Bilingualism: Language and Cognition

cambridge.org/bil

Reversal rewards drive language switching during observational learning: Evidence from a dual-brain EEG study

Junjun Huang[1,2], Mengjie Lv[1,2], Yingyi Xiang[1,2], Shuang Liu[1,2], Yujing Shen[1,2], John W. Schwieter[3,4] and Huanhuan Liu[1,2]

[1]Research Center of Brain and Cognitive Neuroscience, Liaoning Normal University, Dalian, China; [2]Key Laboratory of Brain and Cognitive Neuroscience, Dalian, China; [3]Language Acquisition, Multilingualism, and Cognition Laboratory / Bilingualism Matters, Wilfrid Laurier University, Waterloo, Canada and [4]Department of Linguistics and Languages, McMaster University, Hamilton, Canada

Research Article

Cite this article: Huang, J., Lv, M., Xiang, Y., Liu, S., Shen, Y., Schwieter, J.W. and Liu, H. (2025). Reversal rewards drive language switching during observational learning: Evidence from a dual-brain EEG study. *Bilingualism: Language and Cognition* **28**, 1227–1238. https://doi.org/10.1017/S1366728925000070

Received: 7 August 2024
Revised: 22 November 2024
Accepted: 28 December 2024
First published online: 4 February 2025

Keywords:
language-switching; reversal rewards; observational learning; dual EEG; adaptive control hypothesis

Corresponding author:
Huanhuan Liu;
Email: abcde69503@126.com

Abstract

Research on the cognitive neural mechanisms of language control often overlooks the role of rewards. To investigate how reversal rewards affect bilingual language switching during observational learning, we conducted a dual-brain electroencephalography (EEG) study. Participants, classified as direct learners or observers, performed a voluntary language-switching task under dynamic reward conditions. Our results demonstrated that both direct learners and observers exhibited high correct acquisition rates for the switch and non-switch behaviors in both pre- and post-reversal phases. Notably, direct learners and observers showed reduced switch costs in the post-reversal phase, highlighting enhanced language control efficiency. EEG analyses revealed that direct learners exhibited late positive component (LPC) switch costs in both pre- and post-reversal phases, while observers showed LPC switch costs only in the post-reversal phase. These findings support the Adaptive Control Hypothesis by highlighting the adaptability of language control mechanisms in response to dynamic reward environments during direct and observational learning.

Highlights

- This study assesses how reversal rewards affect language control in social contexts.
- Both direct learners and observers showed similar improvement in language control.
- Reversal rewards are crucial for adaptive language control in observational learning.

1. Introduction

In social contexts, bilinguals frequently switch between their languages to meet relevant communicative demands. This process requires inhibiting interference from their non-target language and monitoring potential language conflicts to maintain effective use of the target language, a cognitive ability known as language control (Abutalebi et al., 2008; Crinion et al., 2006; Declerck et al., 2020; Green & Abutalebi, 2013; Liu et al., 2020). Enhanced bilingual language control is crucial for efficient language switching. Previous studies have primarily examined the cognitive neural mechanisms involved in language switching through language control (de Bruin et al., 2018; Liu et al., 2020; Liu et al., 2022), language proficiency (Abutalebi et al., 2013; Luque & Morgan-Short, 2021) and the relationship between language control and cognitive control (Anderson et al., 2018; Iluz-Cohen & Armon-Lotem, 2013). Rewards, as potent motivational factors, play a crucial role in modulating cognitive control (Botvinick & Braver, 2015; Chiew & Braver, 2014; Yee & Braver, 2018). While prior research has demonstrated that rewards enhance cognitive control, the relationship between rewards and language control among bilinguals remains underexplored. This is particularly the case in observational learning, a key aspect of real-world bilingual communication. In the present study, we aim to address this gap by examining the effects of reversal rewards on bilingual language control during observational learning.

1.1. Reversal rewards

Numerous studies have demonstrated that rewards significantly enhance cognitive control by improving attention and boosting motivation (Botvinick & Braver, 2015; Chiew & Braver, 2014; Yee & Braver, 2018). For example, in task-switching contexts, rewards have been shown to improve task-switching abilities as well as optimize behavioral performance by reinforcing effective strategies and behaviors (Fröber & Dreisbach, 2016; Kleinsorge & Rinkenauer, 2012;

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Shen & Chun, 2011). Reversal rewards are frequently used to study how rewards influence cognitive control, involving adaptive changes in reward conditions in which individuals must inhibit previously learned responses and adopt new ones when reward patterns change. In a typical reversal reward task, participants are trained to associate a specific stimulus with a reward. Once this association is established, the conditions are reversed and the participant must learn to associate a different stimulus with the reward to continue obtaining positive outcomes (Bartolo & Averbeck, 2020; Ghahremani et al., 2010; Izquierdo et al., 2017; McAlonan & Brown, 2003). Frings et al. (2020) proposed Binding and Retrieval in Action Control theory, which posits that action control integrates stimuli, responses and effect features into an event archive. Under this assumption, a reward for a stimulus will build binding, and a reversal reward will foster the dynamic nature of action control.

Several studies have advanced our understanding of reversal rewards. For instance, Wang et al. (2023) employed a probabilistic Go/NoGo reversal reward task in which participants learned associations between two tactile stimuli and responses over a series of trials, followed by a reversal reward. Participants were asked to determine the optimal response to each tactile stimulus through several trials to obtain rewards. During the reversal reward task, the first ten trials after the reversal were recognized as reversal naïve (post-reversal phase), while the last ten trials before the reversal or task completion were identified as reversal expert (pre-reversal phase). Analyses of behavioral performance during reversal rewards revealed that participants rapidly adjusted their responses and quickly acquired a high correct rate of switch and non-switch behaviors. Notably, their correct acquisition rate was higher in the pre-reversal phase compared to the post-reversal phase.

Moreover, some electroencephalography (EEG) studies on reversal rewards have underscored the essential role of the P300 component in facilitating cognitive adaptation to changing reward conditions (Donaldson et al., 2016; Von Borries et al., 2013). The P300 is a well-established event-related potential (ERP) component that is critical in processing unexpected but task-relevant information. This component is stimulus-locked and typically peaks between 300 and 700 ms after the presentation of an outcome, primarily at posterior/parietal sites. The P300 reflects the reallocation of attention to salient stimuli, particularly those that are task-relevant, infrequent or unexpected, thereby facilitating the updating of stimulus representations (Donchin & Coles, 1988; Fonken et al., 2020; Glazer et al., 2018; Polich, 2007). Overall, reversal reward tasks reveal the capacity of individuals to adapt to changing rewards, illustrating key mechanisms of cognitive control. This adaptability is essential not only for general cognitive tasks but may also extend itself to specific domains, such as bilingual language control, where switching between languages requires adjustments to contextual demands.

1.2. Language control

Effective language control is crucial for bilinguals to switch between their languages. This process entails activating the target language while suppressing the non-target language to ensure smooth switching and effective communication (Green, 1998; Iluz-Cohen & Armon-Lotem, 2013; Schwieter & Sunderman, 2008). The adaptive control hypothesis (ACH) posits that bilinguals switch between different languages based on contextual demands. During this switching process, various executive functions are engaged to perform the operations necessary for effective language control (Green & Abutalebi, 2013). The voluntary language switching paradigm is widely used to investigate bilingual language control. In this paradigm, participants name pictures or numbers in a series of trials and switch between their languages at will rather than being cued (de Bruin et al., 2018; Gross & Kaushanskaya, 2015; Jiao et al., 2022). Trials in which the language differs from the preceding trial are termed switch trials, while those using the same language consecutively are referred to as non-switch trials. Performance on switch trials is generally less accurate and slower than on non-switch trials, a discrepancy known as the switch cost. This task allows for the examination of how bilinguals manage their languages in naturalistic settings (Heikoop et al., 2016; Paap et al., 2017; Verhoef et al., 2009).

It is well-established that the N2 component is sensitive to language switching and is commonly associated with inhibitory control or conflict monitoring. It is predominantly localized in the prefrontal region of the scalp and peaks approximately 200–350 ms after stimulus onset. For instance, Kang et al. (2020) found that switch trials elicited a more pronounced N2 effect than non-switch trials during the lexical selection phase, indicating increased inhibition of cross-language interference during switch trials. However, this pattern has not consistently been revealed in other work on language switching. Several studies have examined the late positive component (LPC), which typically appears 400–650 ms after stimulus presentation and is predominantly distributed in the posterior scalp region. The LPC is considered an indicator of stimulus–response remapping during language switching, reflecting language control in word selection and retrieval (Jackson et al., 2001; Liu et al., 2016; Martin et al., 2013). For example, Timmer et al. (2017) found that switch trials elicited greater LPC effects than non-switch trials during the word selection phase of language production.

