Bial

Behavioural Brain Research 236 (2013) 332–343

Contents lists available at SciVerse ScienceDirect

Behavioural Brain Research

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

Research report

Event-related potential correlates of emergent inference in human arbitrary relational learning

Ting Wang, Simon Dymond *

Department of Psychology, Swansea University, Singleton Park, Swansea SA2 8PP, United Kingdom

HIGHLIGHTS

- Examined functional–anatomical correlates of emergent relational inference.
- Relations trained between either words and pseudowords or arbitrary symbols.
- EEG was recorded during presentations of related and unrelated stimulus pairs.
- Faster, more accurate responses on symmetry and equivalence trials.
- ERPs were significant at mainly frontal–parietal and occipital sites.

ARTICLE INFO

Article history:
Received 2 July 2012
Received in revised form 17 August 2012
Accepted 21 August 2012
Available online xxx

Keywords:
Categories
Concepts
Relational inference
Stimulus equivalence
Relatedness
Directly trained
Event-related potentials

ABSTRACT

Two experiments investigated the functional–anatomical correlates of cognition supporting untrained, emergent relational inference in a stimulus equivalence task. In Experiment 1, after learning a series of conditional relations involving words and pseudowords, participants performed a relatedness task during which EEG was recorded. Behavioural performance was faster and more accurate on untrained, indirectly related symmetry (i.e., learn AB and infer BA) and equivalence trials (i.e., learn AB and AC and infer CB) than on unrelated trials, regardless of whether or not a formal test for stimulus equivalence relations had been conducted. Consistent with previous results, event related potentials (ERPs) evoked by trained and emergent trials at parietal and occipital sites differed only for those participants who had not received a prior equivalence test. Experiment 2 further replicated and extended these behavioural and ERP findings using arbitrary symbols as stimuli and demonstrated time and frequency differences for trained and untrained relatedness trials. Overall, the findings demonstrate convincingly the ERP correlates of intra-experimentally established stimulus equivalence relations consisting entirely of arbitrary symbols and offer support for a contemporary cognitive-behavioural model of symbolic categorisation and relational inference.

1. Introduction

Traditionally, psychologists and philosophers alike have been interested in the seemingly unique human propensity to classify, categorise and organise linguistic stimuli. Categorisation and concept formation abilities have long been considered to be defining features of symbolic behaviour that often cannot readily be traced to a history of direct learning. Early behavioural approaches, for instance, emphasised the role of principles of reinforcement, discrimination and generalisation in concept learning based on direct learning [1]. Recently, behavioural psychology has developed a fruitful and rigorous approach to the study of categorisation and symbolic behaviour, called derived relational responding, which is based on Sidman's [2] stimulus equivalence paradigm. Historically, the phenomenon of stimulus equivalence dates back to ancient Greece [3] and was studied by experimental psychologists, such as stimulus–response (S–R) theorists [4] for decades until the demise of S–R psychology [5]. However, it was not until the early 1970s that Sidman rediscovered the topic and set about devising a coherent set of experimental procedures and terminology with which to study it [6].

Research on stimulus equivalence and other forms of derived relational responding has generated considerable interest because it may provide a novel approach to the investigation of unlearned or emergent categorisation skills involving physically distinct, arbitrarily related stimuli. The basic finding shows that when verbally-able humans learn a series of interconnected conditional discriminations, the stimuli often become related to one another in ways not explicitly trained. For instance, if choosing Stimulus B in

* Corresponding author. Tel.: +44 1792 295602; fax: +44 1792 295679.
E-mail addresses: t.wang@swansea.ac.uk (T. Wang), s.o.dymond@swansea.ac.uk (S. Dymond).

0166-4328/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbr.2012.08.033

the presence of Stimulus A is taught (i.e., A–B), and choosing Stimulus C in the presence of Stimulus A (i.e., A–C) is also taught, it is likely that untrained relations will emerge between B and A, C and A (called *symmetry*), B and C, and C and B (called *combined symmetry and transitivity*, or *equivalence*), in the absence of any feedback. When these relations emerge, a stimulus equivalence relation is said to have formed among the relata [6,7]. These untrained, but nonetheless predictable, derived stimulus relations have been the focus of concerted research attention precisely because they are not readily explained by traditional behavioural principles of discrimination and stimulus generalisation. Neither B nor C, for instance, have a history of differential reinforcement with regard to each other (a defining feature of discrimination learning), therefore, neither should control selection of the other. Also, the outcomes cannot be accounted for on the basis of generalisation because the stimuli are all physically dissimilar or by simpler conditioning processes because that would involve an appeal to backward conditioning, which is weak at best [8].

The research interest generated by derived relations such as stimulus equivalence partly stems from the absence of an unequivocal demonstration of symmetry, transitivity, and equivalence in nonhumans [9]. There is minimal evidence of transitivity in rats [10], pigeons [11,12], chimpanzees [13], and monkeys [14]; of symmetry in rats [15], chimps [16], baboons and rhesus monkeys [17]; and widely contested, and as yet un-replicated, demonstrations of equivalence in chimpanzees [18] and sea lions [19]. The difficulties encountered obtaining positive test outcomes in nonhuman research on derived relations contrasts with the apparent ease with which humans, even young infants [20], have in passing tests for equivalence relations. This has lead many to propose that deriving stimulus relations may be a species-specific human ability underpinned by, or otherwise related to, language [9,21–23].

Indeed, contemporary behavioural theories of human language and cognition contend that derived relations such as stimulus equivalence may provide a new approach to studying symbolic behaviour and category formation because of the similarity between symmetry and equivalence test outcomes and the inherent bidirectional nature of word-referent relations in natural language [6,21,22]. According to these accounts, the relational nature of language, as studied with derived stimulus relations, provides the basis of modern accounts of language that overcome the limitations of previous approaches [21]. Yet, if research on derived relations is to inform contemporary behavioural accounts of language, then many of the effects seen in natural language research should be capable of being replicated with intra-experimentally acquired relations [21,24]. Support for this position comes from behavioural studies showing that reaction times to equivalently related word pairs are significantly faster than non-equivalently related word pairs [25]. Such findings show that equivalence relations may produce effects similar to semantic priming effects observed in the cognitive literature when real words are used [26].

Related research on the neural mechanisms underlying equivalence relations has recently begun to identify potential overlap with brain regions responsible for language and semantic processing. For instance, neuroimaging studies highlight a key role for frontal–subcortical and frontal–parietal networks in the emergence of derived relations [27–31]. One study [27] found activation in inferior frontal (dorsolateral) and inferior parietal regions for both trained (i.e., A–B and B–C) and derived (i.e., B–A, C–B, A–C, and C–A) relations. Schlund et al. [28,29] extended these findings by conducting event related analyses of BOLD signal change correlated with blocks of trained, derived relations and matching control relations. They found that accurate responding on trained and derived trials was correlated with bilateral activation in inferior parietal lobule, dorsolateral and ventrolateral inferior frontal regions, and in area of the thalamus and globus pallidus. Activation in the inferior parietal region during stimulus equivalence tasks is consistent with studies that have examined transitive inference and other forms of inference-based relational learning [32,33]. The recruitment of the hippocampus during transitivity and equivalence relations, and the demonstration of activation in the parahippocampus to cross-class (unrelated) control tasks further reveal hippocampal involvement in relational inference tasks such as tests for stimulus equivalence relations that do not rely on a serial order structure [29]. Recently, prefrontal, medial frontal, and intraparietal cortices were activated during tests for symmetry, transitivity and equivalence, with additional activation in the precuneus and posterior parietal cortex during transitivity and equivalence testing [30]. Taken together, neuroimaging studies of stimulus equivalence indicate a role for a distributed frontal–parietal, hippocampal system in integrating and processing non-adjacent stimuli in a manner resembling that reported by previous studies using similar tasks [28,33].

To date, only a handful of event-related potential (ERP) studies have used the stimulus equivalence paradigm to determine the functional–anatomical correlates of untrained, inferential categorisation. In the first such study, Barnes-Holmes et al. [34] sought to examine whether the N400 ERP [35], evoked by semantic incongruity, was modulated by related and unrelated non-words learned intra-experimentally. First participants were exposed to training with a series of conditional relations involving non-existing words ('pseudowords', referred to here for the purposes of clarity with alphanumerics: A1–B1–C1–D1, A2–B2–C2–D2; participants were never exposed to these labels) until a high mastery criterion was achieved (*M* trials completed = 398). Next, pairs of stimuli were presented while EEG was recorded in a two-word lexical decision task with combinations of class–class (i.e., directly trained [A1–B1], symmetry [B1–A1], transitivity [A1–C1], and equivalence [C1–A1]) and class–nonclass trials (i.e., A1-C2), along with trials involving novel nonsense words (class–nonsense, nonsense–class, and nonsense–nonsense). Stimuli consisted entirely of pronounceable, orthographically regular six-letter pseudowords [36]. Reaction time and error data showed participants responded significantly faster and made fewer errors with directly and indirectly related pairs than to unrelated pairs. Generally, ERPs revealed a greater negativity indicative of the N400-like effect during unrelated pairs (i.e., class–nonclass and novel trial types) than to directly trained or equivalent pairs. The directly trained, equivalent, and nonequivalent trials selected for analysis by Barnes-Holmes et al. [34] closely resemble the directly related, indirectly related, and unrelated word pairs often examined in ERP studies of semantic priming [26,37] and provide preliminary support for a derived relations model of semantic-like processing of pseudowords.

