Kent Academic Repository

Full text document (pdf)

Citation for published version

Black, J. and Williams, David M. and Ferguson, Heather J. (2018) Imagining counterfactual worlds in autism spectrum disorder. Journal of Experimental Psychology: Learning, Memory, and Cognition, 44 (9). pp. 1444-1463. ISSN 0278-7393.

DOI

https://doi.org/10.1037/xlm0000500

Link to record in KAR

http://kar.kent.ac.uk/59626/

Document Version

Author's Accepted Manuscript

Copyright & reuse
Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research
The version in the Kent Academic Repository may differ from the final published version.
Users are advised to check **http://kar.kent.ac.uk** for the status of the paper. **Users should always cite the published version of record.**

Enquiries
For any further enquiries regarding the licence status of this document, please contact:
researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at **http://kar.kent.ac.uk/contact.html**

KAR Kent Academic Repository

University of Kent

Imagining Counterfactual Worlds in Autism Spectrum Disorder

Jo Black

David Williams

Heather J Ferguson

University of Kent, UK

Correspondence to:

Heather Ferguson
School of Psychology
Keynes College
University of Kent
Canterbury
Kent CT2 7NP
England, UK

email: h.ferguson@kent.ac.uk
Tel: +44 (0) 1227 827120 Fax: +44 (0) 1227 827030

Abstract

Two experiments are presented which explore online counterfactual processing in autism spectrum disorder (ASD) using eye-tracking. Participants' eye movements were tracked while they read factual and counterfactual sentences in an anomaly detection task. In Experiment 1, the sentences depicted everyday counterfactual situations (e.g. *If Joanne had remembered her umbrella, her hair would have been dry/wet when she arrived home*). Sentences in Experiment 2 depicted counterfactual versions of real world events (e.g. *If the Titanic had not hit an iceberg, it would have survived/sunk along with all the passengers*). Results from both experiments suggest that counterfactual understanding is undiminished in adults with ASD. In fact, participants with ASD were faster than TD participants to detect anomalies within realistic, discourse-based counterfactuals (Experiment 1). Detection was comparable for TD and ASD groups when understanding could be grounded in knowledge about reality (Experiment 2), though the two groups employed subtly different strategies for responding to and recovering from counterfactual inconsistent words. These data argue against general difficulties in global coherence and complex integration in ASD.

Key words: Counterfactual reasoning, autism, eye-tracking, reading, anomaly detection

Introduction

Our tendency to undo events in the past and imagine 'what might have been' is an example of engaging in counterfactual reasoning (e.g. "if only I had woken up on time, I would have made it to my job interview"), which can help to avoid making similar mistakes in the future (e.g. avoiding oversleeping by setting an extra alarm). Understanding counterfactual statements arguably requires the comprehender to hold two mental representations – one of the implied 'real world' (where I did not wake up on time and missed my job interview) and one of the counterfactual, imagined world (where I *did* wake up on time and made it to my job interview; Mental Model Theory, Byrne & Tasso, 1999; Johnson-Laird & Byrne, 1991). Indeed, on this view, to represent the counterfactual consequent (that I did make it to my job interview), it is necessary to inhibit the representation of the real world (that I did not wake up on time and missed my job interview; see Byrne, 2016).

Although the cognitive basis of counterfactual processing has only recently been investigated empirically, evidence has been largely supportive of the Mental Model Theory (for reviews, see Ferguson, in press, Kulakova & Nieuwland, 2016b). For example, Ferguson and Sanford (2008; see also Ferguson, Sanford, & Leuthold, 2008 Experiment 1) tested counterfactual processing using an eye-tracked anomaly detection reading task. Previous research has established that when readers encounter a word that is anomalous with the sentence context, they make longer fixations on that word and more regressions back, compared to when the word is congruent with the sentence context (e.g. Braze, Shankweiler, Ni, & Palumbo, 2002; Ni, Fodor, Crain, & Shankweiler, 1998; Rayner, Warren, Juhasz, & Liversedge, 2004). This shows that readers are sensitive to discourse-level inconsistencies in text. Ferguson and Sanford (2008) presented participants with factual and counterfactual scenarios across two sentences. The first sentence established the context as either factual (e.g. "If cats are hungry…") or counterfactual (e.g. "If cats were vegetarians…"), and the

second sentence presented a hypothetical event that was either consistent or inconsistent with the context, dependent on a critical word (e.g. feeding a cat a bowl of *fish* vs. *carrots*). Thus, scenarios were fully crossed to describe either real world inconsistent, real world consistent, counterfactual world inconsistent, or counterfactual world consistent, events. Results showed that readers rapidly accommodated the novel counterfactual world, spending longer reading critical words that were inconsistent with the described counterfactual world and taking longer to go past this region (as they tried to make sense of the anomaly) than when critical words were consistent with the counterfactual world. Nonetheless, counterfactual world inconsistencies were detected slower than factual world inconsistencies, because readers had to inhibit real-world knowledge before they could detect counterfactual inconsistencies (as Mental Model Theory would predict). Interestingly, in an event-related brain potentials (ERPs) study, Nieuwland and Martin (2012) observed comparable anomaly detection brain responses within factual and counterfactual contexts, suggesting that factual knowledge does not always disrupt counterfactual processing.

Experiments subsequent to Ferguson and Sanford (2008) have employed diverse methods, such as reaction times, eye-tracking, and ERPs. These studies have found subtle differences in the time course that neurotypical adults process different types of counterfactual sentences, from those sentences that describe everyday events (de Vega, Urrutia, & Riffo, 2007; de Vega & Urrutia, 2012; Ferguson, 2012; Ferguson & Cane, 2015) to those sentences that describe surreal events (Ferguson & Sanford, 2008; Ferguson et al., 2008; Nieuwland, 2013). It is possible that these differences are due to the different types of counterfactual being tested. For example, processing counterfactuals may be less cognitively effortful when they describe plausible alternatives to reality (e.g., counterfactuals about everyday events) compared to counterfactuals that depict novel scenarios (e.g., counterfactuals about non-realistic events). Indeed, in production studies people are more

likely to produce 'close to fact' counterfactuals, and do not tend to create reality violating alternatives (e.g. Lewis, 1973; McMullen & Markman, 2002; Seeleu, Seeleu, Wells, & Windschitl, 1995). Furthermore, the methods used in previous research differ in terms of how language comprehension progresses; eye-tracking allows text to be presented at once, thus reading is more natural. However, ERPs are typically recorded via word-by-word presentation, meaning that readers cannot revisit earlier parts of the text to resolve difficulties. The current study fills a gap in the literature on counterfactual processing by presenting two experiments that examined different types of counterfactual structures, but using a common (eye-tracking) methodology, which allows accurate comparison of effects across experiments.

