *64*(5), 574–593.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D’Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory: An aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*(3), 119–128. <https://doi.org/10.1016/j.bandl.2010.12.001>
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain’s default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*(1), 1–38. <https://doi.org/10.1196/annals.1440.011>
- Caffman, E., Shulman, E. P., Steinberg, L., Claus, E., Banich, M. T., Graham, S., & Woolard, J. (2010). Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. *Developmental Psychology*, *46*(1), 193–207. <https://doi.org/10.1037/a0016128>
- Cohen, J. (1960). Kappa: Coefficient of concordance. *Educational and Psychological Measurement*, *20*(37).
- Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal*, *29*(3), 162–73. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8812068>
- D’Argembeau, A. (2013). *On the role of the ventromedial prefrontal cortex in self-processing: The valuation hypothesis*. *Frontiers in human neuroscience*, *7*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3707083/>
- Damasio, A. R., & Damasio, H. (1983). The anatomic basis of pure alexia. *Neurology*, *33*(12), 1573–83. <https://doi.org/10.1212/WNL.33.12.1573>
- Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–668. <https://doi.org/10.1037/0033-2909.125.6.627>
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. <https://doi.org/10.1016/j.tics.2011.04.003>
- Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *Journal of Personality and Social Psychology*, *31*(4), 674–685.
- Dweck, C. S., & Bempechat, J. (1983). Children’s theories of intelligence: Consequences for learning. In S. Paris, G. Olson, & H. Stevenson (Eds.), *Learning and motivation in the classroom* (pp. 239–256). Hillsdale: Erlbaum.
- Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., ... Flowers, D. L. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron*, *44*(3), 411–22. <https://doi.org/10.1016/j.neuron.2004.10.019>
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel’s meta-analysis. *Review of Educational Research*, *71*(3), 393–447. <https://doi.org/10.3102/00346543071003393>
- Ehri, L. C., Satlow, E., & Gaskins, I. (2009). Grapho-phonemic enrichment strengthens keyword analogy instruction for struggling young readers. *Reading & Writing Quarterly*, *25*(2/3), 162–191. <https://doi.org/10.1080/10573560802683549>
- Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(28), 7900–7905. <https://doi.org/10.1073/pnas.1602413113>
- Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron*, *24*(1), 205–218. [https://doi.org/10.1016/S0896-6273\(00\)80833-8](https://doi.org/10.1016/S0896-6273(00)80833-8)
- Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, *8*(9), 700–711.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(27), 9673–9678. <https://doi.org/10.1073/pnas.0504136102>
- Gabrieli, J. D. E. (2009). Dyslexia: A new synergy between education and cognitive neuroscience. *Science (New York, N.Y.)*, *325*(5938), 280–283. <https://doi.org/10.1126/science.1171999>
- Graves, W. W., Boukrina, O., Mattheiss, S. R., Alexander, E. J., & Baillet, S. (2017). Reversing the standard neural signature of the word-nonword distinction. *Journal of Cognitive Neuroscience*, *29*(1), 79–94. <https://doi.org/10.1111/j.1540-8191.2009.00972.x>
- Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex*, *20*(8), 1799–1815. <https://doi.org/10.1093/cercor/bhp245>
- Graves, W. W., Grabowski, T. J., Mehta, S., & Gordon, J. K. (2007). A neural signature of phonological access: Distinguishing the effects of word frequency from familiarity and length in overt picture naming. *Journal of Cognitive Neuroscience*, *19*(4), 617–631. <https://doi.org/10.1162/jocn.2007.19.4.617>
- Graves, W. W., Grabowski, T. J., Mehta, S., & Gupta, P. (2008). The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. *Journal of Cognitive Neuroscience*, *20*(9), 1698–1710. <https://doi.org/10.1162/jocn.2008.20113>
- Gupta, P., Lipinski, J., Abbs, B., Lin, P.-H., Aktunc, E., Ludden, D., ... Newman, R. (2004). Space aliens and nonwords: Stimuli for

- investigating the learning of novel word-meaning pairs. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc.*, 36(4), 599–603. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15641405>
- Gusnard, D. A., Raichle, M. E., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*, 2(10), 685–694. <https://doi.org/10.1038/35094500>
- Harackiewicz, J. M. (1979). The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, 37(8), 1352–1363. Retrieved from [https://www.researchgate.net/profile/Judith\\_Harackiewicz/publication/232431854\\_The\\_Effects\\_of\\_Reward\\_Contingency\\_and\\_Performance\\_Feedback\\_on\\_Intrinsic\\_Motivation/links/5700687f08aea6b77469b858.pdf](https://www.researchgate.net/profile/Judith_Harackiewicz/publication/232431854_The_Effects_of_Reward_Contingency_and_Performance_Feedback_on_Intrinsic_Motivation/links/5700687f08aea6b77469b858.pdf)
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, 106(3), 491–528.