A further study [38] compared ERPs to pairs of identical stimuli (i.e., reflexivity trials) and related and unrelated equivalence stimuli presented in a two-choice matching to sample task. Occipital P2 and frontal N2 components were evoked during reflexivity trials (where the matched stimuli are identical), while a later component, a parietal P3, was observed during equivalence test trials. Subtracted ERP components (unrelated-reflexivity and unrelated-equivalence) revealed a significant dN400-like effect with similar scalp topography to that reported in previous studies on stimulus equivalence and relational matching [34,39,40], thus providing further evidence of semantic-like processing during tests for equivalence.

A final study on the ERP correlates of stimulus equivalence [41] compared related vs. unrelated stimulus pairs in small groups of participants who received EEG recording either before or after matching to sample tests for equivalence. Only those participants who received EEG recording after equivalence tests showed an N400-like effect, as revealed by difference waves (unrelated-related). These findings are limited, however, by the small sample

size ($n = 8$) and the absence of statistical analysis of both the behavioural and electrophysiological data.

Thus, the small body of existing ERP studies on the functional–anatomical correlates of stimulus equivalence-based categorisation have several potential limitations and inconsistencies that detract from their findings. First, the effect of prior equivalence relations testing has not been systematically investigated, despite variations in the administration of such tests. Barnes-Holmes et al. [34] recorded EEG prior to a formal test for equivalence relations in order to rule out competing explanations in terms of mediated priming (as the equivalent stimuli, C1–A1 for example, would have been "paired" together on screen during the equivalence test). Barnes-Holmes et al. did not, however, compare ERPs in groups with and without a prior equivalence test. Although Haimson et al. [41] administered a behavioural equivalence test before and after EEG recording, their findings are limited by the small sample size and lack of statistical analysis of the ERP data. Yorio et al. [38] recorded EEG during one exposure to the test for equivalence relations in a matching-to-sample format identical to training. The present study will systematically compare the effects of prior equivalence testing on ERPs to equivalent and nonequivalent stimuli with two groups of participants.

Second, EEG will be recorded during a relatedness decision task distinct from the task used to train conditional relations [41]. During the relatedness decision task, participants will be exposed to prime-target stimulus pairs and instructed to decide whether or not they are related. This procedural modification removes the need to deceive participants that some of the stimuli are foreign words and nonsense words [34] and makes the ERPs evoked by within- and between-class trials the main comparison of interest.

Third, the relatedness decision task employed in the present study permits, for the first time, within-participant comparisons to be made between ERPs to related and unrelated real words. Such comparisons are an important test of the derived relations model of semantic-like processing. Fourth, the size of the intra-experimentally established stimulus relations has varied across studies. Haimson et al. [41] employed the greatest number of relations (3, 6-member: A1–B1–C1–D1–E1–F1, A2–B2–C2–D2–E2–F2, A3–B3–C3–D3–E3–F3), while Barnes-Holmes et al. [34] and Yorio et al. [38] adopted 2, 4-member (A1–B1–C1–D1, A2–B2–C2–D2) and 2, 3-member relations (A1–B1–C1, A2–B2–C2), respectively. The size of the relations used is important, not only in ensuring that as many participants as possible meet criterion and pass tests, but also if derived relations are to provide a model of the kinds of semantic processing seen with real world categories that contain many distantly related members.

Finally, the sample sizes of previous studies [e.g., 41] may have been either too small to identify significant between-group differences or underpowered to detect important within-group differences in related and unrelated trial-types. Therefore, the present study adopted an intermediate size relational network (4, 3-member; A1–B1–C1, A2–B2–C2, A3–B3–C3, A4–B4–C4) consisting entirely of unrelated real words and pseudowords trained and tested in a manner previously shown to be effective [41].

Thus, the present study sought to systematically investigate the ERP correlates of equivalence relations prior to, and following, matching to sample behavioural testing for stimulus equivalence with a large sample of participants. We predicted faster accuracy and fewer errors on related compared to unrelated trials involving pseudowords, with a graded effect across degrees of relatedness (e.g., trained and symmetry trials should be faster than equivalence trials, etc.). We also predicted significant differences in ERPs evoked by related trials when EEG was recorded prior to a formal match to sample test for stimulus equivalence. Specifically, we expected a late parietal P3 component to be evoked during related symmetry and equivalence trials consistent with previous findings [38].

2. Method

2.1. Participants

Forty-six, right-handed [42] students (22 males, 24 females), ranging in age from 19 to 29 years ($M = 22.33$, $SD = 2.87$) with normal or corrected-to-normal vision were randomly assigned to one of two groups: EEG recording before (Pre-Equiv.; $n = 26$) or after equivalence testing (Post-Equiv.; $n = 20$). The groups did not differ in age ($W = 62.5$, $Z = -.285$, $p = 0.776$). Prior to the study, all participants confirmed they were in good physical health and had no history of epilepsy or drug or alcohol dependence. The Department of Psychology Ethics Committee at Swansea University approved the study. Participants were compensated with £10 at the end of the experiment and could withdraw, without penalty, at any time.

2.2. Apparatus

All tasks were operated using a Pentium® 4 processor Dell® PC linked to a 7-key Cedrus® (RB-730) response box. The first key was removed and covered, and the second and last key were labelled "NO" and "YES", respectively. All stimuli were presented on an 18-in. flat screen monitor with a viewing distance of 80 cm, whereby each stimulus subtended a visual angle of approximately 3.2–4° in width and 0.9° in height.

During the training (and testing) and relatedness tasks, the stimuli consisted of pseudowords ('vartle', 'sinald', 'drager', 'rettes', 'troper', 'gedeer', 'wollef', and 'casors' [36]) and English words (school, nature, coffee, and doctor [43]). In addition, the following words were used during the practice trials that preceded the relatedness task: chair, table, door, fish, cat, nose, finger, eye, summer, and winter.

2.3. Procedure

2.3.1. Phase 1: single-comparison A–B and A–C training

In this phase, a sample and one "comparison" stimulus (i.e., the positive comparison) were presented on every trial. Participants were instructed that their task was to match English words to foreign words and that they should first read the sample and click on it, followed by the comparison to be presented in one of the corners of the screen.

The samples always consisted of one of the A stimuli (A1, A2, A3 or A4) and the comparisons were one of the B (B1, B2, B3 and B4) or C (C1, C2, C3 and C4) stimuli, respectively. Trials commenced with presentation of the sample in the centre of the screen, which remained in place throughout the trial. After a mouse click to the sample, a single "comparison" stimulus was presented 500 ms later in one of the four corners of the screen. Clicking on the comparison stimulus was followed by the word "Correct", which was presented below the sample in the centre of the screen for 1 s. Each of the eight trials (A1–B1, A2–B2, A3–B3, A4–B4, A1–C1, A2–C2, A3–C3, and A4–C4) was presented twice in a block of 16 trials, and participants had to achieve 100% correct to complete this phase.

2.3.2. Phase 1.1: multiple-comparison A–B and A–C training

In this phase, one sample and four comparisons were presented on every trial. Participants were again instructed to match English words to foreign words and to first read the sample and click on it, which would then be followed by presentations of the four comparisons in each of the corners of the screen.

Each sample was conditionally related to two comparisons, one B stimulus and one C stimulus. For instance, in the presence of the A1 sample, B1 was the positive comparison, and B2, B3, and B4 were negative comparisons. Similarly, in the presence of the A1 sample, C1 was the positive comparison, and C2, C3, and C4 were negative comparisons. Each of the eight trials (A1–B1, A2–B2, A3–B3, A4–B4, A1–C1, A2–C2, A3–C3, and A4–C4) was presented four times in a block of 32 trials, and participants had to achieve a minimum of 90% correct (i.e., 29/32) to complete this phase. Participants were re-exposed to this phase until criterion was met.

2.3.3. Phase 1.2: multiple-comparison A–B and A–C training (reduced feedback)

This phase was identical to Phase 1.1 except that feedback only followed half of all trials. If the 90% minimum criterion was not met, participants were re-exposed to Phase 2. Participants were told they would not receive feedback on all of their choices.

2.3.4. Phase 2: derived symmetry and equivalence relations testing

Prior to the onset of this phase, participants were informed they would not receive feedback after each choice but that it was still possible to get every choice correct based on what was previously learned. In this phase, symmetry (B–A and C–A) and equivalence (B–C and C–B) test trials were interspersed with baseline A–B and A–C training trials in a block of 64 trials. Each of the eight baseline training trials (A1–B1, A2–B2, A3–B3, A4–B4, A1–C1, A2–C2, A3–C3, and A4–C4) was presented four times. Feedback followed half of the trials. Each of eight symmetry test trials (B1–A1, B2–A2, B3–A3, B4–A4, C1–A1, C2–A2, C3–A3, and C4–A4) and the eight equivalence test trials (B1–C1, B2–C2, B3–C3, B4–C4, C1–B1, C2–B2, C3–B3,

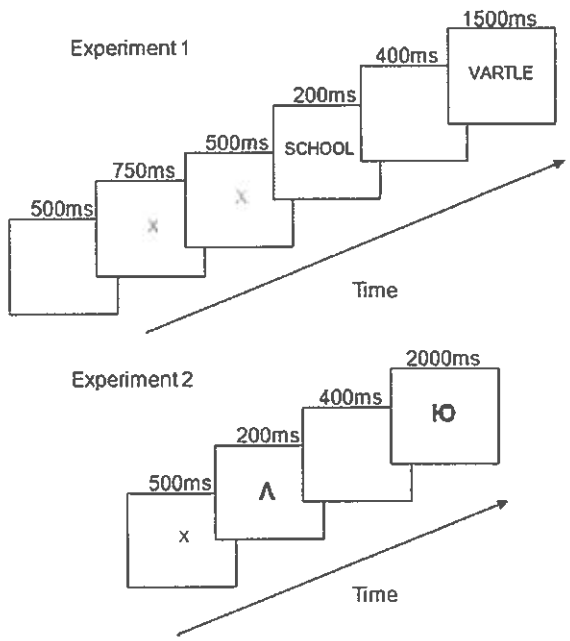

Fig. 1. Timings of the relatedness decision tasks used in Experiments 1 and 2. Participants were instructed to read the first word silently (Experiment 1) or observe the first symbol (Experiment 2) and to decide as fast as they can whether or not the second word or symbol was related to the first by pressing one of two marked keys.

and C4–B4) was presented twice. No feedback followed any of the symmetry or equivalence test trials. Participants were deemed to have passed the test for derived relations if they produced at least 80% accuracy on each of the three trial types (baseline, symmetry and equivalence).