1.3. The present study

This study examines how rewards influence language control in social contexts through a dual-brain EEG approach. Specifically, it addresses gaps in the existing literature by introducing a novel paradigm that integrates reversal rewards with voluntary language switching in an observational learning context. This approach permits an examination of both the role of rewards in bilingual language control and the impact of social learning on these processes. We implemented an observational learning setup where participants voluntarily chose to switch or not to switch languages in order to gain a high reward and avoid a low reward. This setup simulated a social situation where individuals observed others' behavior and adapted their own decisions accordingly (Goubert et al., 2011; Rautio et al., 2024). In our study, participants were paired with a companion who observed their actions and rewards while making similar choices from a set of pictures. We classify individuals who adapted their language-switching behavior based on their own reward outcomes as direct learners. Individuals who adapted their behavior based on the reward outcomes observed from their companions were classified as observers.

To examine the dynamics of language-switching behaviors throughout reversal rewards, we divided participants' behavioral performance into pre-reversal and post-reversal phases based on previous studies (Ghahremani et al., 2010; Wang et al., 2023). To determine whether direct learners and observers acquired and adjusted behaviors following reversal rewards, we calculated the correct acquisition rates for direct learners and observers. We hypothesize that direct learners will dynamically adjust their

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

language-switching behaviors during the two reversal phases. Additionally, we anticipate that observers, influenced by observational reversal rewards, will exhibit behaviors similar to those of direct learners. Next, to examine the presence of switch costs between direct learners and observers in the pre- and post-reversal phases, we compared their reaction times (RTs). Consistent with the ACH, we predict that both direct learners and observers will exhibit significant switch costs across both phases. These switch costs reflect the dynamic adjustments that are required to adapt to changing reward conditions, illustrating how language control mechanisms respond flexibly to contextual demands. Finally, to elucidate the mechanisms underlying the effects of reversal rewards on language switching in direct learners and observers, we analyzed their brain responses to reversal rewards and compared two key language control components, namely the N2 and LPC. We hypothesize that observers will exhibit language-switching behaviors similar to direct learners, as indicated by N2 and LPC effects, though the patterns of language control may differ. By introducing different reward conditions, we seek to provide empirical insight into bilingual communication in social/interactive scenarios to improve experimental validity in bilingual research and to offer scientific guidance for language teaching and learning.

2. Method

2.1. Participants

The sample size was calculated as 28 using G*Power 3.1.9.7 (Faul et al., 2007) with the following parameters: *F*-tests > ANOVA: Repeated measures, within factors, Effect size $f = .27$, $\alpha = .05$, Power (1-β error probability) = .8, Number of groups = 1, Number of measurements = 2, Correlation among measures = .5 and Non-sphericity correct $\in = 1$. To ensure sufficient power, 64 unbalanced bilinguals from Liaoning Normal University were recruited and randomly assigned into 32 same-gender pairs. In each pair, participants were then randomly assigned to be direct learners (Participant A) who received direct rewards or observers (Participant B) who observed rewards. Participants were all right-handed, had normal vision, and had no language or cognitive impairments. The experiment was approved by the research ethics committee at the same university, and participants provided their written informed consent before participating.

Due to data quality issues, we excluded some data, resulting in a final sample of 29 pairs of participants (48 females, 10 males; M_{age} = 23 years and SD_{age} = 2 years). All participants were native Chinese speakers who began learning English between the ages of 5–13 years. To gather information about participants' proficiency in their first language (L1) and second language (L2), we asked them to rate their reading, writing, speaking and listening abilities in both languages on a six-point scale, with "6″ indicating very fluent and "1″ indicating no fluency. Analyses using paired-samples *t*-tests revealed that participants scored significantly higher in their L1 proficiency than in their L2 proficiency across the four language skills. Additionally, to assess their objective proficiency in English, participants completed the Oxford Placement Test (OPT), a validated placement tool published by Oxford University Press (Allan, 2004). Due to time constraints, an abbreviated version of the OPT was administered, including 25 multiple-choice questions and 25 cloze test items. The maximum score was 100 and participants achieved an average of 52.41 ± 10.57 (see Supplementary Table S1). These results align with previous studies of Chinese–English unbalanced bilinguals with intermediate L2 proficiency (Liu et al., 2016). The self-ratings and OPT were both completed prior to the formal experiment.

2.2. Materials

The experimental stimuli consisted of a total of 66 black-and-white line drawings, with 6 used for practice and 60 in the formal experiment. These images were selected from a standardized set of pictures by Snodgrass and Vanderwart (1980). All depicted items were neutral nouns and contained no Chinese–English cognates or cross-language homographs. We referred to the Chinese Lexical Database (Cai & Brysbaert, 2010) and the English Lexical Database (Brysbaert & New, 2009) for word frequency information. Paired-sample *t*-tests revealed no significant difference in word frequency between the L1 and L2 picture names, despite considerable standard deviations (L1: $M = 86.93$, $SD = 152.23$; L2: $M = 102.23$, $SD = 145.59$; $t_{(59)} = -1.18$, $p = .24$).

2.3. Design and procedure

Throughout the entire experiment, each pair of participants was seated in front of the same computer screen. To prevent visual interference between the participants, an opaque foam board (length × width: 1 meter × .5 meters) divided the screen into two equal sections. The left side of the screen was visible only to direct learners (Participant A), while the right side was visible only to observers (Participant B). Two EEG recording systems were used to ensure synchronous data collection, with pulses sent simultaneously to each EEG device via a controller. Two identical ANT amplifiers, each with independent grounding, were connected to two separate computers and operated using the same software interface to ensure synchronization between the two sets of electrodes. RTs for picture naming were recorded using two PSTSR devices connected to microphones. These measures were implemented to maintain consistency in the experimental environment and to ensure the accuracy and reliability of data collection.

Before beginning the experiment, participants familiarized themselves with the Chinese and English names of the pictures on the computer. Direct learners (Participant A) were informed that they could freely choose the language (Chinese or English) in which to name the pictures during each task. Notably, each block of the experiment involved 12 reward rule reversals, with six providing high rewards for switch behaviors and the remaining six for non-switch behaviors. Direct learners were asked to choose between switch and non-switch behaviors to maximize rewards based on previous reward feedback. As illustrated in Figure 1a, high rewards were initially provided for non-switch behaviors. After a reward reversal, high rewards were subsequently given only for switch behaviors. To avoid predictability, reward reversals occurred randomly every four to six trials. This range was chosen to challenge adaptability without overwhelming cognitive load, ensuring the task remained feasible. For the analyses, the first two trials after a reward reversal were defined as the post-reversal phase (naïve phase), and the remaining trials were defined as the pre-reversal phase (expert phase).

Before commencing the picture naming task, observers (Participant B) were instructed to listen to the direct learner's (Participant A) naming responses and to observe their corresponding reward feedback. Participant B's task was to infer which behaviors (switch or non-switch) were associated with high rewards.