The current study also fills a gap in the study of developmental disorders, because it investigates counterfactual processing among adults with autism spectrum disorder (ASD), in comparison to age- and IQ-matched neurotypical adults. ASD is a developmental disorder diagnosed on the basis of behavioural difficulties with social-communication, and a restricted and repetitive repertoire of behaviours and interests (American Psychiatric Association, 2013). At the cognitive level, people with ASD manifest impairments in executive functioning (Adams & Jarrold, 2012; Williams & Jarrold, 2013) and the ability to imagine novel scenarios (Lind & Williams, 2012; Lind, Williams, Bowler, & Peel, 2014). It would be important to know whether counterfactual processing is also impaired in this disorder, because counterfactual thoughts are pervasive in adult mental life and play an integral role in higher-level cognition. A recent ERP study on a neurotypical sample has shown that individual differences in social skills play an important part in establishing counterfactual worlds and linking them with real world knowledge online (Kulakova & Nieuwland, 2016a). Yet very little work has explored this directly, using rigorous methods in a clinical sample. What little research that exists has focused on children or adolescents with ASD, and

examined biases in counterfactual thinking, with mixed results. For example, Scott, Baron-Cohen and Leslie (1999) found that, relative to Typically Developing (TD) children, children with ASD were unimpaired in their ability to respond to syllogistic reasoning questions based on counterfactual reasoning (e.g. "All cats bark. Rex is a cat. Does Rex bark?"). However, the ASD group's performance became significantly impaired when they were instructed to create a picture of the story in their mind; in contrast, this instruction enhanced performance among the TD children. This may suggest that counterfactual processing is impaired in ASD only when demands on imagination are high. More recently, Morsanyi and Handley (2012) found impairments in counterfactual reasoning within a fantasy context among adolescents with ASD. Specifically, Morsanyi and Handley found that TD adolescents benefitted from the presentation of contrary-to-fact statements in a fantasy context, whereas adolescents with autism did not. These difficulties in ASD could be the result of an underlying difficulty disengaging from real world knowledge, among other reasons (discussed below). However, drawing firm conclusions from the existing research on counterfactual processing in ASD is not possible, because the literature is limited in several respects.

First, all previous studies have focused on a narrow age-group of children/adolescents. However, ASD is a lifelong developmental disorder and therefore it is not clear whether the observed differences and difficulties in processing counterfactual sentences in childhood persist into adulthood, or whether the development of counterfactual processing ability is merely delayed in ASD. Second, the majority of previous research in this area has focused on the *production* of counterfactual statements (Begeer, Terwogt, Lunenburg, & Stegge, 2009), or has required explicit responses to an investigator's questions (e.g. Leevers & Harris, 2000; Peterson & Bowler, 2000). This is problematic because it combines ambiguous task demands and social interaction with an unfamiliar experimenter, which may lead to deficits among participants with ASD on experimental tasks not because

of difficulties processing counterfactuals per se, but because of extraneous task demands on social interaction and inference. Alternatively, it may be that people with ASD (adults and/or children) have difficulties with *generating* counterfactuals, but not with *comprehending* them. Finally, all previous studies of counterfactual processing in ASD have employed 'offline', response-based, methods that do not assess the real-time processing of counterfactual information, and are therefore not sensitive to subtle preferences or delays in factual/counterfactual processing.

To address these gaps in the literature, the two experiments reported here employed online eye-tracking methods to explore the *comprehension* of counterfactual statements in adults with ASD during natural reading. Eye movements during reading provide an excellent online tool to test real-time processing of counterfactuals, because eye movement data can be examined using a variety of spatially and temporally sensitive measures, allowing identification of exactly where in the sentence difficulties emerge. Crucially, eye-tracking measures have been integral to psycholinguistic research for over 30 years (see Rayner, 1998; 2009), meaning that a great deal is known about different eye movement patterns and the aspects of language processing that they represent.

In Experiment 1, participants with and without ASD were presented with hypothetical (but realistic) counterfactual scenarios (as used in Ferguson, 2012) in an anomaly detection task. Participants read sentences that described everyday counterfactual events (e.g. "If Joanne had remembered her umbrella, her hair would have been dry/wet when she arrived at work"), where the critical word (underlined in the example) was consistent or inconsistent with the counterfactual context. Crucially, these counterfactuals involved a minimal change from reality (i.e. they did not violate real-world knowledge), and comprehension could be supported by participants' everyday experiences, which places relatively low demands on inhibitory control and imagination. Passages depicting factual versions of events (using

"Because") were also included as a baseline measure of contextual integration. Our central predictions in Experiment 1 were as follows. Because the materials map directly onto people's real-life experiences and do not alter existing knowledge of the real-world, we did not expect to observe between-group differences (i.e., a significant effect of Group) in the detection of contextually anomalous critical words within the counterfactual sentences. Rather, we predicted that both TD and ASD participants would detect inconsistent words in factual and counterfactual scenarios to an equivalent extent. However, we predicted that both groups would experience more reading disruption for inconsistent words within factual scenarios than counterfactual scenarios (replicating Ferguson, 2012; Ferguson & Cane, 2015), since the implied factual version of events within a counterfactual context would be activated alongside the counterfactual event. In other words, we expected to observe a context x consistency interaction, reflecting a larger difference in reading times for consistent versus inconsistent conditions within a factual than counterfactual context.

Experiment 1

Method

Participants

Ethical approval for this study was obtained from the University of Kent Psychology Research Ethics Committee. Twenty five adults with ASD (18 males) and 25 age-, sex-, and IQ-matched TD adults took part (see Table 1), all of whom gave written, informed consent before participating. IQ was assessed in all participants using the Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999). Participants in the ASD group had all received formal diagnoses of autistic disorder (n = 6), Asperger's Syndrome (n = 18), or Pervasive Developmental Disorder Not-Otherwise Specified (PDD-NOS; n = 1), according to DSM-V or ICD-10 criteria (American Psychiatric Association, 2013; World Health Organization,

1993), and diagnostic reports were verified by the researcher. Current ASD features were assessed in 22 of the 25 participants in the ASD group by a trained, research-reliable assessor using the Autism Diagnostic Observation Schedule-Generic (ADOS, Lord et al., 2000).

All participants completed the Autism-spectrum Quotient (AQ, Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), a 50-item self-report questionnaire that assesses ASD/ASD-like features. Mean scores for the AQ in each group are shown in Table 1. All participants were over the age of 18 and none had diagnoses of dyslexia or intellectual disability. All participants reported English as their native language. Participants in the TD group did not report any current psychiatric diagnoses.

Insert Table 1 here.

Materials

Anomaly Detection Reading Task

Sentences were modified from Ferguson (2012), as in (1) below. Thirty two experimental items were created, each with four conditions (counterfactual consistent, counterfactual inconsistent, factual consistent, factual inconsistent). Each item consisted of a single sentence, where the antecedent in the first clause introduced a factual ("Because…") or a realistic counterfactual ("If…") scenario, and the consequent in the second clause described an event, where the critical word was either consistent or inconsistent with the preceding context. All scenarios described physical state events (e.g. wet/dry hair) to eliminate any possible confound associated with group differences in mental state processing.

(1) a. If Joanne had remembered her umbrella, her hair would have been <u>dry</u> when she arrived at work. *(counterfactual consistent)*

b. If Joanne had remembered her umbrella, her hair would have been wet when she arrived at work. *(counterfactual inconsistent)*

c. Because Joanne had remembered her umbrella, her hair was dry when she arrived at work. *(factual consistent)*

d. Because Joanne had remembered her umbrella, her hair was wet when she arrived at work. *(factual inconsistent)*

Critical words were matched across conditions for length, written frequency (using the MRC Psycholinguistics Database; Wilson, 1988), and semantic relatedness with the preceding context (using Latent Semantic Analysis; Landauer & Dumais, 1997)[1]. Experimental items were also pre-tested for cloze probability and critical word plausibility. Cloze probability was tested by a total of 48 students from the University of Kent using an online questionnaire platform (Qualtrics). Items were presented one at a time, truncated before the critical word, and participants were instructed to complete the sentence with the first sensible word coming to mind. Cloze probability was computed as the percentage of trials that elicited the intended consistent or inconsistent critical words, and the ANOVA was conducted allowing generalization to items. Participants were significantly less likely to complete sentences with inconsistent (M = 1%) critical words than sentences with consistent critical words (M = 56%; $F(1, 31) = 164.1$, $p < .001$, $\eta_p^2 = .84$), and context (counterfactual *vs.* factual) did not modulate this difference ($F = 1.42$).