2.3.5. Phase 3: relatedness decision task

In this phase, EEG was recorded while participants performed a relatedness decision task using two-button response box. Participants first received a practice block of 18 trials with real words not used in the other experimental phases and were instructed to press the marked "Yes" or "No" key as fast and accurately as possible if the second word is either related or unrelated to the first word. Trials were 3850 ms in duration and commenced with a blank screen for 500 ms, followed by a red 'X' in the centre of the screen for 750 ms, a green 'X' for 500 ms, and then by the prime for 200 ms and, after 400 ms, by the target, which was presented for 1500 ms (see Fig. 1).

Prime-target pairs were presented from directly trained (e.g., A1–B1 [Within-class], A1–B2 [Between-class]), symmetry (e.g., B1–A1 [Within-class], B1–A2 [Between-class]), equivalence (e.g., B1–C1 [Within-class], B1–C2 [Between-class]), and related and unrelated real word trial types (see Table 1). Within-class trials were those consistent with the training received (i.e., related), between-class trials were those inconsistent with the training received (i.e., unrelated), and each involved presentations of directly trained, symmetry, and equivalence trials (Fig. 2). "Yes" and "No" responses were equated across the tasks (Table 1), and each trial type was presented twice in a randomised block of 144 trials.

Participants in the Pre-Equiv. group received *Phase 1*, *Phase 1.1*, *Phase 1.2*, *Phase 3*, and *Phase 2*, while participants in the Post-Equiv. group received *Phase 1*, *Phase 1.1*, *Phase 1.2*, *Phase 2*, and *Phase 3*, in that order.

2.3.6. EEG recording and analysis

Voltage scalp recordings were obtained using an ActiveTwo BioSemi™ electrode system (BioSemi, Amsterdam, The Netherlands) from 34 sintered Ag–AgCl electrodes (32 scalp electrodes) arranged according to the 10–20 system (Fp1–2, AF3–4, F3–4, F7–8, Fz, FC1–2, FC5–6, T7–8, C3–4, Cz, CP1–2, CP5–6, P3–4, P7–8, Pz, PO3–4, O1–2, Oz, plus a Common Mode Sense active electrode and Driven Right Leg passive electrode). The EEG was digitised at 2048 Hz and acquired through a low-pass filter of 100 Hz and a high-pass filter of 0.16 Hz. All impedances were kept below 20 kΩ throughout recording.

Data were analysed using BESA® (Brain Electrical Source Analysis) Version 5.1.8 (MEGIS Software GmbH, Germany). Epochs of 900 ms with a pre-stimulus baseline of 200 ms were analysed. Data were filtered with a .30 Hz low-pass filter (6 dB/octave) and a high-pass filter of 40 Hz (12 dB/octave), and baseline correction was performed in relation to a pre-stimulus baseline of 200 ms before the epoch of interest (target presentation). Mean amplitude was measured for two epochs of 250–350 ms and 350–550 ms, respectively, following target onset. Trials on which the horizontal electro-oculogram amplitude exceeded 150 μV, and those on which the vertical electro-oculogram threshold exceeded 250 μV were rejected from analysis [44]. Prepotent RTs (i.e., <200 ms/>2000 ms) were excluded from the ERP analysis. Behavioural RT data and average ERPs for within- and between-class baseline, symmetry, and equivalence relations were obtained, and mean amplitudes were subjected to mixed model analysis of variance (ANOVA) on the between-subject factor of group and the within-subject factors of trial type and electrode sites. The False Discovery Rate (FDR) was applied as the correction method for multiple comparisons [45]. Independent samples *t*-tests were conducted on Between- and Within-class, and related/unrelated real words trials. Error data were analysed with Wilcoxon Signed-Rank Tests.

Table 1
Categories of within-class (i.e., related) and between-class (i.e., unrelated) prime-target pairs and correct responses arranged as directly trained, symmetry, equivalence, related and unrelated real words trial-types that were presented in the relatedness decision task in Experiment 1.

Within-class			Between-class		
Prime	Target	Response	Prime	Target	Response
Directly trained			Directly trained		
A1	B1	Yes	A1	B4	No
A1	C1	Yes	A1	C4	No
A2	B2	Yes	A2	B3	No
A2	C2	Yes	A2	C3	No
A3	B3	Yes	A3	B2	No
A3	C3	Yes	A3	C2	No
A4	B4	Yes	A4	B1	No
A4	C4	Yes	A4	C1	No
Symmetry			Symmetry		
B1	A1	Yes	B1	A4	No
C1	A1	Yes	C1	A4	No
B2	A2	Yes	B2	A3	No
C2	A2	Yes	C2	A3	No
B3	A3	Yes	B3	A2	No
C3	A3	Yes	C3	A2	No
B4	A4	Yes	B4	A1	No
C4	A4	Yes	C4	A1	No
Equivalence			Equivalence		
B1	C1	Yes	B1	C4	No
C1	B1	Yes	C1	B4	No
B2	C2	Yes	B2	C3	No
C2	B2	Yes	C2	B3	No
B3	C3	Yes	B3	C2	No
C3	B3	Yes	C3	B2	No
B4	C4	Yes	B4	C1	No
C4	B4	Yes	C4	B1	No
Related real words			Unrelated real words		
N1	Y1	Yes	N1	Y12	No
N2	Y2	Yes	N2	Y11	No
N3	Y3	Yes	N3	Y10	No
N4	Y4	Yes	N4	Y9	No
N5	Y5	Yes	N5	Y8	No
N6	Y6	Yes	N6	Y7	No
N7	Y7	Yes	N7	Y6	No
N8	Y8	Yes	N8	Y5	No
N9	Y9	Yes	N9	Y4	No
N10	Y10	Yes	N10	Y3	No
N11	Y11	Yes	N11	Y2	No
N12	Y12	Yes	N12	Y1	No

3. Results

Six participants in the Pre-Equiv. group failed to achieve mastery criterion during the equivalence test and were excluded from further analysis. All participants in the Post-Equiv. group achieved criterion, leaving a final $n = 20$ in each group. Table 2 shows the mean number (and standard deviation) trials to criterion in *Phase 1*, *Phase 1.1*, and *1.2*, and the percentage correct during baseline, symmetry and equivalence trials in *Phase 2*. The proportion of eligible trials included in the ERP analysis was 90.9% Directly Trained (DT), 87.8% Symmetry (SY), 67.5% Equivalence (EQ), 82.1% Within-class (Within), 84.3% Between-class (Between), 91.7% (related real

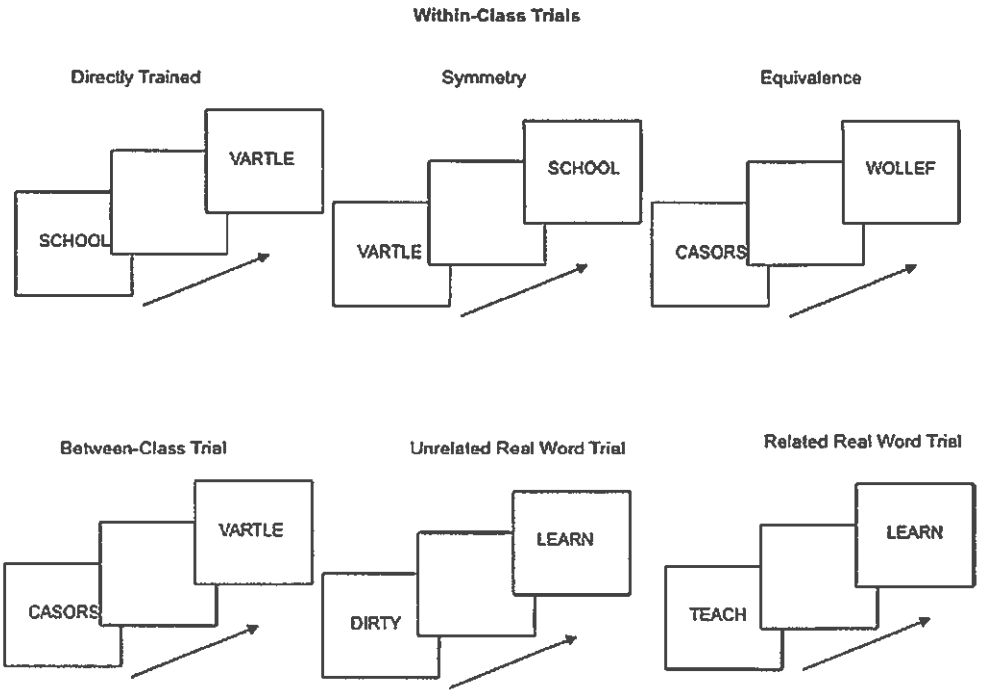

Fig. 2. Examples of within-class (i.e., related) directly trained, symmetry and equivalence trials (upper panels) and between-class (i.e., unrelated) and related and unrelated real world relatedness decision trials (lower panels) presented in Experiment 1.

words), and 96.4% (unrelated real words) of trials for the Pre-Equiv. Group, and 93.7% (DT), 94.4% (SY), and 83.4% (EQ), 90.5% (Within), 91.4% (Between), 95.8% (related real words), and 93.1% (unrelated real words) of trials for the Post-Equiv. Group, respectively.