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

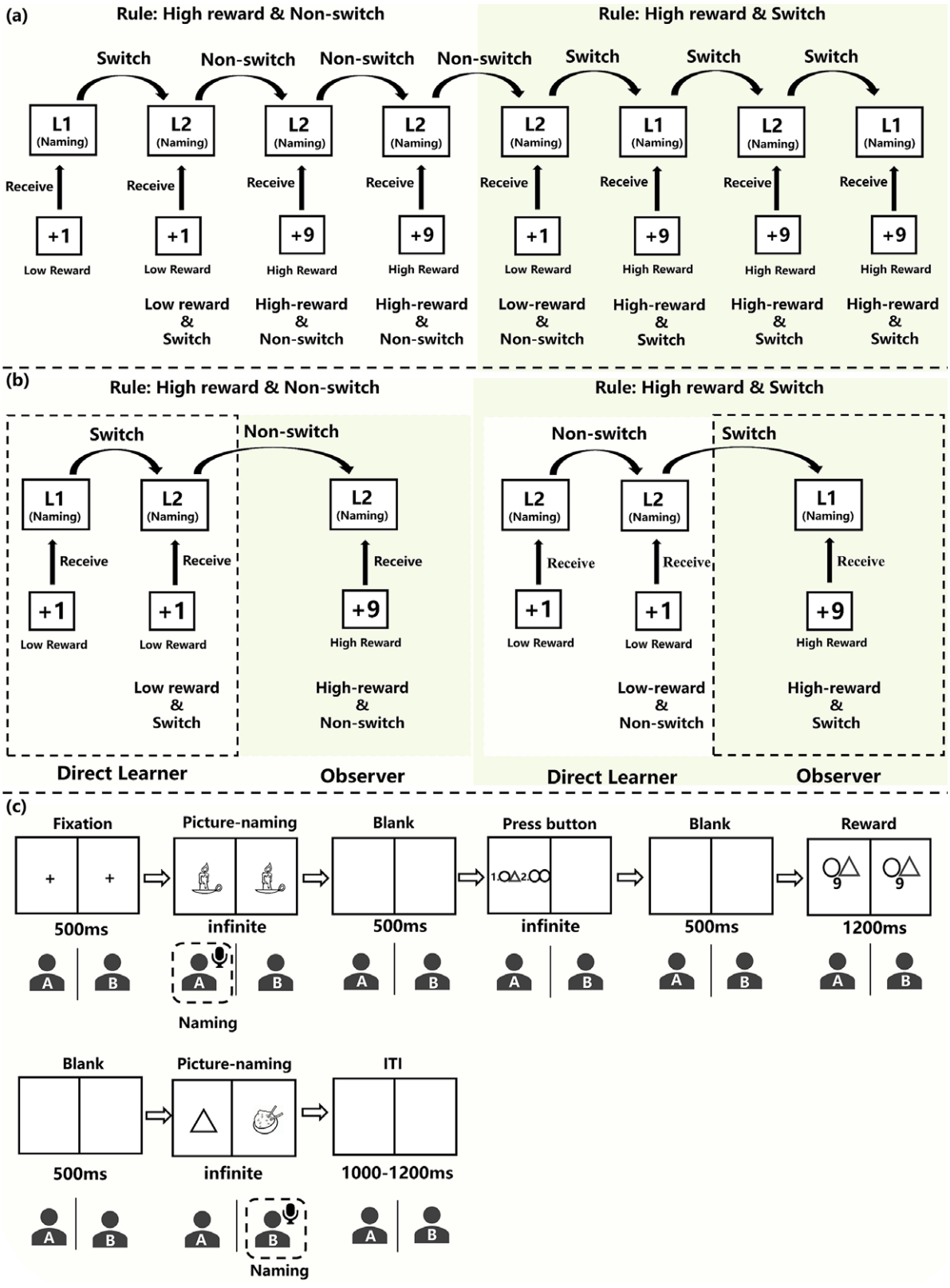


Figure 1. Illustration of the experimental design.
Notes. (a, b) denotes the settings for direct learners and observers regarding rewards and language-switching behaviors, respectively. (c) shows the flowchart of the experiment.

After each observation, observers were required to name a new picture aloud using the reward information they gathered to guide their own responses. As depicted in Figure 1b, through observation, the observer initially noted the high rewards associated with non-switch behaviors but subsequently experienced a sudden reversal, where switch behaviors became necessary to obtain high rewards. The division of trials in the pre-reversal and post-reversal phases for observers was identical to that of direct learners. Moreover,

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

observers were informed that their reward feedback would not be displayed on-screen but would be recorded by the experimenter. Both direct learners and observers were encouraged to respond quickly and accurately to maximize their rewards. Participants received feedback of either 9 points for high rewards or 1 point for low rewards, but their actual compensation for participating in the study was a fixed amount of 80 yuan for each participant.

All participants completed 14 practice trials to familiarize themselves with the procedure. The experiment consisted of 4 blocks, each with 1 practice trial and 60 formal trials. Participants' responses were categorized as switch or non-switch trials based on whether the language of the current trial differed from the previous trial (direct learners) or observed language (observers). Specifically, for direct learners, a trial was considered a switch if they used a different language than in the previous trial; otherwise, it was considered a non-switch trial (see Figure 1a). For observers, a trial was considered a switch if their current language use differed from what they observed from the direct learner; otherwise, it was a non-switch trial (see Figure 1b). Regardless of whether bilinguals switched from L1 to L2 or from L2 to L1, the current trial was considered a switch trial. This design is based on bilingual research indicating that multiple languages share a common cognitive neural system rather than relying on separate representations for each language (Blanco-Elorrieta & Pylkkänen, 2016, 2017). Therefore, differences between the two languages were not considered in subsequent data analyses.

As shown in Figure 1c, each trial began with a central fixation point (+) presented simultaneously on both sides of the screen for 500 ms. Subsequently, a picture appeared for naming, with direct learners freely choosing to name the picture in either their L1 or L2 aloud with no time constraints. This was followed by a 500 ms blank screen. Direct learners indicated whether their naming behavior differed from their previous trial by pressing the 1 or 2 key, which represented switch or non-switch behaviors. These key assignments were counterbalanced across participants. Notably, switch and non-switch behaviors were represented by two combinations of patterns: a circle and a triangle or two circles. These associations were also counterbalanced across participants. After a 500 ms blank screen, the reward feedback for the direct learner's picture naming and key pattern were displayed for both participants. Feedback of nine points indicated a high reward, whereas 1 point indicated a low reward. Finally, observers were required to name a new picture aloud and a random blank screen of 1000–1200 ms was presented before the next trial.

2.4. Behavioral data analyses

Due to the high accuracy of the picture naming task (> 98%), we did not analyze error rates. In the analysis of RTs, we excluded the following data: pictures named incorrectly, responses inconsistent with the reward condition, key press errors, the first practice trial of each block, and RTs exceeding 2.5 SDs from the mean (direct learners: 4.56%; observers: 4.75%). We first calculated the correct acquisition rates for direct learners and observers in the learning phase (pre-reversal and post-reversal) and sequence type (switch trials and non-switch trials). Next, we built a generalized linear mixed-effects model using R software (version 4.3) to calculate whether there was a difference between the correct acquisition rates of direct learners and observers under the two variables separately. Learning phase (pre-reversal and post-reversal) and sequence type (switch trials and non-switch trials) were treated as fixed effects, and participants were treated as random effects. Finally, to investigate the impact of reversal rewards on direct learners' and observers' language switching costs, we constructed generalized linear mixed-effects models for RTs. Learning phase (pre-reversal and post-reversal) and sequence type (switch trials and non-switch trials) were considered fixed effects, and participants were regarded as random effects. We ultimately removed random slopes that led to convergence and overfitting. Additionally, paired-sample *t*-tests were used to compare the language switching costs (i.e., non-switch trials subtracted from switch trials) before and after reversal for both direct learners and observers.

2.5. EEG acquisition and preprocessing

Using the 64-channel electrode cap extended from the international 10–20 system, EEG data were recorded. All electrodes were referenced online to CPz and then re-referenced offline to the average of the left and right mastoids, with impedance kept below 5 kΩ. The sampling rate of the EEG recording was 1000 Hz, which was later reduced to 500 Hz during offline processing. EEG activity was filtered online within the range of .1 to 100 Hz, followed by offline high-pass filtering at .1 Hz and low-pass filtering at 30 Hz. Data with excessive eye movements or other motion artifacts were visually inspected and manually removed. Independent component analysis (ICA) in EEGLAB was utilized to remove ocular artifacts. Each trial was segmented into epochs from 200 ms before stimulus onset to 800 ms after stimulus onset. Additionally, EEG data exceeding ±80 μV were automatically removed.

2.6. ERP analyses

The ERP analyses primarily focused on two phases: reward feedback and picture naming. During the reward feedback-locked phase, we analyzed the P300 component. Consistent with previous literature (Von Borries et al., 2013; Yeung et al., 2005), our analyses targeted five electrode sites: Fz, FCz, Cz, CPz and Pz. We defined the total time window of interest as 200–800 ms post-stimulus. Within each time window, we used linear mixed-effects models with reversal reward as the fixed effect and participants as the random effect to compare neural differences between the first reward trial after reversal and the second reward trial after reversal.

During the picture naming-locked phase, we focused on two-time windows: N2 (Liu et al., 2016; Liu et al., 2020; Misra et al., 2012) and LPC (Jackson et al., 2001; Liu et al., 2016; Liu et al., 2018; Timmer et al., 2019). The N2 was defined as 250–350 ms and LPC as 370–600 ms. Consistent with prior research (Declerck et al., 2020; Liu et al., 2018; Liu et al., 2023), the electrodes of interest for the N2 included: F3, F1, Fz, F2, F4, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2 and C4. Regarding the LPC, the electrodes of interest were: F3, F1, Fz, F2, F4, FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4, P3, P1, Pz, P2 and P4. Amplitude maps were generated based on the average values from these 25 electrodes. To interpret the dynamics of language-switching behaviors throughout the reversal rewards, we classified participants' behavioral performance into two phases, pre-reversal and post-reversal, based on previous work (Ghahremani et al., 2010; Wang et al., 2023). The first two trials after reward reversal were characterized as the post-reversal phase, while the remaining trials were described as the pre-reversal phase. Notably, the EEG data analyses for both direct learners and observers in the picture naming phase were based on their correct acquisition rate. For each time window analysis, we employed linear mixed-effects models

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

with learning phase and sequence type as fixed effects and participants as a random effect.

3. Results

3.1. Behavioral results

3.1.1. Correct acquisition rate

We investigated whether direct learners and observers, when confronted with reversal rewards, promptly adjusted their behaviors to optimize rewards. We refer to this behavior as the correct acquisition rate. Among direct learners (see Figure 2a), our analysis revealed a correct acquisition rate of 89.51% in the post-reversal switch condition (1246 trials, M = 42.97 ± 4.52), 88.03% in the pre-reversal switch condition (1838 trials, M = 63.38 ± 5.59), 89.22% in the post-reversal non-switch condition (1242 trials, M = 42.83 ± 4.56) and 89.08% in the pre-reversal non-switch condition (1860 trials, M = 64.14 ± 5.13). Observers exhibited a correct acquisition rate of 90.37% (1258 trials, M = 43.38 ± 4.45) in the post-reversal switch condition, 89.61% (1871 trials, M = 64.52 ± 5.53) in the pre-reversal switch condition, 86.28% in the post-reversal non-switch condition (1201 trials, M = 41.41 ± 4.99) and 93.20% in the pre-reversal non-switch condition (1946 trials, M = 67.10 ± 3.47) (see Figure 2b).