[1] There was a non-significant trend for longer words in inconsistent compared to consistent conditions (6.34 *vs.* 5.63), $F(1, 124) = 2.90$, $p = .091$, $\eta_p^2 = .02$, and marginally lower log Kucera and Frances written frequency scores in inconsistent compared to consistent conditions (1.48 *vs.* 1.74), $F(1, 122) = 3.52$, $p = .063$, $\eta_p^2 = .03$. Therefore, word length and word frequency were included as random factors in the eye-tracking analyses on the critical word to control for any such differences across conditions. No differences were found in semantic relatedness, $F = .5$, $p = .5$.

Critical word plausibility was assessed by a different set of 84 students from the University of Kent using Qualtrics. Items were presented in full, and participants were asked to rate the plausibility of each sentence using a five-point sliding scale from -2 (highly implausible) to +2 (highly plausible). The ANOVA was conducted by items. Sentences containing inconsistent critical words were rated as significantly less plausible (M = -1.46) than sentences containing consistent words (M = 1.46; $F(1, 31) = 2046.9$, $p < .001$, $\eta_p^2 = .99$), and this effect was larger within a factual (-1.50 *vs.* 1.56) than counterfactual (-1.42 *vs.* 1.36) context ($F(1, 31) = 10.8$, $p < .005$, $\eta_p^2 = .26$).

Four presentation lists were created, with each list containing thirty-two experimental items, eight in each of the four conditions. Sentences from Experiments 1 and 2 (a total of 64 sentences) were randomly interleaved with each other and an additional 64 filler items. These filler items were all single sentences and did not contain any anomalies, or conditional structures. Half described everyday events (e.g. “Hubert was fanatical about soap operas and never missed an episode”) and half described general knowledge facts (e.g. “New York is a big city, it has lots of taxis and restaurants"). This created a random order such that each participant only saw each experimental sentence once, in one of the four conditions. Participants were randomly assigned to read each list. All sentences were presented over two lines of text, with one blank line in between. The critical word was never the first or last word in a line, and appeared in the same position across all conditions. Comprehension questions followed half of the experimental (i.e., 16 each for Experiments 1 and 2) and filler (i.e. 32) trials.

Inhibitory Control Task

The Stroop task (Stroop, 1935) was used as a measure of individuals’ inhibitory control. It consisted of 50 incongruent trials, 50 congruent trials, and 50 non-colour word neutral trials.

Responses were recorded using a five-button serial response box. Interference scores were calculated by subtracting response times (RTs) on the neutral trials from incongruent trials. RTs under 200ms or greater than 2.5 standard deviations from each participants' mean RT were removed as outliers. Results showed marginally impaired inhibitory control in the ASD versus TD group (91 *vs.* 61msec; see Table 1).

Imagination Task

The scene construction task described by Hassabis, Kumaran and Maguire (2007), and used by Lind et al. (2014) among individuals with ASD, was used as a measure of imagination. Participants were asked to imagine three vivid scenes in their mind's eye and describe them out loud. Participants completed a questionnaire after each description to report their experiences in the task, including difficulty, vividness of the scene, etc. The recordings of the descriptions were transcribed and blind coded by the first author for the number of people and objects mentioned; sensory descriptions; spatial descriptions; and thoughts, emotions and actions described. The third author also coded 10% of the transcriptions to check reliability of coding and inter-rater reliability was found to be good with intraclass correlation of .89 ($F(14, 14) = 11.42$, $p < .001$). Scores from the descriptions and questionnaires were then combined as described in Hassabis et al. to provide an index of imagination (an "experiential index score"), with higher scores indicating greater imagination ability. Results showed significantly impaired imagination in the ASD versus TD group (34.7 *vs.* 40.8; see Table 1).

Procedure

For the anomaly detection reading task, participants' gaze location and movement from their dominant eye was recorded using an EyeLink 1000 Plus eye-tracker (viewing was binocular). All sentences were presented in size 20 Arial font style on a VDU screen, 60cm from the

participants' eyes. Prior to the experiment, the procedure was explained and participants were instructed to read at their normal rate. Participants were seated at the eye-tracker and a chin rest was used to stabilize participants' head position. The eye-tracker was calibrated using a nine point procedure. Before each sentence, participants performed a drift correction using a central fixation point. Once this calibration check was completed accurately, the experimenter advanced the screen to display the next item. Adjustments to the calibration were made whenever necessary. After reading each sentence, participants clicked a button on the mouse that either led to the presentation of a comprehension question (after 50% of trials) or the next trial.

The entire testing procedure took approximately two hours to complete, and participants always completed the eye-tracking task first. The AQ, WASI, Imagination, and Stroop tasks were subsequently completed in a randomised order. Participants with ASD returned on a separate occasion to take part in the ADOS. Testing took place in a quiet research lab at the University of Kent.

Results

Methods of Analysis The experimental passages were divided into three regions for analysis. The consistent or inconsistent critical word was always presented mid-sentence:

	Pre-critical	Critical	Post-critical
"If Joanne had remembered her umbrella,	her hair would have been	wet/ dry	when she arrived at work."

An automatic procedure pooled fixations shorter than 80ms with larger adjacent fixations, excluded fixations shorter than 40ms that were not within three characters of another fixation and truncated fixations longer than 1200ms. A baseline comparison of global

eye movement data across all trials revealed that TD and ASD groups did not differ in average fixation duration (208ms *vs.* 213ms, p=.53), number of fixations per sentence (19.0 *vs.* 19.9, p=.51), or saccade amplitude (3.02° *vs.* 3.0°, p=.83), thus we can be confident that there were no sampling differences between the two groups.

Two early measures of language processing are reported. First-pass reading time is the sum of the duration of fixations made on first entering a region of text until an eye-movement exits the region to either the left or right. Regression path is the sum duration of all fixations from first entering a region on the left, to exiting the region on the right (this includes any initial regressions back to previous regions). These early measures provide an indication of the difficulty experienced when participants initially process a region of text. We also analysed one later measure. Total reading time is the sum duration of all fixations made within a region and provides an indication of the overall amount of time spent processing text in that region. Mean reading times were computed for each measure in each of the three regions for all four conditions in each group (see Table 2). For transparency, the full datasets for both experiments are available on the Open Science Framework web pages (see https://osf.io/kw8sx/?view_only=2e57d7ecf7134eababf2c19ffddf20eb).

Insert Table 2 here.

The eye movement data was log-transformed and analysed separately for each region and measure using the lmer function in the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) using R (version 3.3.2, R Core Team, 2016). Each model included fixed effects of Group, Context, and Consistency, with the two levels of each fixed effect deviation coded (-.5 *vs.* .5) to ensure they could be directly compared. In addition, given group differences in imagination and inhibitory control (marginal) these values were centred and entered into the

model as continuous fixed effects to ensure that any effects of group were a result of ASD. Models included the maximal random effects structure, including random effects for participants and items, crossed random slopes for Group, Context and Consistency and random slopes for imagination and inhibitory control within items, and crossed random slopes for Context and Consistency within participants (as suggested by Barr, Levy, Scheepers, & Tily, 2013). In addition, analyses on the critical word included continuous predictors of word length and word frequency. Random effects were only removed where they lead to non-convergence due to overparameterization. Statistical effects from these models are shown in Table 3.