- Haskell, D. W., Fooman, B. R., & Swank, P. R. (1992). Effects of three orthographic/phonological units on first-grade reading. *Remedial and Special Education*, 13(2), 40–49. <https://doi.org/10.1177/074193259201300207>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and Emotion*, 16(3), 231–247.
- Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Powers, A., Mehta, N., ... Casey, B. J. (2014). Adolescent-specific patterns of behavior and neural activity during social reinforcement learning. *Cognitive, Affective, & Behavioral Neuroscience*, 14(2), 683–697. <https://doi.org/10.3758/s13415-014-0257-z>
- Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology*, 35(3), 835–847. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10380873>
- 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. *Psychology Bulletin*, 119(2), 254–284.
- Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology*, 18(4), 411–417.
- Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *NeuroImage*, 18(2), 263–272. [https://doi.org/10.1016/S1053-8119\(02\)00057-5](https://doi.org/10.1016/S1053-8119(02)00057-5)
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159. <https://doi.org/10.2307/2529310>
- Leff, A. P., Crewes, H., Plant, G. T., Scott, S. K., Kennard, C., & Wise, R. J. S. (2001). The functional anatomy of single-word reading in patients with hemianopic and pure alexia. *Brain*, 124(3), 510–521. <https://doi.org/10.1093/brain/124.3.510>
- Lesch, M. F., & Martin, R. C. (1998). The representation of sublexical orthographic-phonologic correspondences: Evidence from phonological dyslexia. *The Quarterly Journal of Experimental Psychology. A: Human Experimental Psychology*, 51(4), 905–938. <https://doi.org/10.1080/713755790>
- Lin, L., Atkinson, R. K., Christopherson, R. M., Joseph, S. S., & Harrison, C. J. (2013). Animated agents and learning: Does the type of verbal feedback they provide matter? *Computers & Education*, 67, 239–249. <https://doi.org/10.1016/j.compedu.2013.04.017>
- Lyon, G. R., Shaywitz, S. E., Shaywitz, B. A., Sl, B. A., Catts, H., Dickman, E., ... Viall, T. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1–14. <https://doi.org/10.1007/s11881-003-0001-9>
- Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences*, 1145, 237–259. <https://doi.org/10.1196/annals.1416.024>
- Maurer, U., Brem, S., Bucher, K., Kranz, F., Benz, R., Steinhausen, H.-C., & Brandeis, D. (2007). Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain: A Journal of Neurology*, 130(Pt. 12), 3200–3210. <https://doi.org/10.1093/brain/awm193>
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293–299. [https://doi.org/10.1016/S1364-6613\(03\)00134-7](https://doi.org/10.1016/S1364-6613(03)00134-7)
- McNorgan, C., Randazzo-Wagner, M., & Booth, J. R. (2013). Cross-modal integration in the brain is related to phonological awareness only in typical readers, not in those with reading difficulty. *Frontiers in Human Neuroscience*, 7, 388. <https://doi.org/10.3389/fnhum.2013.00388>
- Miller, G. A., Bruner, J. S., & Postman, L. (1954). Familiarity of letter sequences and tachistoscopic identification. *The Journal of General Psychology*, 50(1), 129–139. <https://doi.org/10.1080/00221309.1954.9710109>
- Morais, J. (2003). Levels of phonological representation in skilled reading and in learning to read. *Reading and Writing: An Interdisciplinary Journal*, 16(1/2), 123–151. <https://doi.org/10.1023/a:1021702307703>
- Morgan, P. L., & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? *Exceptional Children*, 73(2), 165–183.