Furthermore, we performed a linear mixed-effects model to analyze correct acquisition rates across learning phases and sequence types for direct learners and observers. The results revealed a significant interaction effect only for observers (b = .077, SE = .024, t = 3.22, p = .0018) (see Figure 2b). Specifically, non-switch trials (M = .93 ± .015) showed a higher acquisition rate than switch trials (M = .89 ± .015) during the pre-reversal phase (b = − .036, SE = .017, t = −2.13, p = .036), while switch trials (M = .90 ± .015) exhibited a higher acquisition rate compared to non-switch trials (M = .86 ± .015) in the post-reversal phase (b = .041, SE = .017, t = 2.43, p = .017). Additionally, the correct acquisition rate for non-switch trials significantly decreased from the pre-reversal (M = .93 ± .015) to the post-reversal phase (M = .86 ± .015) (b = − .069, SE = .017, t = −4.10, p < .001).

3.1.2. Effects of reversal rewards on language switching in direct learners

The linear mixed-effects model analysis, which incorporated the learning phase (post-reversal and pre-reversal) and sequence type (switch and non-switch), revealed significant main effects for both variables (Sequence: b = −.07, SE = .010, t = −6.49, p < .001; Learning phase: b = .13, SE = .019, t = 7.09, p < .001). Notably, RTs were faster for non-switch trials (M = 849.41 ± 272.33 ms) compared to switch trials (M = 928.61 ± 304.63 ms) and were faster in the post-reversal phase (M = 874.33 ± 274.80 ms) than in the pre-reversal phase (M = 898.16 ± 270.30 ms).

The interaction between the learning phase and sequence type was statistically significant (b = −.072, SE = .012, t = −6.03, p < .001) (see Figure 3a). A simple effect analysis revealed that, regardless of pre or post-reversal, non-switch trials exhibited faster RTs compared to switch trials (post-reversal switch trials: M = 890.75 ± 293.27 ms > post-reversal non-switch trials: M = 857.87 ± 253.99 ms, b = .031, SE = .013, t = 2.48, p = .016; pre-reversal switch trials: M = 954.35 ± 309.57 ms > pre-reversal non-switch trials: M = 843.87 ± 212.37 ms, b = .10, SE = .012, t = 9.09, p < .001), indicating significant switch costs both before and after reversal reward. For switch trials, RTs were faster in the post-reversal phase (M = 890.75 ± 293.27 ms) compared to the pre-reversal phase (M = 954.35 ± 309.57 ms) (b = −.063, SE = .008, t = −7.33, p < .001). However, for non-switch trials, there was no significant difference in RTs between the pre-reversal phase (M = 843.87 ± 212.37 ms) and the post-reversal phase (M = 857.87 ± 253.99 ms) (b = .01, SE = .008, t = 1.18, p = .24), suggesting a more pronounced impact of reversal rewards on switch trials compared to non-switch trials.

3.1.3. Effects of reversal rewards on language switching in observers

Using a linear mixed-effects model which analyzed the learning phase (pre-reversal and post-reversal) and sequence type (switch and non-switch), we identified significant main effects for both variables (Sequence: b = −.027, SE = .013, t = −2.07, p = .048; Learning phase: b = .12, SE = .020, t = 6.08, p < .001). Notably, observers exhibited faster RTs for non-switch trials (M = 838.47 ± 267.12 ms) compared to switch trials (M = 881.91 ± 317.37 ms) and the pre-reversal phase (M = 849.18 ± 283.13 ms) elicited faster responses than the post-reversal phase (M = 877.13 ± 309.45 ms).

The interaction between the learning phase and sequence type was significant (b = −.10, SE = .013, t = −7.77, p < .001) (see Figure 3b). Subsequent simple effect analyses found that non-switch trials (M = 807.48 ± 223.58 ms) had faster RTs than switch trials (M = 892.98 ± 328.86 ms) (b = .077, SE = .014, t = 5.57,

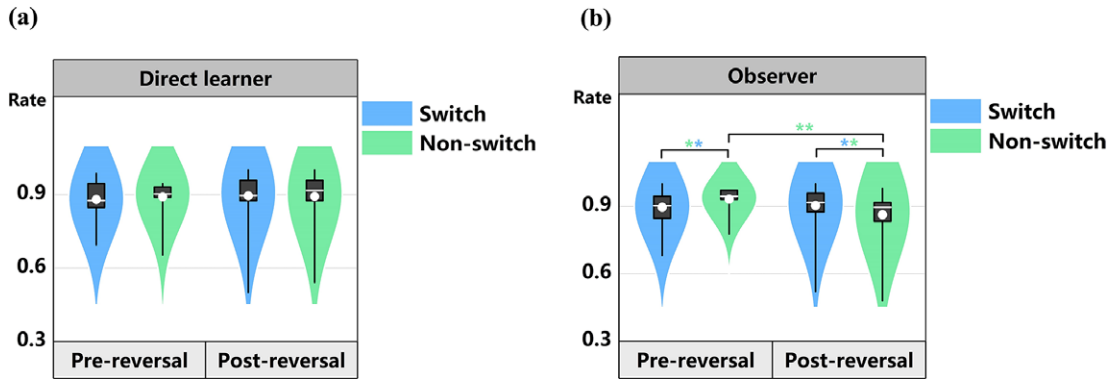


Figure 2. Correct acquisition rates of direct learners (a) and observers (b) in the learning phase and sequence.
Notes. White circles and lines represent the mean and median of participants' responses. Boxplots illustrate the interquartile range (25% and 75%). Colored asterisks (**) indicate significant differences (p < .05) between the two conditions.

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

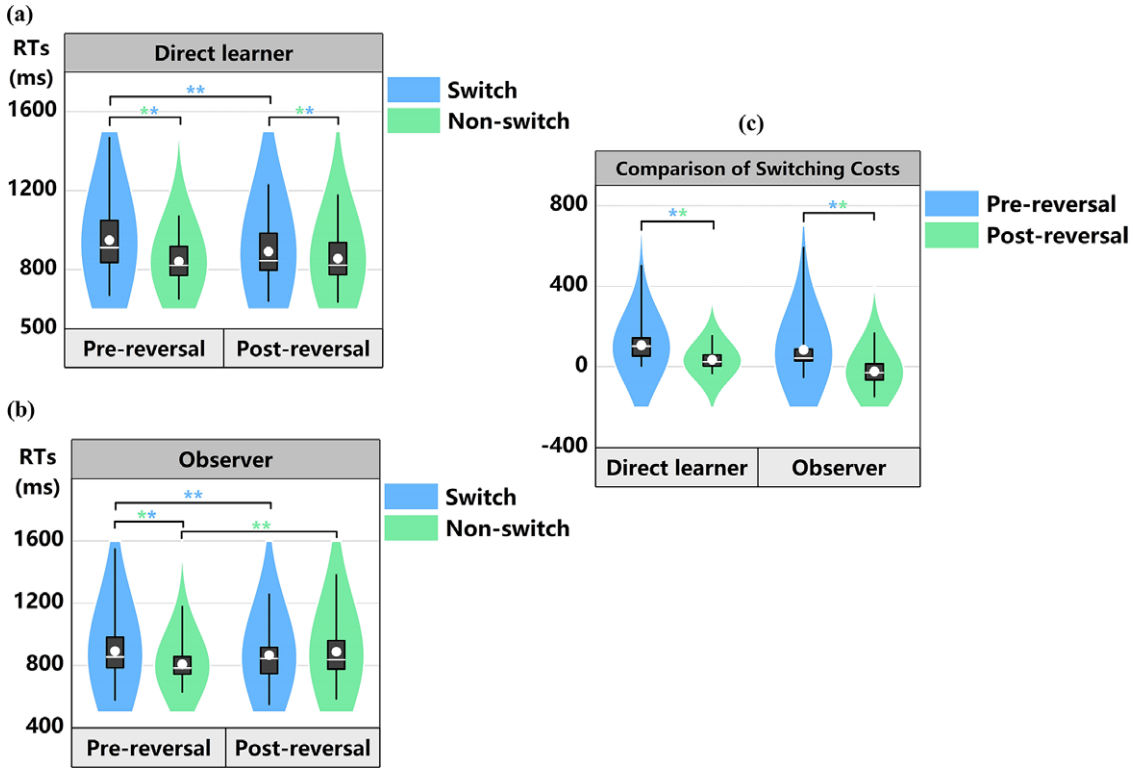


Figure 3. RTs of direct learners (a) and observers (b) in the learning phase and sequence, and switching cost comparisons (c).
Notes. White circles and lines represent the mean and median of participants' responses. Boxplots illustrate the interquartile range (25% and 75%). Colored asterisks (**) indicate significant differences ($p < .05$) between the two conditions.