Insert Table 3 here.

Insert figure 1 here.

First-Pass Reading Time

As expected, given the difference in length between counterfactual and factual contexts (e.g. "her hair would have been" *vs.* "her hair was"), a significant main effect of context was found in the pre-critical region, with longer first-pass reading times for counterfactual compared to factual sentences. In addition, first-pass reading times on the critical word were modulated by individual differences in Imagination, showing that participants with higher imagination scores had longer first-pass reading times in these regions.

The Consistency x Group interaction was significant on the critical word region. Follow-up analyses revealed that, for TD participants, there was no significant difference in first-pass reading times whether the critical word was consistent or inconsistent with the sentence context, *Est.* = .016, *SE* = .012, $t = 1.28$, $p = .2$. However, the ASD group showed marginally longer first-pass reading times when the critical word was inconsistent with the

sentence context relative to when it was consistent, *Est.* = -.025, *SE* = .014, *t* = -1.86, *p* = .063.

Finally, there was a significant Context x Consistency interaction in the post-critical region. Follow-up analyses showed that within a factual context first-pass reading times were marginally longer following an inconsistent critical word than a consistent critical word, *Est.* = -.032, *SE* = .018, *t* = -1.8, *p* = .08. However, within a counterfactual context first-pass reading times did not differ between consistent and inconsistent conditions, *Est.* = .027, *SE* = .02, *t* = 1.32, *p* = .19. Following an inconsistent critical word, first-pass reading times were marginally longer in a factual than counterfactual context, *Est.* = -.041, *SE* = .021, *t* = 1.97, *p* = .058, but did not differ following a consistent critical word, *Est.* = -.021, *SE* = .017, *t* = -1.24, *p* = .22.

Regression Path Reading Time

Similar to the first-pass reading time measure, a significant effect of Context in the pre-critical region showed that participants took longer to move past this region for counterfactual sentences compared to factual, due to length differences between context conditions. Individual differences in Imagination had a significant effect on regression path reading times in the pre-critical and critical regions, reflecting the same positive correlation between imagination scores and reading time as seen in the first-pass reading time measure.

Group had a significant effect on reading times in the critical region, due to significantly longer regression path times in the ASD versus TD group. Moreover, Group interacted significantly with Consistency in this critical region. Similar to the pattern seen in first-pass reading times, no difference was found in regression path times for TD participants whether the critical word was consistent or inconsistent with the sentence context, *Est.* = .009, *SE* = .022, *t* = .41, *p* = .68. However, the ASD group showed significantly longer

regression path times on the critical word when it was inconsistent compared to consistent with the context, *Est.* = -.064, *SE* = .025, *t* = -2.57, $p < .05$.

Finally, a significant effect of Consistency in the post-critical region revealed longer regression path times for inconsistent versus consistent words. A simultaneous interaction between Context and Consistency revealed that this effect of consistency held within both contexts, though was larger following a factual, *Est.* = -.184, *SE* = .025, *t* = -7.22, $p < .001$, than counterfactual context, *Est.* = -.115, *SE* = .033, *t* = -3.53, $p < .001$. However, participants had longer regression path reading times on the post-critical region following a consistent critical word within a counterfactual context than a factual context, *Est.* = -.072, *SE* = .029, *t* = -2.45, $p < .05$. Regression path reading times did not differ between factual and counterfactual contexts following an inconsistent critical word, *Est.* = -.001, *SE* = .03, *t* = .04, $p = .97$.

Total Reading Time

Statistical analyses revealed a significant effect of Consistency in the total time spent reading each region, indicating that participants spent longer reading the sentences when the critical word was inconsistent compared to when it was consistent with the wider context.

As in other measures, the effect of Context was significant in the pre-critical region, showing that participants spent longer reading this region when sentences were counterfactual compared to factual, due to length differences between context conditions. Total reading times on the critical word were also marginally longer in the ASD group than the TD group.

Interestingly, as in first-pass and regression path reading times, the interaction between Context and Consistency was significant on this measure in the post-critical region. Follow-up analyses showed that the overall effect of consistency (inconsistent > consistent)

was only significant within a factual context, *Est.* = -.103, *SE* = .02, *t* = -5.13, *p* < .001, and not within a counterfactual context, *Est.* = -.021, *SE* = .022, *t* = -.98, *p* = .34. Participants did not differ in total time reading this region following a consistent critical word within a counterfactual or factual context, *Est.* = -.025, *SE* = .019, *t* = -1.33, *p* = .19, but spent more time reading when this post-critical region followed an inconsistent word within a factual than a counterfactual context, *Est.* = .06, *SE* = .02, *t* = 2.99, *p* < .01.

Summary

Several important findings were revealed in Experiment 1. First, the consistency effect (inconsistent > consistent) on total reading times in all regions showed that overall, participants accommodated the hypothetical counterfactual world; they detected counterfactual inconsistencies in a comparable timecourse to factual inconsistencies. Second, factual inconsistencies elicited greater disruption to reading compared to counterfactual inconsistencies, as reflected by a larger consistency effect on the post-critical region in first-pass reading times, regression path reading times, and total reading times. Crucially, these patterns were manifest in *both* groups. Thus, when counterfactuals involve a minimal change from reality, counterfactual language processing is comparable in TD adults and adults with ASD. Nevertheless, some differences in processing emerged between the two groups. First, the ASD participants showed longer overall reading times on the critical word compared to the TD participants. More importantly, participants with ASD were *quicker* than TD participants at spotting anomalies across both types of contexts; they had longer first-pass and regression path reading times on inconsistent critical words than consistent critical words, a difference not found among the TD participants. We will return to this point in the General Discussion.

In sum, Experiment 1 provided evidence that high-functioning adults with ASD are able to process counterfactual sentences. However, these were simple counterfactual sentences where the implied factual world was based purely on a temporary discourse, resulting in a minimal change to reality. It may be that people with ASD have more difficulty with counterfactual reasoning when a more substantial change to reality is required. To test this in Experiment 2, participants read sentences that described hypothetical alternatives to known historical events (e.g. "If the Titanic had not hit an iceberg, it would have survived/sunk along with all the passengers"; adapted from Nieuwland & Martin, 2012), where the critical word was consistent or inconsistent with the counterfactual context (but inconsistent/consistent with factual knowledge). Crucially, these counterfactuals required readers to suspend their knowledge of reality and imagine a novel version of the world. Thus, our central prediction in Experiment 2 was that participants with ASD would experience greater interference from their historical knowledge, and that this would disrupt or delay anomaly detection effects for an inconsistent critical word within a counterfactual context. This would be demonstrated by a reduced or delayed inconsistency effect within a counterfactual context for participants with ASD, relative to the TD participants.

Experiment 2

Method

Experiment 2 was run concurrently with Experiment 1. Therefore methodology was as described in Experiment 1, bar the materials which are detailed below.

Materials

Sentences described known historical events, and were translated and modified from Nieuwland and Martin (2012). Thirty two experimental items were created in the same four conditions as Experiment 1, as in (2) below.