3.1. Behavioural results

3.1.1. RTs

A significant main effect showed that the RTs differed across DT, SY and EQ trial-types ($F(2, 76) = 84.57$, $p < 0.001$), with the largest RTs observed on EQ trials (Fig. 3). Follow-up FDR-corrected tests confirmed that RTs differed between DT and SY ($q = 0.029$), DT and EQ, and SY and EQ (both $q < 0.001$) trials. Individual t-tests found significant differences for the Between- and Within-class trial-types in the Pre-Equiv. group ($t(19) = 6.055$, $p < 0.001$) and the Post-Equiv. group ($t(19) = 5.596$, $p < 0.001$). Related and Unrelated real words differed between the Pre-Equiv. ($t(19) = -5.041$, $p < 0.001$) and Post-Equiv. groups ($t(19) = -2.315$, $p = 0.032$), respectively. Generally, RTs were faster on trained and symmetry trials than on equivalence and between trials for both groups, with Pre-Equiv. participants marginally faster than Post-Equiv. participants.

3.1.2. Errors

Significantly greater errors were made on DT and EQ ($W = 0$, $Z = -3.830$, $p < 0.001$) and on SY and EQ trials ($W = 2$, $Z = -3.537$, $p < 0.001$) in the Pre-Equiv. group, with differences between related real words and unrelated real words trials approaching significance ($p = 0.087$). The Post-Equiv. group made significantly more errors on DT and EQ trials ($W = 14$, $Z = -2.987$, $p = 0.003$) and on SY and EQ trials ($W = 9$, $Z = -3.476$, $p = 0.001$; Fig. 3). No other differences were found.

3.2. EEG results

A series of mixed model ANOVAs were calculated across three frequency ranges (i.e., alpha, beta, and theta). Significant results for mean amplitude were typically found in the 4–8 Hz (theta) range, 350 ms after target onset, whereas peak mean amplitude differences were typically found in the theta range 250 ms after target onset. Accordingly, the main findings will be presented for each of the two epochs (250–350 and 350–550 ms, respectively) in the 4–8 Hz range for Pz and Oz only (Fig. 4).

Table 2
Mean (and standard deviation) number of trials to criterion in Phase 1, Phase 1.1 and 1.2 and the percentage correct during Baseline, Symmetry and Equivalence trials in Phase 2 for all participants in Experiment 1 and 2, including those participants that failed the equivalence test in Experiment 2.

	Phase 1 Trials to criterion	Phase 1.1 Trials to criterion	Phase 1.2 Trials to criterion	Phase 2 Baseline %	Phase 2 Symmetry %	Phase 2 Equivalence %
Experiment 1						
Pre-Equiv.	16.0 (0)	51.7 (20.8)	38.2 (16.5)	98.1 (3.4)	96.8 (4.3)	97.1 (5.5)
Post-Equiv.	16.0 (0)	80 (56.6)	37.7 (12.9)	97.7 (3.5)	97.5 (4.3)	95.3 (6.4)
Experiment 2						
All	16.0 (0)	155.1 (90.4)	32.0 (0)	98.1 (2.4)	99.0 (2.3)	96.1 (5.4)
Fails	16.0 (0)	166.4 (89.2)	55.5 (49.9)	90.8 (14.8)	86.7 (17.7)	61.6 (28.2)

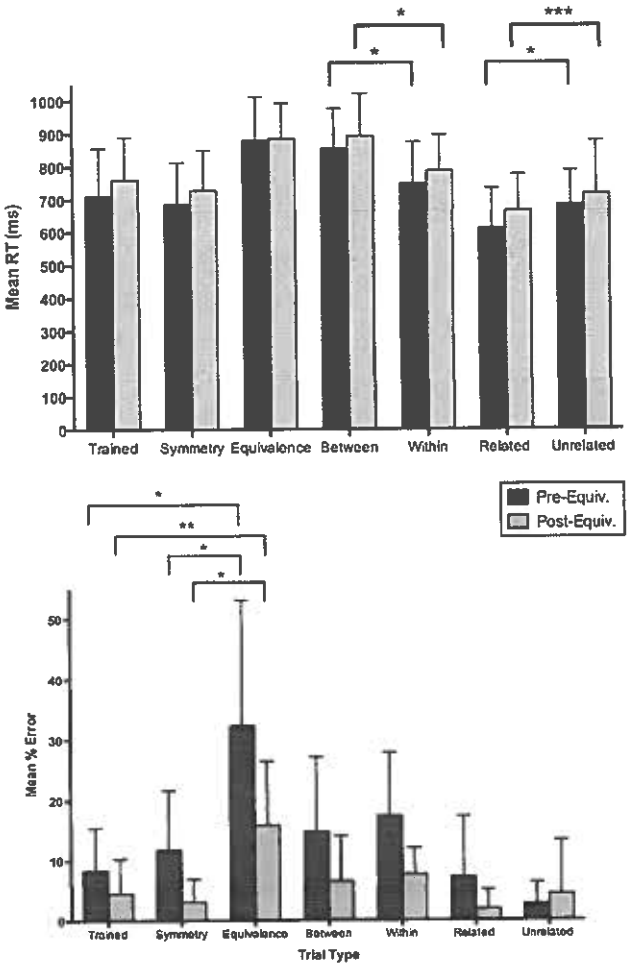

Fig. 3. Mean RT (upper panel) and error data (lower panel) for directly trained, symmetry, equivalence, between, within, and related and unrelated real words Experiment 1. Both groups responded faster and more accurately to directly trained and symmetry trials, with slower RTs and higher errors shown to equivalence and between-class trials. Bars show standard error of the mean. $^{*}p < 0.001$, $^{**}p < 0.01$, $^{***}p < 0.05$.

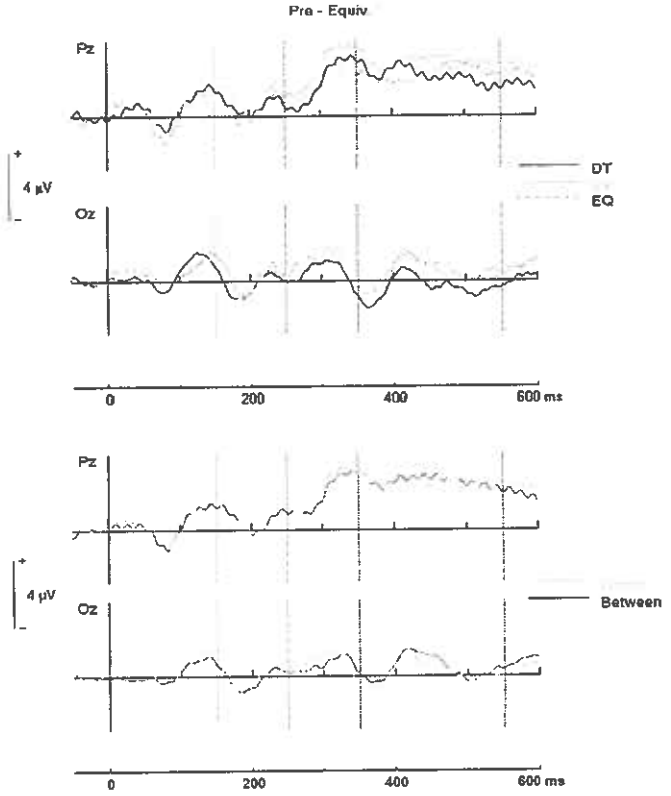

Fig. 4. Grand average waveforms from directly trained (DT), symmetry (SY) and equivalence (EQ) trials types (upper panels) and within and between-class trials (bottom panels) at Pz and Oz sites for the Pre-Equiv. group in Experiment 1. Greater positivity was observed for SY trials during the 250–350 ms epoch, with greater negativity observed for EQ trials during 350–550 ms epoch. Greater positivity was observed for within-class trials during the 250–350 ms epoch, and some evidence of a N400-like effect was observed during 350–550 ms epoch at Oz.

3.2.1. 250–350 ms

Peak mean amplitude was calculated with group as between-subject variable, trial type (DT, SY, EQ) and sites (CP1–2, CP5–6, P7–8, P3–4-Z, PO3–4, O1–2-Z) as repeated variables (Fig. 4). There was a significant main effect for trial type ($F(2, 76) = 3.34$; $p = 0.041$) and a significant interaction with sites ($F(26, 988) = 1.694$; $p = 0.017$). Follow-up tests indicated significant differences between SY and EQ ($q = 0.012$) and between DT and SY at CP1 only ($q = 0.039$). The difference between the two groups approached significance ($p = 0.08$).