$p < .001$) during the pre-reversal phase. However, in the post-reversal phase, there was no significant difference in RTs between non-switch trials (M = 889.58 ± 320.10 ms) and switch trials (M = 865.08 ± 298.42 ms) (b = −.023, SE = .015, t = −1.54, p = .13), indicating the presence of switch costs in the pre-reversal phase but not in the post-reversal phase. However, RTs were faster in the post-reversal phase (M = 865.08 ± 298.42 ms) than in the pre-reversal phase (M = 892.98 ± 328.86 ms) (b = −.024, SE = .009, t = −2.62, p = .008) for switch trials and RTs were faster in the pre-reversal phase (M = 807.48 ± 223.58 ms) than in the post-reversal phase (M = 889.58 ± 320.10 ms) (b = .076, SE = .009, t = 8.36, $p < .001$) for non-switch trials, suggesting a greater impact of reversal rewards on switch trials compared to non-switch trials.

3.1.4. Comparison of language switching costs between direct learners and observers

To explore whether language switch costs differed between direct learners and observers and whether they were influenced by reversal rewards, we conducted paired-samples t-tests to compare the differences between pre- and post-reversal RTs. The results, illustrated in Figure 3c, revealed that both direct learners (pre-reversal: M = 108.02 ± 17.44 ms, post-reversal: M = 35.68 ± 9.12 ms) and observers (pre-reversal: M = 84.74 ± 21.80 ms, post-reversal: M = 35.68 ± 13.12 ms) exhibited diminished switch costs in the post-reversal condition compared to the pre-reversal condition (direct learners: $M_{A\text{-}B}$ = −72.34 ms, SE = 14.97, t = −4.82, $p < .001$; observers: $M_{A\text{-}B}$ = −106.82 ms, SE = 25.75, t = −4.15, $p < .001$).

3.2. ERP Results

3.2.1. Results of reversal reward feedback

Using mixed-effects models, we compared reward trials at the feedback phase, the first reward trial after reversal with the second reward trial after reversal. As depicted in Figure 4a, the first reward trial following reversal elicited a larger P300 amplitude in direct learners compared to the second reward trial post-reversal, indicating the influence of reversal reward effects on direct learners ($M_1^{st}{}_{reward\ post\ -\ reversal}$ = 5.55 ± .62 μV, $M_2^{nd}{}_{reward\ post\ -\ reversal}$ = 1.52 ± .62 μV, b = −4.02, SE = .39, t = −10.13, $p < .001$). Similarly, for observers (Figure 4b), the P300 amplitude was greater during the first reward trial following reversal compared to the second reward trial after reversal, suggesting that observational learning was also influenced by the effects of reversal rewards ($M_1^{st}{}_{reward\ post\ -\ reversal}$ = 1.23 ± .62 μV, $M_2^{nd}{}_{reward\ post\ -\ reversal}$ = .57 ± .62 μV, b = −.66, SE = .33, t = −1.99, p = .047).

3.2.2. ERP results of the effects of reversal rewards on language switching in direct learners

A linear mixed-effects model analysis was performed to investigate the impact of the learning phase (post-reversal and pre-reversal) and sequence type (switch and non-switch) on both N2 and LPC. No significant effects were found for N2. In contrast, LPC showed a significant interaction (b = −1.44, SE = 0.50, t = −2.86, p = .0043) (see Figure 5a). Specifically, both post-reversal and pre-reversal phases exhibited switch costs (post-reversal switch: M = 7.55 ± .94 μV > post-reversal non-switch: M = 6.76 ± .94 μV, b = .79, SE = .39, t = 2.02, p = .044; pre-reversal non-switch: M = 7.25 ± .92 μV > pre-reversal switch: M = 6.60 ± .92 μV, b = −.65, SE = .32,

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

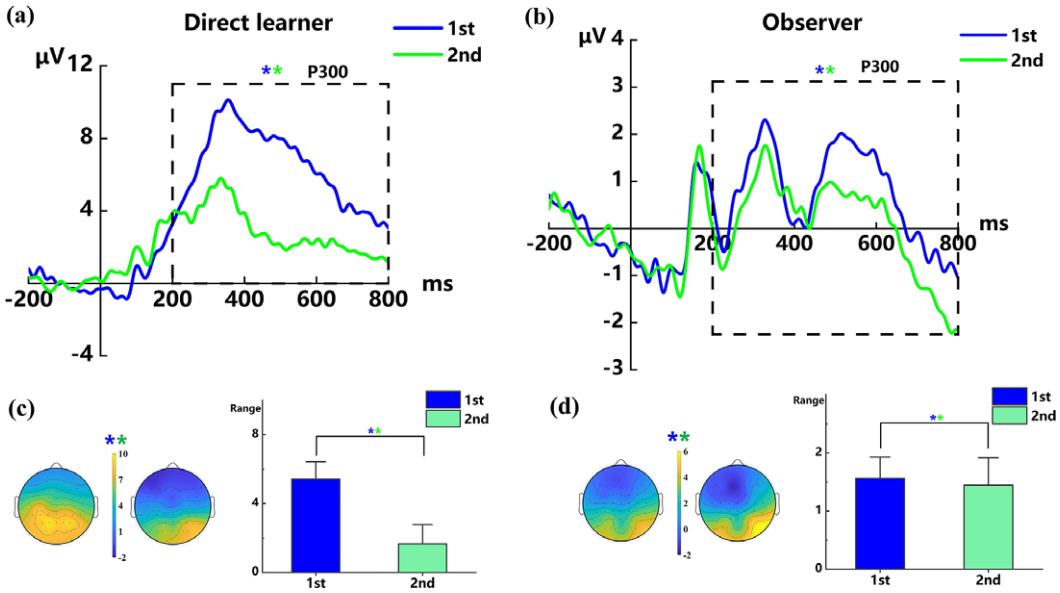


Figure 4 ERP waveforms, topographic maps and bar graphs of direct learners (a, c) and observers (b, d) in response to reversal rewards.
Notes. Colored asterisks (**) indicate significant differences between the two conditions ($p < .05$). Bar graphs show the mean voltages of the P300 across sites for direct learners and observers in the first reward post−reversal and second reward post−reversal conditions. Error bars show the standard error of means. 1st = first reward post−reversal; 2nd = second reward post−reversal.

$t = -2.05$, $p = .041$). Additionally, the post-reversal phase exhibited a more pronounced LPC effect compared to the pre-reversal phase for switch trials (post-reversal: $M = 7.55 \pm .94$ μV > pre-reversal: $M = 6.60 \pm .92$ μV, $b = .95$, $SE = .36$, $t = 2.67$, $p = .0077$), whereas no significant difference was found between pre-reversal and post-reversal phases for non-switch trials (post-reversal: $M = 6.76 \pm .94$ μV, pre-reversal: $M = 7.25 \pm .92$ μV, $b = -.49$, $SE = .36$, $t = -1.37$, $p = .17$).

3.2.3. ERP results of the effects of reversal rewards on language switching in observers

A linear mixed-effects model analysis was conducted to examine the effects of the learning phase (post-reversal and pre-reversal) and sequence type (switch and non-switch) on N2 and LPC. The results showed a significant main effect of sequence type on N2 amplitude. Specifically, switch trials exhibited larger N2 amplitudes ($M_{switch} = -.99 \pm .99$ μV) compared to non-switch trials ($M_{non\text{-}switch} = -1.67 \pm .99$ μV; $b = -.68$, $SE = .24$, $t = -2.82$, $p = .0048$). Furthermore, as shown in Figure 5b, there was an interaction between sequence type and learning stage on the LPC ($b = -1.04$, $SE = .50$, $t = -2.08$, $p = .037$). A simple effects analysis indicated that the post-reversal phase showed a switch cost, with switch trials exhibiting larger LPC amplitudes compared to non-switch trials (switch: $M = 2.57 \pm .91$ μV > non-switch: $M = 1.79 \pm .92$ μV, $b = .78$, $SE = .39$, $t = 2.02$, $p = .044$). However, there was no significant difference in LPC amplitudes between switch and non-switch trials during the pre-reversal phase (switch: $M = 3.10 \pm .90$ μV, non-switch: $M = 3.36 \pm .90$ μV, $b = -.26$, $SE = .32$, $t = -.82$, $p = .41$). Additionally, for non-switch trials, the pre-reversal phase elicited larger LPC amplitude compared to the post-reversal phase (pre-reversal: $M = 3.36 \pm .90$ μV > post-reversal: $M = 1.79 \pm .92$ μV, $b = -1.57$, $SE = .36$, $t = -4.42$, $p < .001$). However, there was no significant difference between post-reversal and pre-reversal phases for switch trials (post-reversal: $M = 2.57 \pm .91$ μV, pre-reversal: $M = 3.1 \pm .90$ μV, $b = -.53$, $SE = .35$, $t = -1.50$, $p = .134$).