(2) a. If the Titanic had not hit an iceberg, it would have survived along with all the passengers. *(counterfactual consistent)*

b. If the Titanic had not hit an iceberg, it would have sunk along with all the passengers *(counterfactual inconsistent)*

c. Because the Titanic had hit an iceberg, it had sunk along with all the passengers *(factual consistent)*

d. Because the Titanic had hit an iceberg, it had survived along with all the passengers *(factual inconsistent)*

As in Experiment 1, critical words were matched across conditions for length, frequency, and semantic relatedness with the preceding context, and no significant differences were found (all Fs < 1.8, ps > .18)[2]. Cloze probability of experimental items was rated by 44 undergraduate students at the University of Kent, and sentence plausibility was rated by 42 different undergraduate students at the University of Kent. ANOVAs were conducted allowing generalization to items. As expected, cloze probability of consistent critical words (M = 55%) was significantly higher than inconsistent critical words (M = 1.2%, $F(1, 31) = 234.4$, $p < .001$, $\eta_p^2 = .88$), and context did not modulate this difference (F = 0.19). Plausibility ratings were significantly higher for sentences where the critical word was

[2] Although there were no significant differences in word length or frequency across conditions, these variables were included as random factors in the analysis of the critical word region for consistency with Experiment 1.

consistent (M = 1.46) than inconsistent (M = -1.63; $F(1, 31) = 1177.9$, $p < .001$, $\eta_p^2 = .97$), and this effect was larger within a factual (1.58 *vs.* -1.84) than counterfactual context (1.35 *vs.* -1.43), $F(1, 31) = 18.9$, $p < .001$, $\eta_p^2 = .38$.

General Knowledge Assessment

Participants' general knowledge for the historical events described in Experiment 2 was assessed using a questionnaire. One question for each experimental item (e.g. "What happened in New York on September 11th 2001"?) was presented on a computer, and participants were instructed to type in a short answer, or write "I don't know". Participants were highly accurate, averaging 95% of questions answered correctly (SD = 6.5%), and there was no significant difference in accuracy between the ASD (M = 93%, SD = 5.59) and TD groups (M = 94%, SD = 7.48), $t(48) = -0.49$, $p = .629$, $d = 0.15$.

Results

Methods of Analysis As in Experiment 1, stimuli were divided into three regions for analysis, with the consistent/inconsistent word always appearing mid-sentence:

	Pre-critical	Critical	Post-critical
"If the Titanic had not hit an iceberg,	it would have	survived/ sunk	along with all the passengers."

Eye movement data was prepared and analysed as in Experiment 1. The resulting reading items for each measure, region and condition are shown in Table 4, and statistical effects are shown in Table 5.

Insert Table 4 here.

Insert Table 5 here.

First-pass Reading Time

A significant effect of Context was found on first-pass reading time in all three regions, reflecting longer first-pass reading times for counterfactual compared to factual sentences. In the pre-critical region this is likely due to differences in length between counterfactual and factual contexts, whereas for the critical and post-critical regions this is likely to reflect additional early processing on the critical word within a counterfactual context.

A significant Context × Consistency × Group interaction was found on the critical word region. Within a factual context, neither group showed a significant difference between consistent and inconsistent critical words (TD: *Est.* = -.015, *SE* = .017, *t* = -.88, *p* = .38; ASD: *Est.* = -.032, *SE* = .025, *t* = -1.27, *p* = .21). More importantly, within a counterfactual context TD participants showed a significant consistency effect on the critical word (inconsistent > consistent), *Est.* = -.046, *SE* = .017, *t* = -2.67, $p < .05$, whereas participants with ASD did not, *Est.* = .008, *SE* = .025, *t* = .32, *p* = .75. In addition, first-pass reading times on the critical word were modulated by individual differences in Imagination, reflecting longer first-pass reading times in participants with higher imagination scores.

Insert Figure 2 here

Regression Path Reading Time

Similar to first-pass reading times, there was a significant effect of Context on regression path reading times in the pre-critical and post-critical regions. This showed that participants took longer to move past these regions for counterfactual compared to factual sentences, and is likely to reflect length differences in the pre-critical region. The context effect in the post-

critical region, however, may additionally indicate slower integration of the counterfactual premise.

More importantly, a significant effect of Consistency was found on the critical word region and the post-critical region. Participants took longer to move past these regions when the word was inconsistent with the preceding context compared to when it was consistent. In addition, regression path reading times on the critical word showed a positive correlation with participants' imagination scores.

Total Reading Time

A significant effect of Consistency was found in all three regions, with participants spending longer reading each region when the critical word was inconsistent compared to when it was consistent. In addition, a significant effect of Context was found in the pre-critical and critical regions, showing that participants spent longer overall reading the antecedent clause and critical word when it described a counterfactual version of the world as opposed to factual information.

There was a significant Consistency x Group interaction in the pre-critical region. TD participants spent significantly longer reading this region when the sentence contained an inconsistent critical word compared to a consistent critical word, *Est.* = -.084, *SE* = .018, t = -4.7, $p < .001$, but this difference missed significance in participants with ASD, *Est.* = -.031, *SE* = .018, t = -1.74, p = .09.

Total reading times on the critical word were significantly longer in the ASD group than the TD group, and people with higher imagination scores had longer total reading times on the critical and post-critical regions.

Summary

In Experiment 2, readers successfully used the counterfactual context to process incoming events according to an alternative version of the world (evidenced by the consistency effect on total reading times in all regions). Interestingly, anomaly detection effects were comparable between factual and counterfactual contexts, suggesting that readers did not experience interference from the 'real world' when processing the counterfactual sentences. Indeed, detection of the inconsistency emerged early in Experiment 2, with effects appearing on the critical word itself (in longer regression path reading times), suggesting that participants rapidly noticed the contextual inconsistency and responded by re-reading prior text. Similar to Experiment 1, ASD participants showed longer total reading times on the critical word compared to the TD participants. Most importantly, adults with ASD were able to process non-real counterfactual events in an equivalent timeframe to TD participants, although some subtle between-group differences emerged in the way that counterfactual inconsistent words were responded to and recovered from (we discuss these subtle differences in the General Discussion).

General Discussion

In two eye-tracking reading experiments, we sought to understand how individuals with ASD interpret counterfactual events online. In Experiment 1, items described counterfactual situations that modified the outcome of everyday events (e.g. *If Joanne had remembered her umbrella, her hair would have been dry/wet*). In Experiment 2, items set up hypothetical alternatives to known historical events (e.g. *If the Titanic had not hit an iceberg it would have survived/sunk*). While the counterfactual events in both experiments were logically true, sentences in Experiment 1 incurred a minimal change from reality (i.e. greater overlap between hypothetical and counterfactual worlds), but understanding in Experiment 2 required readers to suspend their knowledge of reality and imagine a novel version of the world.

Before discussing how the results inform our understanding of counterfactual processing in ASD, it is important to consider how the results relate to previous studies of counterfactual processing in neurotypical individuals. Overall, the patterns of performance shown by TD participants in each experiment replicate those from previous research in TD adults. In Experiment 1, where judgements about the consistency of critical words were based on constraints from the local discourse, TD readers elicited anomaly detection responses from the post-critical region, with increased regression path reading times on this region, and longer total reading times across the entire sentence. In addition, readers experienced some interference from the alternative model of the world, showing a reduced (or absent) consistency effect in a counterfactual, relative to factual, context on the post-critical region (across all three reading measures). In contrast, in Experiment 2, where counterfactual/factual worlds were more clearly distinct and understanding could be grounded in knowledge about reality, anomalies disrupted reading from the critical word itself, with longer regression path reading times on this region, and increased total reading times across the entire sentence. Here, readers did not show signs of conflict with the alternative world; the consistency effect was comparable within factual and counterfactual contexts.