- O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. *Current Opinion in Neurobiology*, 14(6), 769–776. <https://doi.org/10.1016/j.conb.2004.10.016>
- Orsolini, M., Fanari, R., Tosi, V., De Nigris, B., & Carrieri, R. (2006). From phonological recoding to lexical reading: A longitudinal study on reading development in Italian. *Language and Cognitive Processes*, 21(5), 576–607. <https://doi.org/10.1080/01690960500139355>
- Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1), 46–59. <https://doi.org/10.1002/hbm.20131>
- Pan, X., Hu, Y., Li, L., & Li, J. (2009). Evaluative-feedback stimuli selectively activate the self-related brain area: An fMRI study. *Neuroscience Letters*, 465(1), 90–94. <https://doi.org/10.1016/j.neulet.2009.08.070>
- Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1/2), 8–13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62(2), 816–47. <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping*, 30(10), 3299–3308. <https://doi.org/10.1002/hbm.20752>
- Roeltgen, D. P., Sevush, S., & Heilman, K. M. (1983). Pure Gerstmann's syndrome from a focal lesion. *Archives of Neurology*, 40(1), 46–47. <https://doi.org/10.1001/archneur.1983.04050010066019>
- Rumsey, J. M., Horwitz, B., Donohue, B. C., Nace, K., Maisog, J. M., & Andreason, P. (1997). Phonological and orthographic components

- of word recognition: A PET-rCBF study. *Brain: A Journal of Neurology*, 120 (Pt. 5), 739–759. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9183247>
- Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., & Cox, R. W. (2009). A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *NeuroImage*, 44(3), 839–848. <https://doi.org/10.1016/j.neuroimage.2008.09.037>
- Sandak, R., Mencl, W. E., Frost, S. J., Rueckl, J. G., Katz, L., Moore, D. L., ... Pugh, K. R. (2004). The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cognitive, Affective, & Behavioral Neuroscience*, 4(1), 67–88. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15259890>
- Schlaggar, B. L., & McCandliss, B. D. (2007). Development of neural systems for reading. *Annual Review of Neuroscience*, 30, 475–503. <https://doi.org/10.1146/annurev.neuro.28.061604.135645>
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., ... Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics*, 104(6), 1351–1359. <https://doi.org/10.1542/peds.104.6.1351>
- Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry*, 57(11), 1301–1309. <https://doi.org/10.1016/j.biopsych.2005.01.043>
- Smith, D. V., & Delgado, M. R. (2015). Reward processing. In A. W. Toga (Ed.), *Brain mapping: An encyclopedic reference* (Vol. 3). New York: Academic Press.
- Somerville, L. H., Kelley, W. M., & Heatherton, T. F. (2010). Self-esteem modulates medial prefrontal cortical responses to evaluative social feedback. *Cerebral Cortex (New York, N.Y.: 1991)*, 20(12), 3005–3013. <https://doi.org/10.1093/cercor/bhq049>
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic axis of the human brain*. Stuttgart: Thieme.
- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. E. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences*, 100(5), 2860–2865. <https://doi.org/10.1073/pnas.0030098100>
- Torgesen, J. K. (2006). Recent discoveries from research on remedial interventions for children with dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 521–537). New York: Wiley.
- Tressoldi, P. E., Vio, C., & Iozzino, R. (2007). Efficacy of an intervention to improve fluency in children with developmental dyslexia in a regular orthography. *Journal of Learning Disabilities*, 40(3), 203–209. <https://doi.org/10.1177/00222194070400030201>
- Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron*, 41(2), 281–292. [https://doi.org/10.1016/S0896-6273\(03\)00848-1](https://doi.org/10.1016/S0896-6273(03)00848-1)
- Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience*, 18(6), 1029–1043. <https://doi.org/10.1162/jocn.2006.18.6.1029>
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, 16(3), 765–780. <https://doi.org/10.1006/nimg.2002.1131>
- van Duijvenvoord, A. C. K., Zanolie, K., Rombouts, S. A. R. B., Raijmakers, M. E. J., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(38), 9495–9503. <https://doi.org/10.1523/JNEUROSCI.1485-08.2008>
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 45(1), 2–40. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14959801>
- Vellutino, F., Scanlon, D., & Spearing, D. (1995). Semantic and phonological coding in poor and normal readers. *Journal of Experimental Child Psychology*, 59(1), 76–123. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0022096585710041>
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143–156. <https://doi.org/10.1016/j.neuron.2007.05.031>
- Vink, M., Pas, P., Bijleveld, E., Custers, R., & Gladwin, T. E. (2013). Ventral striatum is related to within-subject learning performance. *Neuroscience*, 250, 408–416. <https://doi.org/10.1016/j.neuroscience.2013.07.034>
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, 9(4), 325–329. <https://doi.org/10.1111/1467-9280.00064>
- Wechsler, D. (2001). *Wechsler Test of Adult Reading*. San Antonio: Psychological Corporation.
- Wentink, H. W. M. J., Van Bon, W. H. J., & Schreuder, R. (1997). Training of poor readers' phonological decoding skills: Evidence for syllable-bound processing. *Reading and Writing*, 9(3), 163–192. <https://doi.org/10.1023/A:1007921805360>
- Williams, J. P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology*, 72(1), 1–15.
- Zanolie, K., Teng, S., Donohue, S. E., van Duijvenvoorde, A. C., Band, G. P., Rombouts, S. A., & Crone, E. A. (2008). Switching between colors and shapes on the basis of positive and negative feedback: An fMRI and EEG study on feedback-based learning. *Cortex*, 44(5), 537–547. <https://doi.org/10.1016/j.cortex.2007.11.005>
- Zanolie, K., van Leijenhorst, L., Rombouts, S. A. R. B., & Crone, E. A. (2008). Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia*, 46(1), 117–126. <https://doi.org/10.1016/j.neuropsychologia.2007.08.009>