4. Discussion

In this study, we have examined the effects of reversal rewards on language control in direct learners and observers. Our findings indicate that both direct learners and observers displayed highly accurate acquisition rates of the switch and non-switch behaviors in both pre- and post-reversal phases. Observers exhibited higher correct acquisition rates in non-switch behaviors than switch behaviors during the pre-reversal phase, but higher correct acquisition rates in switch behaviors than non-switch behaviors during the post-reversal phase. Furthermore, both direct learners and observers exhibited reduced switch costs in the post-reversal phase compared to the pre-reversal phase. Electrophysiological results revealed that direct learners exhibited LPC switch costs in both pre- and post-reversal phases, with reversal rewards significantly impacting switch trials. Observers showed LPC switch costs only in the post-reversal phase, with reversal rewards notably affecting non-switch trials. These findings highlight the importance of reversal rewards in adaptive language control during observational learning.

4.1. Reversal rewards impact acquisition rates of switch and non-switch behaviors for direct learners and observers

Both direct learners and observers correctly and rapidly performed language-switching behaviors across successive reversal rewards, achieving high correct acquisition rates in both pre- and post-reversal phases. This finding indicates that both direct learners and observers can flexibly adapt their language-switching behaviors in response to reversal rewards, with observers demonstrating this adaptability through observational learning. In this process, both direct and observational reversal rewards act as external stimuli that drive language control (Fröber & Dreisbach, 2016; Izquierdo et al., 2017), promoting adaptive behaviors aligned with a volatile environment. Additionally, individuals modify their actions when previously learned, highly-rewarded behaviors are reversed. The

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

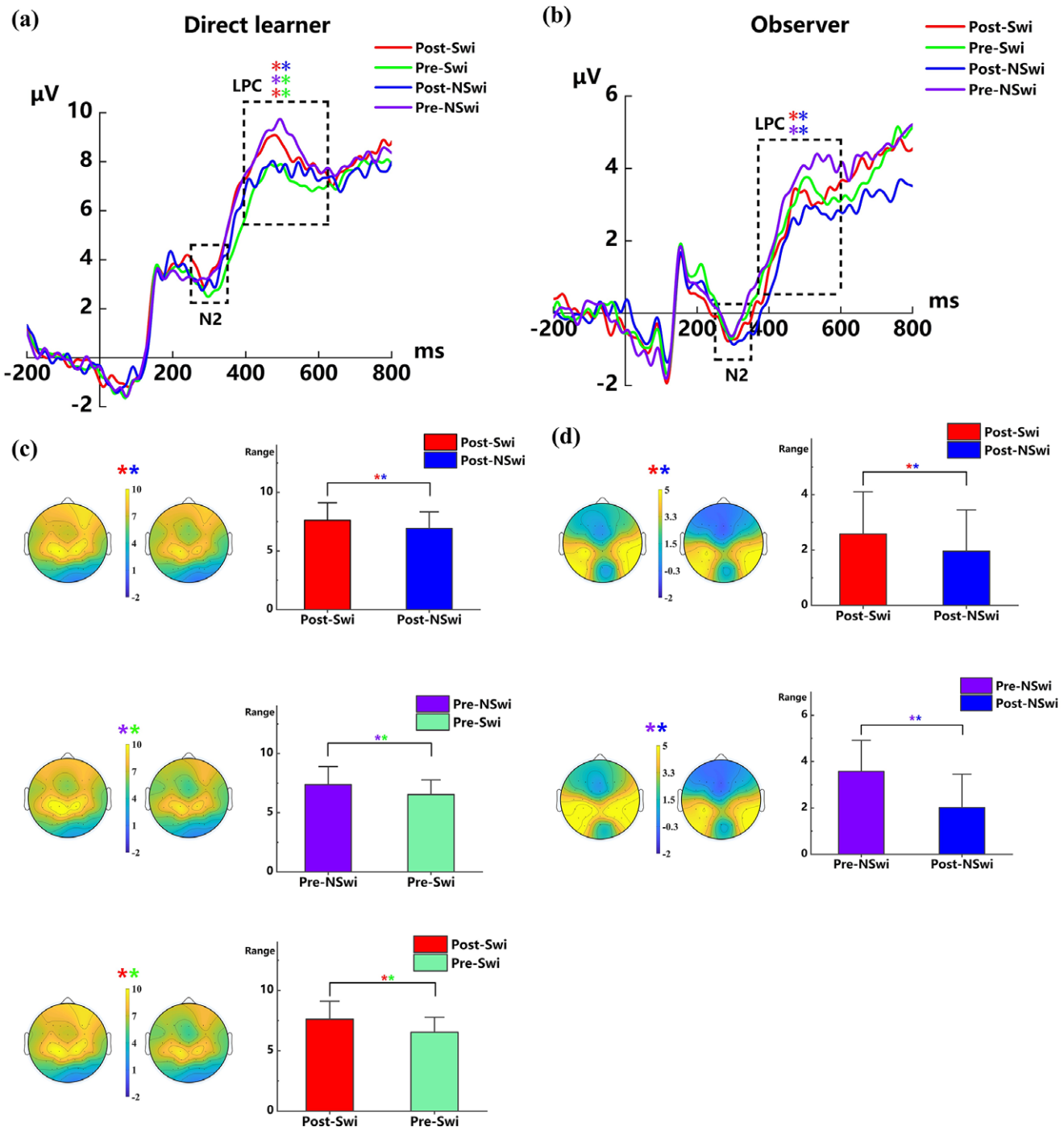


Figure 5. ERP waveforms, topographic maps, and bar graphs illustrate the impact of reversal rewards on language switching for direct learners (a, c) and observers (b, d). *Notes.* The N2 time window is defined as 250–350 ms. The LPC time window is defined as 370–600 ms. Colored asterisks (**) indicate significant differences between two conditions ($p < .05$). Bar graphs show mean voltages of the LPC across sites for direct learners and observers in the learning phase (post-reversal and pre-reversal) × sequence type (switch, non-switch) conditions. Error bars show the standard error of means. Post-Swi = Post-reversal Switch; Pre-Swi = Pre-reversal Switch; Post-NSwi = Post-reversal Non-switch; Pre-NSwi = Pre-reversal Non-switch.

performance of both direct learners and observers can be explained by Binding and Retrieval in Action Control theory (Frings et al., 2020). Specifically, direct learners are influenced by direct rewards, establishing a binding between stimulus and response, while observers rely on observational rewards informed by observed stimulus–response associations. Our findings further support Expected Value Control Theory (Shenhav et al., 2013), which posits that choices to take actions are influenced by calculating potential rewards (gains) and the required effort (costs). In our study, both direct learners and observers maximized their benefits by exerting language control based on these calculations. Both direct learners and observers chose behaviors that offered the greatest rewards: opting for switch behaviors with high rewards over non-switch behaviors with low rewards (higher cost + higher gain > lower cost + lower gain), or selecting non-switch behaviors with high rewards over switch behaviors with low rewards (lower cost + higher gain > higher cost

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

+ lower gain). Our findings are consistent with previous research on reversal rewards. In studies where a dichotomous choice is linked to probabilistic reward outcomes (Bartolo & Averbeck, 2020; Farashahi et al., 2017), both primates and humans tend to associate and acquire the choice with the highest reward and dynamically adjust their behaviors in reversed environments. Our study extends these findings by incorporating observational scenarios, revealing the impact of reversal rewards on behavioral flexibility in social situations. The acquisition of language-switching behaviors by observers aligns with findings from a study by Najar et al. (2020), which showed that observing others' rewards influenced the preferences of individuals' behavior. Similarly, Ihssen et al. (2016) demonstrated that individuals can overcome outcome uncertainty and acquire highly rewarded behaviors through observation. Our results further support these findings and extend their implications to the domain of language control.

In comparing the correct acquisition rates of language-switching behaviors between direct learners and observers in pre- and post-reversal phases, we found that observers displayed non-switch trials at a higher correct acquisition rate than switch trials in the pre-reversal phase. However, in the post-reversal phase, switch trials elicited a higher correct acquisition rate than non-switch trials. This difference can be attributed to the fact that direct learners in novel environments acquire language-switching behaviors solely through the binding between rewards and language-switching, whereas observers learn by observing others' behaviors and outcomes (Burke et al., 2010; Charpentier et al., 2020; Peterburs et al., 2021). This dynamic nature of such behavior illustrates the effect of reversal rewards on the acquisition of language-switching behaviors in social situations.

4.2. Switch costs in language control among direct learners and observers

Direct learners exhibited significant switch costs in both pre- and post-reversal phases, whereas observers displayed significant switch costs only in the pre-reversal phase. Our interpretation of this finding is that, initially, observers experience language control demands similar to direct learners when managing language switching. The absence of significant switch costs in the post-reversal phase suggests that observers efficiently adapted to new reward conditions through observational learning. By observing the behaviors and outcomes of direct learners, observers rapidly adjusted their language-switching strategies, minimizing the language control costs associated with direct learning. This finding is consistent with previous research indicating that observers estimate reward values by combining past experiences with expectations of task success. For example, Harris et al. (2018) showed that individuals can make choices for others by assessing their goals in the current situation. In this process, individuals learn from the rewards obtained by others, which leads to confidence and anticipation, resulting in similar response patterns.