Therefore, in the TD group, anomaly detection effects appeared one region earlier in the sentence in Experiment 2 than in Experiment 1, and context (factual/counterfactual) did not influence anomaly detection in Experiment 2 as it did in Experiment 1. Anomaly detection on the critical word itself in Experiment 2 is likely due to the stimuli, which tapped real world knowledge and could be easily retrieved and updated from long-term memory. In contrast, stimuli in Experiment 1 manipulated discourse based anomalies, so required readers to generate new temporary representations of both factual and counterfactual worlds, which competed with each other in memory, and delayed anomaly detection responses to the post-critical region. Moreover, the timing with which counterfactual inconsistencies were detected

between the two experiments replicates the timing effects seen in previous research (Ferguson, 2012; Ferguson & Cane, 2015; Nieuwland & Martin, 2012). This shows that the different counterfactual constructions being tested (and the specific constraints they incur), rather than methodological differences between previous studies (i.e. eye-tracking *versus* ERPs), can account for differences in the timecourse of counterfactual understanding [see Ferguson & Cane (2015) for further discussion of paradigm sensitivity].

Regarding the key issue of counterfactual understanding in ASD, results from both experiments showed that adults with ASD do not experience gross difficulties understanding counterfactual sentences online; effects of consistency were clear in both groups. This extends findings from a recent eye-tracking reading study showing comparable lexical access in ASD and TD individuals (Howard, Liversedge, & Benson, 2017). Intriguingly, however, between-group differences emerged in the timecourse with which the counterfactuals were processed. In Experiment 1 participants with ASD detected factual and counterfactual inconsistencies even earlier than the TD group, with anomaly detection responses emerging in early reading measures on the critical word itself (on first-pass and regression path reading times). In Experiment 2, both TD and ASD readers showed anomaly detection responses upon encountering the critical word (on regression path time), though subtle differences emerged in the way the two groups responded to and recovered from these inconsistencies. Specifically, the TD group first spent longer first-pass reading inconsistent versus consistent critical words in a counterfactual context before regressing back to check the preceding context (reflected in increased regression path times on the critical word and increased total reading times on the pre-critical region). In contrast, participants with ASD did not show an inconsistency effect on first-pass reading times, but they rapidly regressed back to re-read the critical word (reflected in increased total reading times on the critical word) and earlier parts of the sentence. The fact that the TD group, but not the ASD group, showed an anomaly

effect on the 'early' first-pass pass reading time measure may indicate a subtle timing advantage for anomaly detection in the TD group in Experiment 2, though both groups were clearly sensitive to the inconsistency upon first encountering the critical word. Nevertheless, both groups exhibited the same context modulations of anomaly detection (i.e. weaker effects of consistency for counterfactual than factual contexts in Experiment 1 but not Experiment 2), suggesting that TD and ASD participants experienced similar interference from reality.

The fact that in both experiments counterfactual inconsistencies were detected by participants with ASD in a comparable, or even enhanced, timecourse as TD participants is a novel finding, and contrasts with previous research that has shown impaired counterfactual reasoning in children with ASD (e.g. Grant, Riggs, & Boucher, 2004). One potential explanation for this discrepancy is the developmental nature of the disorder; difficulties in childhood often abate in adulthood, with many symptoms of ASD improving throughout the lifespan (Shattuck et al., 2007). Thus, it is possible that individuals with ASD simply experience a protracted period of developing fully functional counterfactual thinking, which extends beyond that of TD individuals (see Beck & Riggs, 2013; 2014). Alternatively, the explicit, response-based methodologies employed in previous studies may have driven response biases and errors (see Rubio-Fernández (2013) for a similar argument for Theory of Mind processing). For example, people with ASD are more susceptible to demand characteristics and so the syllogistic reasoning questions used in several previous studies (e.g. Scott, Baron-Cohen, & Leslie, 1999; Leevers & Harris, 2000; Morsanyi & Handley, 2012) introduced ambiguity about what responses were desired by the experimenter. In contrast, the passive reading task and online eye movement measures employed in the present study meant that counterfactual understanding was tested in real-time under relatively natural constraints, while limiting response biases. Interestingly, these comparable anomaly detection responses were found despite the ASD group scoring significantly lower than the TD group on the

imagination task, and marginally lower on our measure of inhibitory control, and when these measures were controlled for in our statistical analyses. These two cognitive skills are thought to be highly important for successful counterfactual reasoning. This suggests that although the adult ASD sample tested here were compromised on these skills relative to their TD peers (see also Adams & Jarrold, 2012; Williams & Jarrold, 2013), they were sufficiently able to inhibit reality and imagine an appropriate alternative to interpret the logically valid counterfactuals tested here.

What then do these results mean for theories of ASD that implicate impaired contextual integration? The weak central coherence theory links a local processing bias with a lack of global coherence in ASD (Frith, 1989; Frith & Happé, 1994; Happé & Frith, 2006), which would predict that individuals with ASD would be impaired at integrating information within and between sentences (and thus miss anomalies). Another cognitive account, the theory of complex information processing (CIP; Minshew, Goldstein, & Siegel, 1997; Minshew & Goldstein, 1998; Minshew, Williams, & McFadden, 2008) proposes that ASD is the result of an impairment in the 'complex' processes that facilitate information integration or the use of top-down knowledge. Our results suggest that participants with ASD spontaneously maintained global coherence across the entire sentence, showing appropriate anomaly detection responses based on discourse in Experiment 1, and based on a counterfactual premise that modified reality in Experiment 2. Indeed, the fact that discourse-based anomalies in Experiment 1 were detected even faster by participants with ASD than TD participants, suggests *superior* integration skills. These results are consistent with the findings from Au-Yeung, Kaakinen, Liversedge, and Benson's (2015) eye-tracking reading study, where participants with ASD showed no difficulty utilizing the global discourse context to interpret a sentence as ironic (or not). However, it is important to note that the distinction between local/global and simple/complex contexts is currently undefined in the

WCC and CIP theories, making it possible that adults with ASD are able to maintain global coherence within a single sentence. Further studies are needed to test how people with ASD interpret counterfactual events when the counterfactual world needs to be maintained *over time* (i.e. over multiple clauses and sentences), or when the discourse switches between factual and counterfactual versions of the world (as in Ferguson & Cane, 2015).

In conclusion, the two experiments reported here suggest that counterfactual understanding is undiminished in adults with ASD. Participants with ASD rapidly accommodated the hypothetical counterfactual world, and successfully detected contextually inconsistent events within factual and counterfactual contexts, in a comparable timecourse to TD participants. In fact, participants with ASD were faster than TD participants to detect anomalies within realistic, discourse-based counterfactuals, and detection was comparable when understanding could be grounded in knowledge about reality. Both groups experienced interference from the the alternative 'world' when overlap between worlds was high (as in Experiment 1), but not when the two worlds were more clearly distinct (as in Experiment 2). These data demonstrate that people with ASD employ subtly different processing strategies for counterfactual thinking, and argue against general difficulties in global coherence and complex integration.

Acknowledgements

This work was carried out with the support of a grant to HF and DW from the Leverhulme Trust (Ref: RPG-2014-298). The authors would also like to thank Mante Nieuwland for sharing his Spanish counterfactual materials from Nieuwland and Martin (2012).

References

Adams, N. C., & Jarrold, C. (2012). Inhibition in autism: Children with autism have difficulty inhibiting irrelevant distractors but not prepotent responses. *Journal of Autism and Developmental Disorders, 42(6)*, 1052-1063.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*, Washignton, DC.