3.2.2. 350–550 ms

Mean amplitude was calculated using with group as between-subject variable and trial type (DT, SY, and EQ) and sites (CP1–2, CP5–6, P7–8, P3–4-Z, PO3–4, O1–2-Z) as repeated variables (Fig. 4). There were significant main effects for trial type ($F(2, 76) = 21.02$; $p < 0.001$) and group ($F(1, 38) = 23.529$; $p < 0.001$), and two-way interactions between trial type and group ($F(2, 76) = 20.886$; $p < .0001$), between group and sites ($F(13, 494) = 2.793$; $p = 0.001$) and between trial type and sites ($F(26, 988) = 1.615$; $p = 0.027$). Follow up tests showed significant differences between DT and SY ($q = 0.016$), between DT and EQ and between SY and EQ (all $q < 0.0001$). Significant differences between the two groups, between DT and EQ and between SY and EQ were found at all 14 analysed electrode sites. Differences between DT and SY were significant at three sites (Appendix A).

3.3. Within/Between

Mean amplitude was also calculated for Within and Between trials for the 350–550 ms epoch (Fig. 4). A mixed model ANOVA with group as the between-subject variable and trial type (within and between trials) and sites (CP1–2, CP5–6, P7–8, P3–4-Z, PO3–4, O1–2-Z) as repeated variables showed significant main effects for trial type ($F(1, 38) = 4.814$; $p = 0.034$), with significant interactions between trial type and group ($F(1,38) = 12.679$; $p = 0.001$), and between trial type, group and sites ($F(13, 494) = 2.870$; $p = 0.001$). Follow up tests indicated significant differences on Within vs. Between trial types at nine electrode sites (see Appendix A), with CP2 approaching significance ($p = 0.052$). Although the two groups did not differ in ERPs evoked during this epoch ($p = 0.187$), within-class trials were significant at four electrode sites (i.e. OZ, $p = 0.017$; O2, $p = 0.032$; PO4, $p = 0.022$; P4, $p = 0.025$). Finally, a significant main effect was found for related vs. unrelated real words trial types, $F(1,38) = 9.991$; $p = 0.003$, with interactions shown between trial type and electrode sites that approached significance ($p = 0.055$).

4. Discussion

Experiment 1 replicated findings showing that derived relations produce behavioural priming-like effects of reduced RTs and fewer errors to indirectly related symmetry and equivalence pseudoword relations [25,34]. This effect was obtained in an EEG relatedness decision task administered either before (Pre-Equiv.) or after (Post-Equiv.) a matching-to-sample test for stimulus equivalence. In general, the behavioural findings of Experiment 1 showed that the Pre-Equiv. and Post-Equiv. groups performed similarly apart from the absence of significant error differences on the between- and within-trial type comparisons. From the first EEG epoch analysed (250–350 ms), there was a main effect of electrode site and an interaction with trial type in both groups, with peak mean amplitude differing between DT and SY, between SY and EQ trials. During the 350–550 ms epoch, the groups began to diverge. There were main effects and two-way interactions for all within class trial types between the two groups. However, the within/between comparison and the related/unrelated words comparison for this epoch differed regardless of group.

Our findings showing significant ERP differences at nine sites (Fig. 4) on the within/between trial types for the Pre-Equiv. group only are consistent with those of Barnes-Holmes et al. [34]. Presenting the relatedness task prior to a formal test for stimulus equivalence relations may be considered a stronger demonstration of semantic-like processing of indirectly related pseudowords since the stimuli were never directly paired onscreen at any time prior to EEG recording. To this extent, the present findings add to the literature on ERP correlates of derived relations by comparing the presence and absence of a prior relational test on subsequent behavioural and EEG measures. The clear performance advantage of the Pre-Equiv. group over the Post-Equiv. group suggests that the relatedness task functioned as a form of equivalence test itself, and that the Post-Equiv. group's familiarity with the stimuli (having been exposed to them at both training and testing) may have impaired performance. This issue warrants attention, as does the necessity of presenting related and unrelated words as control stimuli. These and other issues formed the focus of the second experiment.

5. Experiment 2

Experiment 2 was similar to Experiment 1 except for the following important differences. First, we sought to further investigate the facilitative effects of presenting the relatedness task prior to stimulus equivalence testing by only employing a Pre-Equiv. group. Second, we removed the related and unrelated control words from the crucial relatedness task in order to reduce the cognitive load. Third, participants were only included in the behavioural and EEG analyses if they produced at least 80% accuracy on directly trained, symmetry and equivalence trial types during the relatedness task and equivalence test phases, respectively. A case can be made that the relatedness task resembles an equivalence test [34], and thus we chose to focus the analysis on participants who met criterion during both phases. Finally, all stimuli were replaced with arbitrary symbols, previously shown to be effective in evoking ERPs during stimulus equivalence tasks [e.g., 38], as a means of further testing the present behavioural model of symbolic categorisation.

5.1. Method

5.1.1. Participants

Thirty-five, right-handed [42] students (16 males, 19 females), ranging in age from 18 to 36 years ($M = 22.6$, $SD = 4.4$) with normal or corrected-to-normal vision participated. Prior to the study, all participants confirmed they were in good physical health and had no history of epilepsy or drug or alcohol dependence. The Department of Psychology Ethics Committee at Swansea University approved the study. Participants were compensated with £10 at the end of the experiment and could withdraw, without penalty, at any time.

5.1.2. Apparatus

Stimuli selected from the Microsoft® Office symbols collection ('Γ', 'π', 'Λ', 'Σ', 'Ξ', 'Φ', 'Ю', 'ξ', 'Δ', 'δ', 'Ж', 'Д') were presented in black on a white background (Fig. 1).

5.1.3. Procedure

The procedure used in Experiment 2 was identical to that of Experiment 1, except for the following important differences. First, only a Pre-Equiv. group was run. Second, during *Phase 3: Relatedness decision task*, Within- and Between-class directly trained, symmetry, and equivalence trials were presented, with each trial type presented four times in a randomised block of 192 trials (Fig. 2). That is, related and unrelated real words were not presented. A balanced number of "Yes" and "No" responses were employed, and participants first received a practice block of 18 trials with real words. Third, trials were 3600 ms in duration and commenced with a red 'X' in the centre of the screen for 500 ms, the prime for 200 ms, a blank screen for 400 ms, followed by the target, which was presented for 2000 ms and was followed a 2500 ms ITI (see Fig. 1).

5.1.4. EEG recording and analysis

All details of EEG recording and analysis were identical to Experiment 1, except that repeated measures ANOVA were conducted and only participants that met criterion during the relatedness task and equivalence test phase were included in the analysis.

6. Results and discussion

Fourteen participants failed to achieve mastery criterion in the equivalence test and eight in the relatedness task, respectively. Their data were excluded from further analysis, leaving a final $n = 13$. Table 2 shows the mean (and standard deviation) number of trials to criterion in Phase 1, Phase 1.1, and Phase 1.2, and the percentage correct during baseline, symmetry and equivalence trials in Phase 2. The proportion of incorrect trials rejected from the behavioural analysis was 1.7% (DT), 1.4% (SY) and 4.4% (EQ) of trials, respectively. The proportion of eligible trials included in ERP analysis was 90.4% (DT), 95.2% (SY), 82.7% (EQ), 89.3% (Within), and 93.9% (Between) of trials, respectively.

6.1. Behavioural results

6.1.1. RTs

A significant main effect showed that the RTs differed across DT, SY and EQ trial-types ($F(2, 24) = 12.149$, $p < 0.0001$). Mean RTs on the DT, SY, and EQ, and Within and Between trial types are shown in Fig. 5. RTs were marginally faster on Within- as compared to Between-trials ($t(12) = 2.565$, $p = 0.025$), on DT over EQ trials ($q = 0.003$), and on SY than on EQ trials ($q = 0.003$). Overall, participants were faster on trained trials, and comparably slower on EQ and Between-class trial types (Fig. 5).

6.1.2. Errors

Participants made the highest proportion of mean errors on EQ trials and the fewest errors on SY trial types (Fig. 5). The proportion of errors made on Within- and Between-trials ($W = 5$, $Z = -2.298$, $p = 0.022$), DT and EQ trials ($W = 12.5$, $Z = -2.088$, $p = 0.037$), and EQ and SY trials ($W = 0$, $Z = -3.077$, $p = 0.002$) differed significantly.

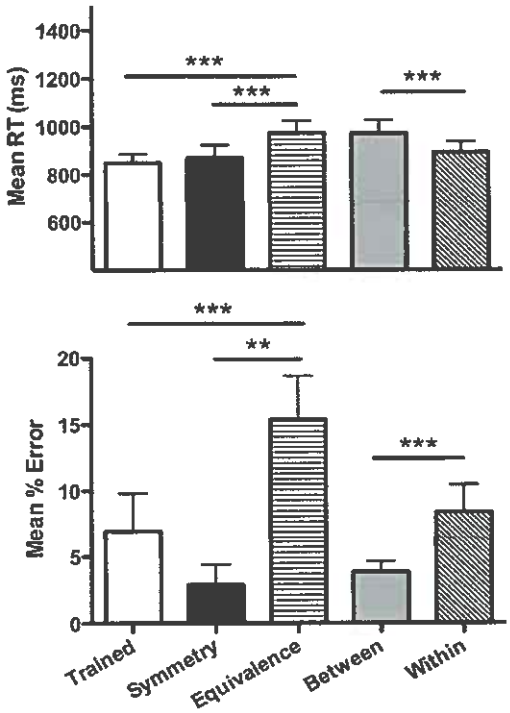

Fig. 5. Mean RT (upper panel) and error data (lower panel) for directly trained, symmetry, equivalence, between and within trial types for Experiment 2. Participants responded faster and more accurately to directly trained and symmetry trials, replicating the findings of Experiment 1, with slower RTs shown to equivalence and between-class trials and greater errors to equivalence and within-class trials. Bars show standard error of the mean. $^{**}p < 0.01$, $^{***}p < 0.05$.

indicating that EQ trials requiring the greatest relational inference resulted in the higher errors (consistent with the findings of Experiment 1).