Both direct learners and observers demonstrated reduced switch costs in the post-reversal phase compared to the pre-reversal phase. The overall reduction in switch costs underscores the dynamic adjustment capabilities and improved efficiency in language control mechanisms, likely due to the integration of new reward information and the refinement of language-switching strategies. This adaptation highlights the flexibility and resilience of language control processes in response to dynamic environmental changes. The ACH suggests that language control is not static, but that it adapts based on task demands and contextual changes (Green & Abutalebi, 2013). Our findings support this hypothesis, demonstrating the adaptive and responsive nature of language control mechanisms.

4.3. Reversal rewards affect the time course of language control

Our EEG findings indicated that both direct and observational rewards elicited enhanced P300 effects, highlighting the sensitivity of P300 to unexpected reward outcomes (Donchin & Coles, 1988; Fonken et al., 2020; Von Borries et al., 2013). Consequently, we infer that reversal rewards drive direct learners and observers to invest more attention and control resources, facilitating dynamic adjustments in their language-switching behaviors.

Furthermore, our results showed that reversal rewards primarily impacted the LPC for direct learners, with post-reversal switch trials eliciting significantly greater LPC effects compared to non-switch trials. Based on previous findings (Liu et al., 2016; Nicholson et al., 2005, 2006), our results suggest that direct learners exhibit more efficient extraction of information during the post-reversal phase. Reversal rewards heightened attentional demands, requiring direct learners to invest more control resources to enhance language control, resulting in pronounced LPC effects and significant behavioral switch costs. Similarly, observers showed increased LPC effects on post-reversal switch trials, indicating that observational reversal rewards also facilitate language control by triggering additional attentional resources. This process helps to inhibit interference from previously established stimulus–response patterns, ensuring effective selection and extraction of target words. Additionally, no significant N2 effects were found in direct learners, while observers showed a main effect of sequence type on N2, with greater amplitudes for switch trials. This finding suggests that reversal rewards prompted earlier cognitive adjustments in observers, who engaged control resources sooner during the task, whereas direct learners relied on more sustained control processes at later stages.

Our findings regarding direct learners and observers elucidate the application of Binding and Retrieval in Action Control theory to language control from the perspective of reversal rewards (Frings et al., 2020). Reversal rewards facilitate the updating of stimulus–response bindings and enhance language control by establishing new stimulus–response associations. This mechanism aligns with the ACH which would argue that reversal rewards, whether direct or observational, affect adaptive control mechanisms during language switching (Green & Abutalebi, 2013). This likely promotes the inhibition of non-target lexical processing during the lexical selection phase, followed by extraction.

5. Limitations and future directions

This study has provided critical insights into how reversal rewards influence language control. However, several limitations should be noted. First, we focused exclusively on reversal rewards, without examining the effects of different reward types, such as intrinsic versus extrinsic rewards, which may differentially shape language switching and cognitive control. Future research should explore these distinctions to deepen our understanding of reward-based language adaptation. Second, the laboratory setting used to collect data in the current study may have limited the ecological validity of our findings. Investigating how reversal rewards influence bilingual language control in naturalistic contexts with greater environmental complexity would enhance the generalizability of our results. Third, the sample's homogeneity in language proficiency and cultural background constrains the broader applicability of our findings. Future studies should seek to include more diverse bilingual populations to determine whether these

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press

effects are consistent across languages and cultures. Finally, while EEG can provide valuable insights into the temporal dynamics of language control, its limited spatial resolution restricts conclusions about the underlying neural mechanisms. Combining EEG with brain imaging techniques such as fMRI (functional magnetic resonance imaging) or MEG (magnetoencephalography) can offer a more comprehensive understanding of the neural basis of language switching.

6. Conclusion

This study provides novel insight into language control mechanisms by integrating reversal rewards with voluntary language-switching paradigms. Both direct learners and observers effectively acquired and adjusted their language-switching behaviors in the context of changing reward conditions. They exhibited pronounced switch costs and LPC effects influenced by reversal rewards, indicating a shared adaptive process in language control. In sum, our study contributes to research on bilingual language control by illustrating how individuals adapt to changing reward environments. The dynamic nature of switch costs and the ability to minimize these costs through learning reflects the complex interplay between language control and environmental feedback. These findings have implications for developing more effective strategies for language learning and control in bilingual contexts, emphasizing the significance of both direct and observational learning processes. The ability of observers to adapt through observation underscores the importance of social learning mechanisms in language control. By integrating observed information, observers can efficiently navigate new reward structures, supporting the broader application of observational learning in language control.

Supplementary material. To view supplementary material for this article, please visit http://doi.org/10.1017/S1366728925000070.

Data availability statement. The datasets generated and analyzed in this study are available in the OSF repository: Liu, H. (2024, July 30). Reversal reward effects on language switching. Retrieved from osf.io/5qmzx.

Author contribution. **Junjun Huang:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization. **Mengjie Lv:** Methodology, Investigation. **Yingyi Xiang:** Methodology, Investigation. **Shuang Liu:** Methodology, Formal analysis. **Yujing Shen:** Methodology, Investigation. **John W. Schwieter:** Writing – original draft, Writing – review & editing, Visualization. **Huanhuan Liu:** Conceptualization, Funding acquisition, Methodology, Formal analysis, Writing – review & editing, Supervision.

Funding statement. This research was supported by grants from the General Program of National Natural Science Foundation of China (32371089), STI 2030—Major Projects 2021ZD0200500, Dalian Science and Technology Star Fund of China (2020RQ055), and the Scientific Research and Innovation Team of Liaoning Normal University.

Competing interest. We have no known conflict of interest to disclose.

References

Abutalebi, J., **Annoni, J. M.**, **Zimine, I.**, **Pegna, A. J.**, **Seghier, M. L.**, **Lee-Jahnke, H.**,…& **Khateb, A.** (2008). Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex*, **18**(7), 1496–1505.

Abutalebi, J., **Della Rosa, P. A.**, **Ding, G.**, **Weekes, B.**, **Costa, A.**, & **Green, D. W.** (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, **49**(3), 905–911.

Allan, D. (2004). *Oxford placement test 2: Test pack*. Oxford: Oxford University Press.

Anderson, J. A., **Chung-Fat-Yim, A.**, **Bellana, B.**, **Luk, G.**, & **Bialystok, E.** (2018). Language and cognitive control networks in bilinguals and monolinguals. *Neuropsychologia*, **117**, 352–363.

Bartolo, R., & **Averbeck, B. B.** (2020). Prefrontal cortex predicts state switches during reversal learning. *Neuron*, **106**(6), 1044–1054.

Blanco-Elorrieta, E., & **Pylkkänen, L.** (2016). Bilingual language control in perception versus action: MEG reveals comprehension control mechanisms in anterior cingulate cortex and domain-general control of production in dorsolateral prefrontal cortex. *Journal of Neuroscience*, **36**(2), 290–301.

Blanco-Elorrieta, E., & **Pylkkänen, L.** (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *Journal of Neuroscience*, **37**(37), 9022–9036.

Botvinick, M., & **Braver, T.** (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology*, **66**(1), 83–113.

Brysbaert, M., & **New, B.** (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, **41**(4), 977–990.

Burke, C. J., **Tobler, P. N.**, **Baddeley, M.**, & **Schultz, W.** (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences*, **107**(32), 14431–14436.

Cai, Q., & **Brysbaert, M.** (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS ONE*, **5**(6), e10729.

Charpentier, C. J., **Iigaya, K.**, & **O'Doherty, J. P.** (2020). A neuro-computational account of arbitration between choice imitation and goal emulation during human observational learning. *Neuron*, **106**(4), 687–699.

Chiew, K. S., & **Braver, T. S.** (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, **14**, 509–529.

Crinion, J., **Turner, R.**, **Grogan, A.**, **Hanakawa, T.**, **Noppeney, U.**, **Devlin, J. T.**, … & **Price, C. J.** (2006). Language control in the bilingual brain. *Science*, **312**(5779), 1537–1540.

de Bruin, A., **Samuel, A. G.**, & **Duñabeitia, J. A.** (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language*, **103**, 28–43.

Declerck, M., **Kleinman, D.**, & **Gollan, T. H.** (2020). Which bilinguals reverse language dominance and why? *Cognition*, **204**, 104384.

Donaldson, K. R., **Oumeziane, B. A.**, **Hélie, S.**, & **Foti, D.** (2016). The temporal dynamics of reversal learning: P3 amplitude predicts valence-specific behavioral adjustment. *Physiology & Behavior*, **161**, 24–32.

Donchin, E., & **Coles, M. G.** (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, **11**(3), 357–374.

Farashahi, S., **Donahue, C. H.**, **Khorsand, P.**, **Seo, H.**, **Lee, D.**, & **Soltani, A.** (2017). Metaplasticity as a neural substrate for adaptive learning and choice under uncertainty. *Neuron*, **94**(2), 401–414.

Faul, F., **Erdfelder, E.**, **Lang, A. G.**, & **Buchner, A.** (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**(2), 175–191.

Fonken, Y. M., **Kam, J. W.**, & **Knight, R. T.** (2020). A differential role for human hippocampus in novelty and contextual processing: Implications for P300. *Psychophysiology*, **57**(7), e13400.