Au-Yeung, S. K., Kaakinen, J. K., Liversedge, S. P., & Benson, V. (2015). Processing of written irony in autism spectrum disorder: An eye-movement study. *Autism Research, 8(6)*, 749-760.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31(1)*, 5-17.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68(3)*, 255-278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67(1)*.

Beck, S. R., & Riggs, K. J. (2013). Counterfactuals and reality. In M. Taylor (Ed.), *The Oxford Handbook of the Development of the Imagination*. New York, USA: OUP.

Beck, S. R., & Riggs, K. J. (2014). Developing thoughts about what might have been. *Child Development Perspectives, 8(3)*, 175-179.

Begeer, S., Terwogt, M. M., Lunenburg, P., & Stegge, H. (2009). Brief report: Additive and subtractive counterfactual reasoning of children with high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39(11)*, 1593-1597.

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research, 31(1)*, 25-44.

Byrne, R. M. (2016). Counterfactual thought. *Annual Review of Psychology, 67*, 135-157.

Byrne, R. M., & Tasso, A. (1999). Deductive reasoning with factual, possible, and counterfactual conditionals. *Memory & Cognition, 27(4)*, 726-740.

de Vega, M., & Urrutia, M. (2012). Discourse updating after reading a counterfactual event. *Psicologica: International Journal of Methodology and Experimental Psychology, 33(2)*, 157-173.

de Vega, M., Urrutia, M., & Riffo, B. (2007). Canceling updating in the comprehension of counterfactuals embedded in narratives. *Memory & Cognition, 35(6)*, 1410-1421.

Ferguson, H. J. (in press). Counterfactuals. In Cummings, C. & Katsos, N. (Eds.), *Handbook of Experimental Semantics and Pragmatics*, Oxford University Press.

Ferguson, H. J. (2012). Eye movements reveal rapid concurrent access to factual and counterfactual interpretations of the world. *The Quarterly Journal of Experimental Psychology, 65(5)*, 939-961.

Ferguson, H. J., & Cane, J. E. (2015). Examining the cognitive costs of counterfactual language comprehension: Evidence from ERPs. *Brain Research, 1622*, 252-269.

Ferguson, H. J., & Sanford, A. J. (2008). Anomalies in real and counterfactual worlds: An eye-movement investigation. *Journal of Memory and Language, 58(3)*, 609-626.

Ferguson, H. J., Sanford, A. J., & Leuthold, H. (2008). Eye-movements and ERPs reveal the time course of processing negation and remitting counterfactual worlds. *Brain Research, 1236*, 113-125.

Frith, U. (1989). *Autism: Explaining the enigma*. Oxford, UK: Basil Blackwell.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition, 50(1-3)*, 115-132.

Grant, C. M., Riggs, K. J., & Boucher, J. (2004). Counterfactual and mental state reasoning in children with autism. *Journal of Autism and Developmental Disorders, 34(2)*, 177-188.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36(1)*, 5-25.

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *Journal of Neuroscience, 27(52)*, 14365-14374.

Howard, P. L., Liversedge, S. P., & Benson, V. (2017). Benchmark eye movement effects during natural reading in autism spectrum disorder. *Journal of Experimental Psychology.Learning, Memory, and Cognition, 43(1), 109-127.*

Johnson-Laird, P. N., & Byrne, R. M. (1991). *Deduction.* Lawrence Erlbaum Associates, Inc.

Kulakova, E., Nieuwland, M.S. (2016a). Pragmatic skills predict online counterfactual comprehension: Evidence from the N400. *Cognitive, Affective, & Behavioral Neuroscience, 16*, 814-824.

Kulakova, E., & Nieuwland, M. S. (2016b). Understanding counterfactuality: A review of experimental evidence for the dual meaning of counterfactuals. *Language and Linguistics Compass, 10(2)*, 49-65.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104(2)*, 211.

Leevers, H. J., & Harris, P. L. (2000). Counterfactual syllogistic reasoning in normal 4-year-olds, children with learning disabilities, and children with autism. *Journal of Experimental Child Psychology, 76(1)*, 64-87.

Lewis, D. (1973). *Counterfactuals*. Blackwell.

Lind, S.E., Williams, D.M., Bowler, D.M., & Peel, A. (2014). Episodic memory and episodic future thinking impairments in high-functioning autism spectrum disorder: An underlying difficulty with scene construction or self-projection? *Neuropsychology, 28*, 55-67.

Lind, S.E. & Williams, D.M. (2012). The association between past and future oriented thinking: Evidence from autism spectrum disorder. *Learning and Motivation, 43(4)*, 231-240.

Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The autism diagnostic observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30(3)*, 205-223.

McMullen, M.N. & Markman, K.D. (2002). Affective impact of close counterfactuals: Implications of possible futures for possible pasts. *Journal of Experimental Social Psychology, 38*, 64-70.

Minshew, N. J., & Goldstein, G. (1998). Autism as a disorder of complex information processing. *Mental Retardation and Developmental Disabilities Research Reviews, 4*(2), 129-136.

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3(4)*, 303-316.

Minshew, N. J., Williams, D. L., & McFadden, K. (2008). Information processing, neural connectivity, and neuronal organization. In A. Zimmerman (Ed.), *Autism: Current theories and evidence* (pp. 381-405). Totowa, NJ: Humana Press.

Morsanyi, K. & Handley, S. J. (2012). Reasoning on the basis of fantasy content: Two studies with high-functioning autistic adolescents. *Journal of Autism and Developmental Disorders, 42(11)*, 2297-2311.

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research, 27(5)*, 515-539.

Nieuwland, M. S. (2013). "If a lion could speak…": Online sensitivity to propositional truth-value of unrealistic counterfactual sentences. *Journal of Memory and Language, 68(1)*, 54-67.

Nieuwland, M. S., & Martin, A. E. (2012). If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension. *Cognition, 122(1)*, 102-109.

Peterson, D. M., & Bowler, D. M. (2000). Counterfactual reasoning and false belief understanding in children with autism. *Autism, 4(4)*, 391-405.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124(3)*, 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62(8)*, 1457-1506.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30(6)*, 1290-1301.

Rubio-Fernández, P. (2013). Perspective tracking in progress: Do not disturb. *Cognition*, *129*, 264-272.

Scott, F. J., Baron-Cohen, S., & Leslie, A. (1999). 'If pigs could fly': A test of counterfactual reasoning and pretence in children with autism. *British Journal of Developmental Psychology, 17(3)*, 349-362.

Seeleu, E.P., Seeleu S.M., Wells, G.L., & Windschitl, P.D. (1995). Counterfactual constraints. In: *What might have been: The social psychology of counterfactual thinking*, ed. N. J. Roese & J. M. Olson. Erlbaum.

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., . . . Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37(9)*, 1735-1747.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18(6)*, 643-662.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence (WASI).* San Antonio, TX: Psychological Corporation.

Williams, D.M., & Jarrold, C. (2013). Assessing planning and set-shifting abilities in autism: Are experimenter-administered and computerised versions of tasks equivalent? *Autism Research*, 6, 461–467.

Wilson, M. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers, 20(1)*, 6-11.

World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva : World Health Organization.