6.2. EEG results

A series of repeated ANOVAs were calculated across three frequency ranges (i.e., alpha, beta, and theta). The main findings will be presented for each of three epochs (150–250, 250–350 and 350–550 ms, respectively) in each of the three frequency ranges, where relevant, for sites Fz, Pz, Cz and Oz (Fig. 6).

6.2.1. 150–250 ms

Mean amplitude, calculated with trial type (DT, SY and EQ) and sites (CP1–2, CP5–6, P7–8, P3–4–Z, PO3–4, O1–2–Z) as factors, revealed a significant interaction in the alpha range, $F(38, 456) = 5.359$, $p < 0.001$, confirming that different relational trials evoked changes in EEG at mainly frontal–central sites (Fig. 6). Follow-up (FDR) tests confirmed significant differences between DT and SY, DT and EQ, and SY and EQ (Appendix B). Greater positivity at Fz and greater negativity at Oz, respectively, was observed during EQ trials. At Pz, the greatest mean amplitude was shown during DT trials, with no separation observed among the trial-types at Cz (Fig. 6, upper panels). In the theta frequency range, a significant interaction was found between trial types and sites, $F(38, 456) = 1.99$, $p = 0.001$. FDR corrections revealed a significant difference between DT and EQ at FC1 ($q = 0.042$).

For peak mean amplitude in the alpha range there was a significant interaction, $F(38, 456) = 2.63$, $p < 0.001$, and a main effect of trial type, $F(2, 24) = 11.911$, $p < 0.001$. Follow up tests revealed significant differences between DT and SY, DT and EQ (both $q = 0.003$), and SY and EQ ($q = 0.033$; see Appendix C).

For peak mean amplitude in the beta range, we found a main effect of trial type, $F(2, 24) = 6.249$, $p = 0.007$, which follow-up tests revealed was driven by differences between DT and EQ ($q = 0.033$) trials. A main effect of trial type was also observed in the theta frequency range, $F(2, 24) = 3.412$, $p = 0.05$, but no significant differences were found at follow-up. Thus, while main effects of relational trial-type were observed in all three frequencies in this early epoch, only the evoked changes in alpha (peak and mean amplitude) were shown to interact significantly with recording location.

6.2.2. 250–350 ms

No differences were found for mean amplitude in all three frequencies (Fig. 6), but there was a significant main effect, $F(2, 24) = 3.664$, $p = 0.041$, and interaction, $F(38, 456) = 1.431$, $p = 0.05$, for peak mean amplitude in the alpha range. FDR-corrected follow up tests showed there were significant differences between DT and EQ and SY and EQ at two electrode sites only (Appendix C).

In the beta frequency range, a significant interaction for peak mean amplitude was found, $F(38, 456) = 2.112$, $p < 0.001$, with significant differences between DT and SY and DT and EQ revealed by follow-up (FDR) tests (Appendix C). Again, peak mean amplitude alpha differed between relational trial types of varying complexity.

6.2.3. 350–550 ms

Mean amplitude in the alpha frequency range reflected significant interactions between trial-types and sites, $F(38, 456) = 4.133$, $p < 0.001$ (Fig. 6). FDR corrections revealed significant differences between DT and SY, between DT and EQ, and between SY and EQ (Appendix B). For mean amplitude theta, a significant interaction was also found, $F(38, 456) = 2.35$, $p < 0.001$, which FDR-corrected follow up tests revealed was due to significant differences between DT and SY, DT and EQ at several electrode sites (Appendix B), and between SY and EQ at P4 ($q = 0.015$). No significant interaction was found for peak mean amplitude in the theta range ($p = 0.109$).

6.3. Within/Between

Peak mean amplitude ERPs in any frequency range evoked during Within/Between trials tended to be undifferentiated during the early (150–250 ms) and mid (250–350 ms) epochs (Fig. 6). However, during the late 350–550 ms epoch, a significant interaction was found for mean amplitude in the beta frequency range, $F(19, 228) = 1.726$, $p = 0.033$, which follow-up tests revealed was significant at Oz only ($p = 0.047$).

Overall, the findings of Experiment 2 replicated and extended those of Experiment 1 in the following ways. First, the behavioural priming effects showing faster RTs on DT and SY trials and greater errors on EQ trials were replicated. For the first time, Experiment 2 showed a significant behavioural difference on within vs. between trials. Second, again for the first time, we found evidence of early (i.e., 150–250 ms) ERP differences in mean amplitude and peak mean amplitude in the alpha range (with beta peak mean amplitude and theta mean amplitude significant between DT and EQ trials). Third, during the 250–350 ms epoch, although no differences between mean amplitude across within class trial types in all three frequency ranges, peak mean amplitude differed in both alpha and beta frequencies. Fourth, during the 350–550 ms epoch, mean amplitude differed in both the alpha and theta range only. The EEG results of Experiment 2 highlighted early and late alpha differences, for all trial type comparisons, and replicated mean amplitude theta differences from Experiment 1. In summary, removing the related/unrelated words from the relatedness decision task, which was conducted prior to formal testing of equivalence relations, produced early and late ERP differences in several frequencies, mainly alpha and theta, at frontal–parietal and central

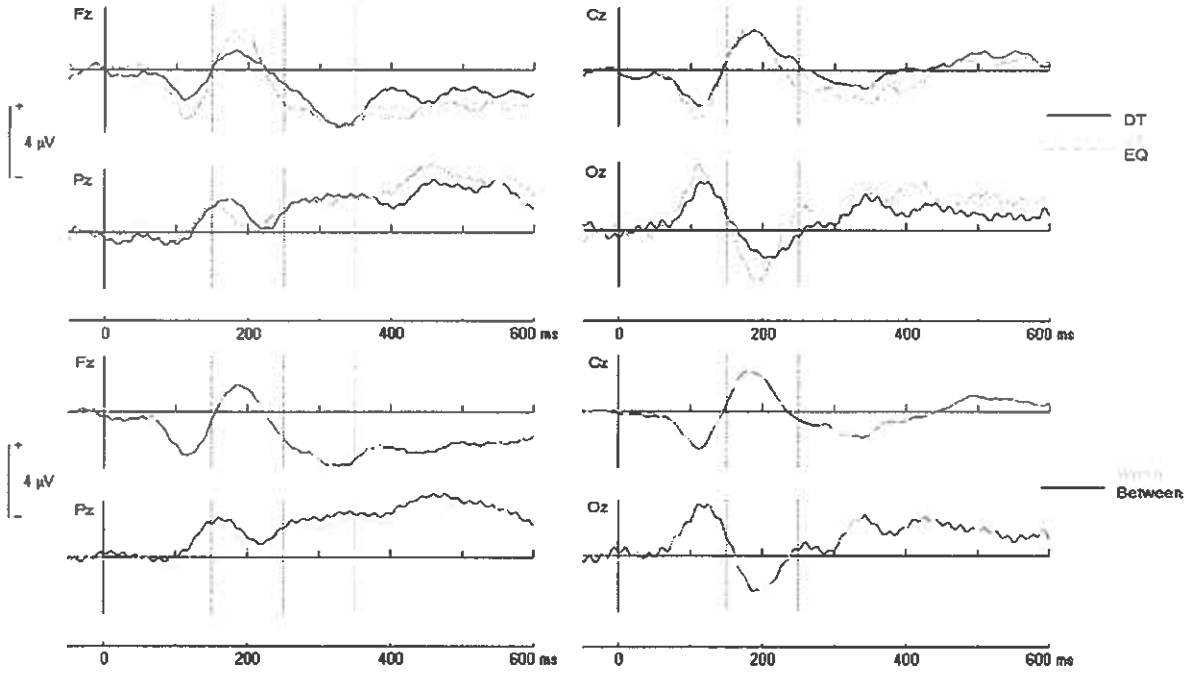

Fig. 6. Grand average waveforms for directly trained (DT), symmetry (SY) and equivalence (EQ), and within- and between-class trial types at Fz, Cz, Pz and Oz sites in Experiment 2. Significant differences were observed in the beta frequency range for DT, SY, EQ and within- and between-class trials at Oz only.

sites, and demonstrates the ERP correlates of intra-experimentally-established derived stimulus relations consisting of symbols.

7. General discussion

The present experiments sought to overcome some of the limitations of previous studies of ERP correlates of emergent relational inference in stimulus equivalence tasks by adopting a blocked training sequence to establish relations consisting of either words and pseudowords (Experiment 1) or arbitrary symbols (Experiment 2). This training was sufficient for the emergence of 4, 3-member stimulus equivalence relations, which were either formally tested before or after EEG was recorded during a relatedness decision task. In Experiment 1, control presentations of related and unrelated real words allowed us to differentiate between ERPs evoked by the intra-experimentally established pseudoword relations and those acquired via natural language processing. Behavioural findings demonstrated a clear priming-like effect of faster RTs and fewer errors on relational trials of less complexity, regardless of whether participants were tested for equivalence relations before or after the relatedness task. EEG findings indicated that only the latter group of participants (i.e., Pre-Equiv.) produced significant differences at nine sites on within/between trials. In addition, a late parietal P3 component was evoked during SY and EQ trials, and at intermediate levels for DT trials, consistent with previous findings [38].