Frings, C., **Hommel, B.**, **Koch, I.**, **Rothermund, K.**, **Dignath, D.**, **Giesen, C.**, & **Philipp, A.** (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences*, **24**(5), 375–387.

Fröber, K., & **Dreisbach, G.** (2016). How sequential changes in reward magnitude modulate cognitive flexibility: Evidence from voluntary task switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**(2), 285.

Ghahremani, D. G., **Monterosso, J.**, **Jentsch, J. D.**, **Bilder, R. M.**, & **Poldrack, R. A.** (2010). Neural components underlying behavioral flexibility in human reversal learning. *Cerebral Cortex*, **20**(8), 1843–1852.

Glazer, J. E., **Kelley, N. J.**, **Pornpattananangkul, N.**, **Mittal, V. A.**, & **Nusslock, R.** (2018). Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *International Journal of Psychophysiology*, **132**, 184–202.

Goubert, L., **Vlaeyen, J. W.**, **Crombez, G.**, & **Craig, K. D.** (2011). Learning about pain from others: An observational learning account. *The Journal of Pain*, **12**(2), 167–174.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, **1**(2), 67–81.

Green, D. W., & **Abutalebi, J.** (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**(5), 515–530.

Gross, M., & **Kaushanskaya, M.** (2015). Voluntary language switching in English–Spanish bilingual children. *Journal of Cognitive Psychology*, **27**(8), 992–1013.

Harris, A., **Clithero, J. A.**, & **Hutcherson, C. A.** (2018). Accounting for taste: A multi-attribute neurocomputational model explains the neural dynamics of choices for self and others. *Journal of Neuroscience*, **38**(37), 7952–7968.

Heikoop, K. W., **Declerck, M.**, **Los, S. A.**, & **Koch, I.** (2016). Dissociating language-switch costs from cue-switch costs in bilingual language switching. *Bilingualism: Language and Cognition*, **19**(5), 921–927.

Ihssen, N., **Mussweiler, T.**, & **Linden, D. E.** (2016). Observing others stay or switch–How social prediction errors are integrated into reward reversal learning. *Cognition*, **153**, 19–32.

Iluz-Cohen, P., & **Armon-Lotem, S.** (2013). Language proficiency and executive control in bilingual children. *Bilingualism: Language and Cognition*, **16**(4), 884–899.

Izquierdo, A., **Brigman, J. L.**, **Radke, A. K.**, **Rudebeck, P. H.**, & **Holmes, A.** (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, **345**, 12–26.

Jackson, G. M., **Swainson, R.**, **Cunnington, R.**, & **Jackson, S. R.** (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, **4**(2), 169–178.

Jiao, L., **Meng, N.**, **Wang, Z.**, **Schwieter, J. W.**, & **Liu, C.** (2022). Partially shared neural mechanisms of language control and executive control in bilinguals: Meta-analytic comparisons of language and task switching studies. *Neuropsychologia*, **172**, 108273.

Kang, C., **Ma, F.**, **Li, S.**, **Kroll, J. F.**, & **Guo, T.** (2020). Domain-general inhibition ability predicts the intensity of inhibition on non-target language in bilingual word production: An ERP study. *Bilingualism: Language and Cognition*, **23**(5), 1056–1069.

Kleinsorge, T., & **Rinkenauer, G.** (2012). Effects of monetary incentives on task switching. *Experimental Psychology*, **59**(4), 216–226.

Liu, H., **Liang, L.**, **Dunlap, S.**, **Fan, N.**, & **Chen, B.** (2016). The effect of domain-general inhibition-related training on language switching: An ERP study. *Cognition*, **146**, 264–276.

Liu, H., **Xie, N.**, **Zhang, M.**, **Gao, X.**, **Dunlap, S.**, & **Chen, B.** (2018). The electrophysiological mechanism of joint language switching: Evidence from simultaneous production and comprehension. *Journal of Neurolinguistics*, **45**, 45–59.

Liu, H., **Tong, J.**, **de Bruin, A.**, **Li, W.**, **He, Y.**, & **Li, B.** (2020). Is inhibition involved in voluntary language switching? Evidence from transcranial direct current stimulation over the right dorsolateral prefrontal cortex. *International Journal of Psychophysiology*, **147**, 184–192.

Liu, H., **Li, W.**, **Zuo, M.**, **Wang, F.**, **Guo, Z.**, & **Schwieter, J. W.** (2022). Cross-task adaptation effects of bilingual language control on cognitive control: A dual-brain EEG examination of simultaneous production and comprehension. *Cerebral Cortex*, **32**(15), 3224–3242.

Liu, D., **Xing, Z.**, **Huang, J.**, **Schwieter, J. W.**, & **Liu, H.** (2023). Genetic bases of language control in bilinguals: Evidence from an EEG study. *Human Brain Mapping*, **44**(9), 3624–3643.

Luque, A., & **Morgan-Short, K.** (2021). The relationship between cognitive control and second language proficiency. *Journal of Neurolinguistics*, **57**, 100956.

Martin, C. D., **Strijkers, K.**, **Santesteban, M.**, **Escera, C.**, **Hartsuiker, R. J.**, & **Costa, A.** (2013). The impact of early bilingualism on controlling a language learned late: An ERP study. *Frontiers in Psychology*, **4**, 815.

McAlonan, K., & **Brown, V. J.** (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioural Brain Research*, **146**(1–2), 97–103.

Misra, M., **Guo, T.**, **Bobb, S. C.**, & **Kroll, J. F.** (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, **67**(1), 224–237.

Najar, A., **Bonnet, E.**, **Bahrami, B.**, & **Palminteri, S.** (2020). The actions of others act as a pseudo-reward to drive imitation in the context of social reinforcement learning. *PLoS Biology*, **18**(12), e3001028.

Nicholson, R., **Karayanidis, F.**, **Poboka, D.**, **Heathcote, A.**, & **Michie, P. T.** (2005). Electrophysiological correlates of anticipatory task-switching processes. *Psychophysiology*, **42**(5), 540–554.

Nicholson, R., **Karayanidis, F.**, **Bumak, E.**, **Poboka, D.**, & **Michie, P. T.** (2006). ERPs dissociate the effects of switching task sets and task cues. *Brain Research*, **1095**(1), 107–123.

Paap, K. R., **Myuz, H. A.**, **Anders, R. T.**, **Bockelman, M. F.**, **Mikulinsky, R.**, & **Sawi, O. M.** (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology*, **29**(2), 89–112.

Peterburs, J., **Frieling, A.**, & **Bellebaum, C.** (2021). Asymmetric coupling of action and outcome valence in active and observational feedback learning. *Psychological Research*, **85**, 1553–1566.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, **118**(10), 2128–2148.

Rautio, I. V., **Holmberg, E. H.**, **Kurup, D.**, **Dunn, B. A.**, & **Whitlock, J. R.** (2024). A novel paradigm for observational learning in rats. *Cognitive Neurodynamics*, **18**(2), 757–767.

Schwieter, J. W., & **Sunderman, G.** (2008). Language switching in bilingual speech production: In search of the language-specific selection mechanism. *The Mental Lexicon*, **3**(2), 214–238.

Shen, Y. J., & **Chun, M. M.** (2011). Increases in rewards promote flexible behavior. *Attention, Perception, & Psychophysics*, **73**, 938–952.

Shenhav, A., **Botvinick, M. M.**, & **Cohen, J. D.** (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, **79**(2), 217–240.

Snodgrass, J. G., & **Vanderwart, M.** (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, **6**(2), 174.

Timmer, K., **Christoffels, I. K.**, & **Costa, A.** (2019). On the flexibility of bilingual language control: The effect of language context. *Bilingualism: Language and Cognition*, **22**(3), 555–568.

Timmer, K., **Grundy, J. G.**, & **Bialystok, E.** (2017). Earlier and more distributed neural networks for bilinguals than monolinguals during switching. *Neuropsychologia*, **106**, 245–260.

Wang, Z., **Nan, T.**, **Goerlich, K. S.**, **Li, Y.**, **Aleman, A.**, **Luo, Y.**, & **Xu, P.** (2023). Neurocomputational mechanisms underlying fear-biased adaptation learning in changing environments. *PLoS Biology*, **21**(5), e3001724.

Verhoef, K., **Roelofs, A.**, & **Chwilla, D. J.** (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, **110**(1), 84–99.

Von Borries, A. K. L., **Verkes, R. J.**, **Bulten, B. H.**, **Cools, R.**, & **De Bruijn, E. R. A.** (2013). Feedback-related negativity codes outcome valence, but not outcome expectancy, during reversal learning. *Cognitive, Affective, & Behavioral Neuroscience*, **13**, 737–746.

Yee, D. M., & **Braver, T. S.** (2018). Interactions of motivation and cognitive control. *Current Opinion in Behavioral Sciences*, **19**, 83–90.

Yeung, N., **Holroyd, C. B.**, & **Cohen, J. D.** (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, **15**(5), 535–544.

https://doi.org/10.1017/S1366728925000070 Published online by Cambridge University Press