We did not find any evidence in Experiment 1 of the N400 evoked component [35,38]. As shown in Fig. 4, a negative-going deflection was observed at Oz approximately 400 ms post-target but was greater for DT trials than SY or EQ. This is somewhat at odds with early accounts of the ERP correlates of derived relations, which posited that N400 would be evoked by presentations of related and unrelated trial types. For instance, Barnes-Holmes et al. [34] found evidence of a negative-going waveform approximating the N400 evoked by non-equivalent priming trial types, while Yorio et al. [38] identified a dN400 effect for a unrelated-equivalence difference contrast that had similar scalp topography to that of Barnes-Holmes and colleagues' findings. The fact that we observed marginally greater negativity for DT trials relative to SY or EQ trials suggests that the presentation of both related (i.e., A1–B1) and unrelated (i.e., A1–B2) directly taught trials may have permitted a greater separation between them and the test trials (which always presented derived, indirectly related trial types). Moreover, significant differences were seen on the Within- and Between-class trial types, but we did not obtain visual evidence of a comparable N400-like waveform. There was, however, some visual indication of a greater negative trend at Oz for DT trials relative to SY or EQ, respectively (Fig. 4).

By comparison, the findings of Experiment 2 provided significant evidence for the behavioural model of semantic-like processing. Behavioural results indicated significantly different RTs and errors on Within- and Between-class trials, and on DT and tested (SY and EQ) trial types. Early and late alpha differences, for all trial types, were observed, as well as a replication of mean amplitude theta differences from Experiment 1. In terms of the particular ERP components identified in Experiment 2, and late P3 component was found at parietal and occipital sites with greater positivity for EQ and SY trials (Fig. 6). No evidence was found for a N400-like component although there was some indication of a late P3-like waveform at Oz. Further research is needed to confirm whether or not this waveform, evoked using a modified version of the relatedness task employed in Experiment 1, resembles that previously identified by Yorio et al. [38].

As outlined in the Introduction, the species-specific ability to discriminate class-consistent, related, stimuli from unrelated stimuli is to establish stimulus-relevant categorisation, which typically evokes a P3 component found at central–parietal regions as a neural marker of categorisation [46–48]. In those studies, visual and auditory attributes or physical characteristics were used to establish classification between and within stimulus categories. Yorio et al. [38] provided the first empirical evidence of a larger late P3-like component in equivalence trial types (i.e., SY and transitivity relations) compared with unrelated stimuli and where the stimuli were neither directly related nor shared any physical similarities.

Our findings extend this by showing that the late P3-like component may also be applied to discrimination of within equivalence trial types. That is, we found a larger late P3-like component in SY trial types, followed by transitivity and equivalence combined trial types at Oz (Fig. 6). Taken together with previous findings [38], the present paper provides further evidence of the functional anatomical correlates of a behavioural model of categorization.

As suggested earlier, the absence of N400 may have resulted from the combinations of trial types presented, but also from the different experimental manipulations employed. For example, Barnes-Holmes et al. used 2, 4-member equivalence classes, whereas Yorio et al. employed 2, 3-member equivalence relations, and both employed a within-subjects design. While 4, 3-member equivalence relations were used in the present paper, the use of a between-subjects design in Experiment 1 made the interpretation of EEG data less straightforward. Moreover, the absence of N400 evoked by real words left one wondering would the same results be found in previous studies that demonstrated N400-like component in pseudo-words (cf. [34,38]), which could be crucial in determining whether they represent a behavioural model of semantic-like processes.

Behavioural findings in Experiment 1 revealed faster RTs and fewer errors on related trials compared to unrelated trials regardless of whether participants were tested for equivalence before or after the relatedness task. Notably, significant EEG differences were only observed in the Pre-Equiv. group that received the relatedness task prior to the formal match to sample test for equivalence relations. Experiment 2 confirmed these behavioural and EEG findings. It is likely therefore that the relatedness task was sensitive in detecting the ERP correlates of related and unrelated stimulus pairs because it may have functioned as a type of test for the formation of equivalence relations [34]. In so doing, the relatedness task may provide a more accurate behavioural model of untrained, indirect semantic relations than a formal match to sample equivalence test. Although no reinforcement or feedback is presented during a match to sample test, the indirectly related sample and comparison stimuli are associated by the selection responses used to match them. In this way, a combination of the simultaneous presentation of indirectly related stimuli and the mediating influence of matching responses is likely to have contributed to the priming-like effect for the behavioural and EEG measures in Experiment 1 (see also [25,34]). Our findings clearly support the use of relatedness or other tasks such as lexical decision or pronunciation tasks to measure priming and record EEG *prior* to a formal match to sample test for equivalence relations.

Interestingly, Yorio et al. [38] used the same matching task in all stages of their study, only recording EEG during the test phase, and found occipital P2 and frontal N2 components evoked during reflexivity trials and a later parietal P3 component during equivalence trials. By using the same task throughout, Yorio and colleagues avoided the aforementioned limitations, presented the test phase once and inferred the presence of equivalence relations at the end of the experiment (only 2 of the 12 participants scored lower than 72.5% during the equivalence test block). Future research should seek to adapt tasks shown to be effective in facilitating high yields in both behavioural studies (e.g., [49]) and neuroimaging investigations of stimulus equivalence [27–31] for use in further understanding the EEG correlates of emergent relations.

Contemporary behavioural models of relational inference argue that the ability to derive equivalence relations is a key process underlying semantic or symbolic behaviour [6,21–24]. For instance, bidirectional relations indicative of symmetry may offer a functional process account of word-referent relations in which a stimulus such as "lemon" is said to stand for or refer to actual lemons. With the addition of the written word 'lemon', the generativity is increased still further. It is this form of control exerted over behaviour that behavioural psychologists consider to underlie symbolic behaviour. In this way, humans come to respond to multiple referents that are functionally substitutable and hence equivalent for one another after minimal training. The present procedures and findings provide further empirical support for this contemporary behavioural approach to semantic or symbolic processing.

There are some potential limitations to the present findings. First, the transition from match to sample to the relatedness task may have disrupted emergent learning. This may partly explain why 23% of participants in the Pre-Equiv. group, for instance, failed tests for stimulus equivalence. As indicated above, future research should seek to address this by employing the same task throughout all stages of the experiment. Second, the size of the relational classes employed in the current study was obviously limited compared to the size of real world categories and concepts. Any tenable behavioural model of symbolic meaning needs to approximate the size and complexity of real world processes of acquiring, storing and retrieving events and facts. Thus, it may be helpful in future studies to employ methods used to study relational memory that have used up to three sets of thirty paired associates to identify neural mechanisms underpinning declarative and inferential reasoning (see [50,51]). Third, it is possible that factors such as intelligence, working memory, and executive functioning may mediate performance on arbitrary relational inference tests such as these. While we did not administer such assessments, there is evidence to suggest that relational ability is correlated with sub-scale IQ scores [52] and that relational training can boost childrens' IQ [53], future research should seek to examine whether factors such as intelligence and executive functioning ability more broadly mediate the behavioural and ERP effects obtained. Finally, the temporal resolution of EEG/ERPs could be combined with the spatial resolution of neuroimaging methods, in particular magnetoencephalography (MEG). To date, studies have used EEG/ERPs and fMRI, but not both, and MEG offers a practical means of achieving high temporal and spatial accuracy (e.g., [54]). In conclusion, the findings of the present study add support to the behavioural model of symbolic inference and have implications for further studies on the ERP correlates of the human ability to classify and categorise related and unrelated stimuli.

Acknowledgements

This research was supported by a grant from the Bial Foundation (78/08). We thank Gary Freegard and Robert Whelan for generous technical and programming assistance.

Appendix A. Critical values of mean amplitude of all interactions from Theta frequency between 350 and 550 ms. in Exp1.

		Electrode Location													
	Trial Type	CP1	CP5	P3	P7	PO3	O1	Pz	Oz	O2	PO4	P8	P4	CP6	CP2
q values	DT vs. SY			.05					.003	.007					
q values	DT vs. EQ	*	.003	.001	.001	.024	.012	.007	.013	.013	.001	.003	*	.004	*
q values	SY vs. EQ	*	.003	*	*	.003	.003	.006	*	*	*	.003	*	.004	.007
p values	Pre- vs. Post-Equiv.	*	*	.005	.002	.013	.002	.006	.003	.011	.008	*	.009	*	*
p values	Within vs. Between in Pre-Equiv. group	.014		.027		.024	.003	.006	.005	*	.001		.001		
p values	Pre- vs. Post-Equiv in within-trial								.017	.032	.022		.025		

* <.0001.

Appendix B. *q* values of mean amplitude of trial type and electrode location interaction in Exp2.

Epoch	Frequency	Trial Type	Electrode Location													
150–250 ms.	Alpha		F3	FC1	CP1	P3	P7	O1	Fz	Cz	Pz	Oz	P8	P4	CP2	FC2	F4
		DT vs. SY							.024							.015	.009
		DT vs. EQ	.012	.018			.008		.003			.047	.015			.017	
		DT vs. EQ		.038			*	.027	.042	.042		.021					
	Theta																
		DT vs. EQ		.042													
350–550 ms.	Alpha																
		DT vs. SY														.040	.014
		DT vs. EQ	.036				.003		.024	.030		.021				.003	.009
		SY vs. EQ					.047									.040	
	Theta																
		DT vs. SY							.044		*			*	.012		.015
		DT vs. EQ	.024		.030	.012			.044		.033						.021
		SY vs. EQ												.015			

* <.0001.

Appendix C. *q* values of peak mean amplitude of trial type and electrode location interaction in Exp2.

Epoch	Frequency	Trial Type	Electrode Location							
150–250 ms.	Alpha		P7	PO3	O1	Oz	O2	P8	C4	F4
		DT vs. SY	.034	.004	.038		.049			
		DT vs. EQ	.034	.003	*	*	.015		.022	.039
		SY vs. EQ		.004	.004	.006			.018	
250–350 ms.	Alpha									
		DT vs. EQ			.018				.033	
		SY vs. EQ			.018				.033	
	Beta									
		DT vs. SY			.036	.006		.045		
		DT vs. EQ			.027	.042		.045		

* <.0001.

References

[1] Keller FS, Schoenfeld WN. Principles of psychology: a systematic text in the science of behavior. Century psychology series. East Norwalk, CT, US: Appleton-Century-Crofts; 1950.

[2] Sidman M. Reading and auditory-visual equivalences. Journal of Speech and Hearing Research 1971;14:5–13.

[3] Aristotle. De Memoria et Reminiscentia [J.I. Beare, Trans.]. In: McKeon R, editor. The basic works of Aristotle. New York: Random House; 1941. p. 607–17.

[4] Hull CL. The problem of stimulus equivalence in behaviour theory. Psychological Review 1939;46:9–30.

[5] Jenkins JJ. Mediated associations: paradigms and situations. In: Cofer CN, Musgrave BS, editors. Verbal behavior and learning: problems and processes. New York: McGraw-Hill; 1963. p. 210–45.

[6] Sidman M. Stimulus equivalence: a research story. Boston: Authors Cooperative; 1994.

[7] Barnes D. Stimulus equivalence and relational frame theory. The Psychological Record 1994;44:91–124.

[8] Hall G. Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. Learning & Behavior 1996;24:233–55.

[9] Dymond S, Roche B, Barnes-Holmes D. The continuity strategy, human behaviour, and behaviour analysis. The Psychological Record 2003;53:333–47.

[10] Davis H. Transitive inference in rats (*Rattus norvegicus*). Journal of Comparative Psychology 1992;106:342–9.

[11] Lipkens R, Kop PFM, Werner M. A test of symmetry and transitivity in the conditional discrimination performances of pigeons. Journal of the Experimental Analysis of Behavior 1988;49:395–409.

[12] Von Fersen L, Wynne CDL, Delius JD, Staddon JER. Transitive inference formation in pigeons. Journal of Experimental Psychology: Animal Behavior Processes 1991;17:334–41.

[13] Boysen ST, Berntson GG, Shreyer TA, Quigley KS. Processing of ordinality and transitivity in chimpanzees (*Pan troglodytes*). Journal of Comparative Psychology 1993;107:208–15.

[14] D'Amato MR, Salmon DP, Loukas E, Tomie A. Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). Journal of the Experimental Analysis of Behavior 1985;44:35–47.

[15] Dymond S, Gomez-Martin S, Barnes D. Multimodal conditional discrimination in rats: some preliminary findings. The Irish Journal of Psychology 1996;17:269–81.

[16] Dugdale NA, Lowe CF. Testing for symmetry in the conditional discriminations of language-trained chimpanzees. Journal of the Experimental Analysis of Behavior 2000;73:5–22.

[17] Sidman M, Rauzin R, Lazar R, Cunningham S, Tailby W, Carrigan P. A search for symmetry in the conditional discrimination of rhesus monkeys, baboons, and children. Journal of the Experimental Analysis of Behavior 1982;37:23–44.

[18] Yamamoto J, Asano T. Stimulus equivalence in a chimpanzee (*Pan troglodytes*). The Psychological Record 1995;45:3–21.

[19] Schusterman RJ, Kastak D. A California sea lion (*Zalophus californianus*) is capable of forming equivalence relations. The Psychological Record 1993;43:823–39.
[20] Luciano C, Gomez Becerra I, Rodriguez Valverde M. The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. Journal of the Experimental Analysis of Behavior 2007;87:349–65.
[21] Hayes SC, Barnes-Holmes D, Roche B, editors. Relational frame theory: a post-skinnerian account of human language and cognition. New York: Plenum; 2001.
[22] Horne PJ, Lowe CF. On the origins of naming and other symbolic behavior. Journal of the Experimental Analysis of Behavior 1996;65:185–241.
[23] Wilkinson KM, McIlvane WJ. Methods for studying symbolic behaviour and category formation: Contributions of stimulus equivalence research. Developmental Review 2001;21:355–74.
[24] Hayes SC, Hayes LJ. Verbal relations and the evolution of behaviour analysis. American Psychologist 1992;47:1383–95.
[25] Hayes SC, Bisset RT. Derived stimulus relations produce mediated and episodic priming. The Psychological Record 1998;48:617–30.
[26] McNamara TP. Semantic priming: perspectives from memory and word recognition. Hove: Psychology Press; 2005.
[27] Dickins DW, Singh KD, Roberts N, Burns P, Downes JJ, Jimmieson P, et al. An fMRI study of stimulus equivalence. Neuroreport 2001;12:405–11.
[28] Schlund MW, Hoehn-Saric R, Cataldo MF. New knowledge derived from learned knowledge: functional-anatomic correlates of stimulus equivalence. Journal of the Experimental Analysis of Behavior 2007;87:287–307.
[29] Schlund MW, Cataldo MF, Hoehn-Saric R. Neural correlates of derived relational responding on tests of stimulus equivalence. Behavioral and Brain Functions 2008;4:6.
[30] Ogawa A, Yamazaki Y, Ueno K, Cheng K, Iriki A. Neural correlates of species-typical illogical cognitive bias in human inference. Journal of Cognitive Neuroscience 2010;22:2120–30.
[31] Whelan R, Schlund MW. Reframing relational frame theory research: gaining a new perspective through the application of novel behavioral and neurophysiological methods. In: Dymond S, Roche B (Eds.), Advances in relational frame theory: research and application. Oakland, CA: New Harbinger; in press.
[32] Acuna BD, Eliassen JC, Donoghue JP, Sanes JN. Frontal and parietal lobe activation during transitive inference in humans. Cerebral Cortex 2002;12:1312–21.
[33] Hinton EC, Dymond S, Von Hecker U, Evans CJ. Neural correlates of relational reasoning and the symbolic distance effect: Involvement of parietal cortex. Neuroscience 2010;168:138–48.
[34] Barnes-Holmes D, Staunton C, Whelan R, Barnes-Holmes Y, Commins S, Walsh D, et al. Derived stimulus relations, semantic priming, and event-related potentials: testing a behavioral theory of semantic networks. Journal of the Experimental Analysis of Behavior 2005;84:417–33.
[35] Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). Annual Review of Psychology 2011;62:621–47.
[36] Massaro DW, Venezky RL, Taylor GA. Orthographic regularity, positional frequency, and visual processing of letter strings. Journal of Experimental Psychology: General 1979;108:107–24.
[37] Weisbrod M, Kiefer M, Winkler S, Maier S, Hill H, Roesch-Ely D, et al. Electrophysiological correlates of direct versus indirect semantic priming in normal volunteers. Cognitive Brain Research 1999;8:289–98.
[38] Yorio A, Tabullo A, Wainselboim A, Barttfeld P, Segura E. Event-related potential correlates of perceptual and functional categories: Comparison between stimuli matching by identity and equivalence. Neuroscience Letters 2008;443:113–8.
[39] Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980;207:203–5.
[40] Nigam A, Hoffman JE, Simons RF. N400 to semantically anomalous pictures and words. Journal of Cognitive Neuroscience 1992;4:15–22.
[41] Haimson B, Wilkinson KM, Rosenquist C, Ouimet C, McIlvane WJ. Electrophysiological correlates of stimulus equivalence processes. Journal of the Experimental Analysis of Behavior 2009;92:245–56.
[42] Oldfield RC. The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia 1971;9:97–113.
[43] Holcomb PJ, Neville H. Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. Language and Cognitive Processes 1990;4:281–312.
[44] Ille N, Berg P, Scherg M. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. Journal of Clinical Neurophysiology 2002;19:113–24.
[45] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B 1995;57:289–300.
[46] Azizian A, Freitas AL, Watson TD, Squires NK. Electrophysiological correlates of categorisation: P300 amplitude as index of target similarity. Biological Psychology 2006;71:278–88.
[47] Duncan-Johnson CC, Donchin E. The P300 component of the event-related brain potential as an index of information processing. Biological Psychology 1982;14:1–52.
[48] Kotchoubey B, Lang S. Event-related potentials in an auditory semantic oddball task in humans. Neuroscience Letters 2001;310:93–6.
[49] Fields L, Doran E, Marroquin ME. Equivalence class formation in a trace stimulus pairing two-response format: effects of response labels and prior programmed transitivity induction. Journal of the Experimental Analysis of Behavior 2009;92:57–84.
[50] Preston AR, Shrager Y, Dudukovic NM, Gabrieli JDE. Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus 2004;14:148–52.
[51] Zeithamova D, Schlichting ML, Preston AR. The hippocampus and inferential reasoning: Building memories to navigate future decisions. Frontiers in Human Neuroscience 2012;6:70.
[52] O'Hora D, Pelaez M, Barnes-Holmes D, Rae G, Robinson K, Chaudhary T. Temporal relations and intelligence: correlating relational performance with performance on the WAIS-III. The Psychological Record 2008;58:569–84.
[53] Cassidy S, Roche B, Hayes SC. A relational frame training intervention to raise Intelligence Quotients: a pilot study. The Psychological Record 2011;61:173–98.
[54] Bonnefond M, Noveck I, Baillet S, Cheylus A, Delpuech C, Bertrand O, et al. What MEG can reveal about inference making: the case of if. . .then sentences. Human Brain Mapping; 2012. http://dx.doi.org/10.1002/hbm.21465.