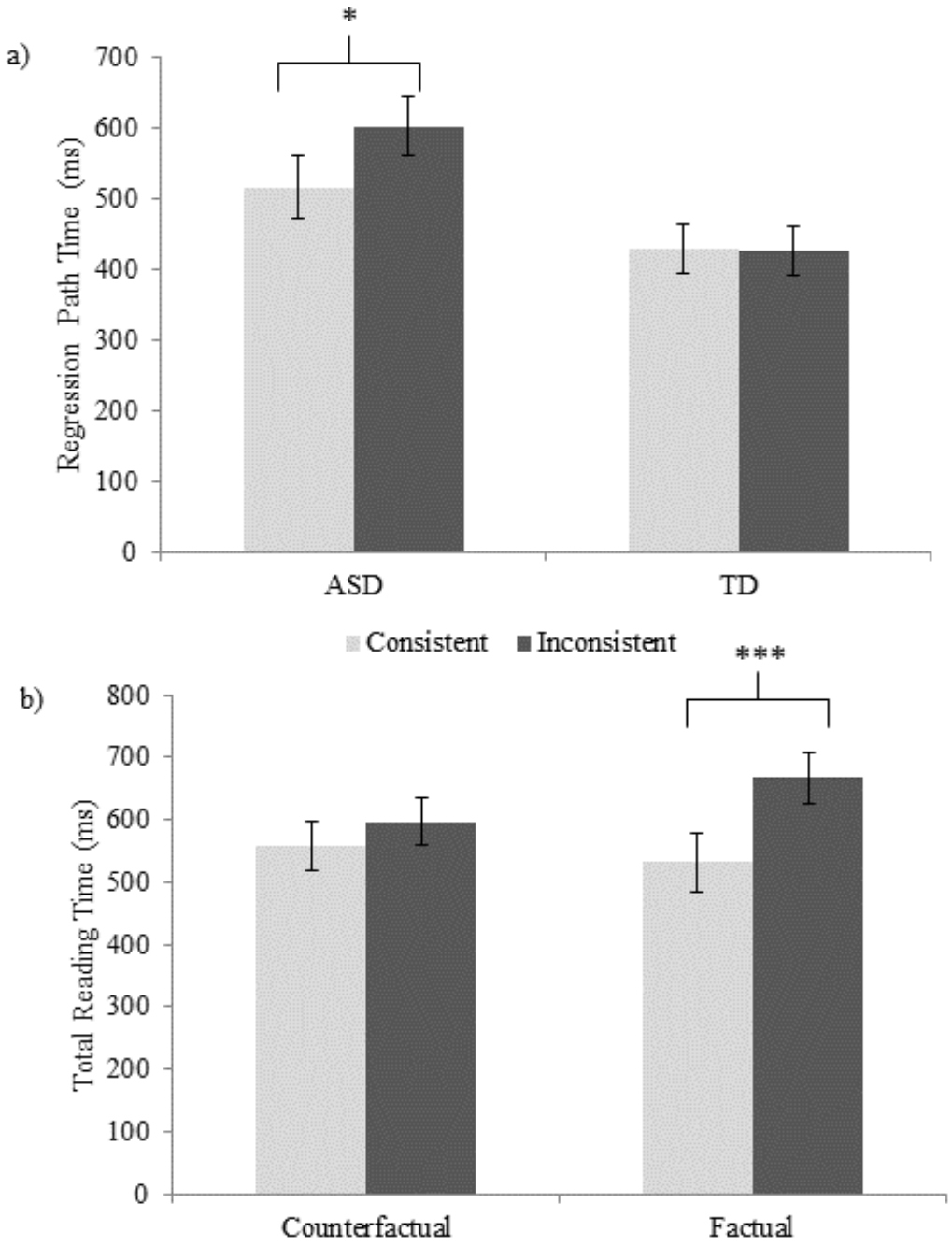

Figure 1. a) Regression Path Time for Region 3, Experiment 1, for ASD and TD participants in consistent and inconsistent conditions. b) Total Reading Time (ms) for Region 4, Experiment 1, after consistent and inconsistent critical words in factual and counterfactual conditions.

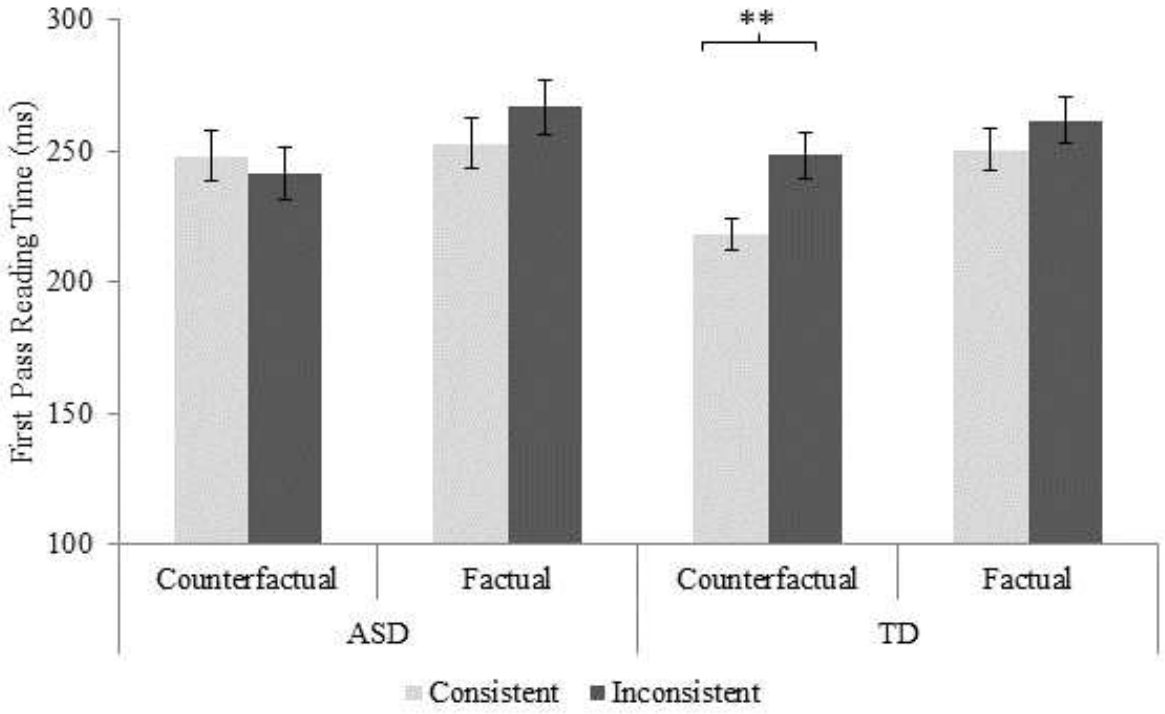

Figure 2. First pass reading times on Region 3, Experiment 2, for ASD and TD participants in all conditions.

Table 1. *Demographic information for ASD and TD groups (M (SD)) with comparison statistics.*

	ASD (n = 25)	TD (n = 25)	t	p	Cohen's d
Sex (m:f)	18:7	18:7			
Age (years)	25.6 (7.96)	26.4 (8.29)	-0.30	.763	.09
Verbal IQ	105 (9.37)	102 (9.35)	1.38	.175	.32
Performance IQ	104 (18.16)	109 (14.94)	-1.04	.305	.30
Overall IQ	105 (13.07)	106 (10.77)	-0.26	.796	.08
Total AQ score	29.7 (9.83)	16.4 (6.12)	5.74	< .001***	1.62
ADOS-2 Module 4 algorithm total	6.32 (3.96)	-			
Interference score (msec, Stroop)	91 (66)	61 (41)	-1.9	.063·	.55
Experiential index score (Imagination)	34.7 (8.58)	40.8 (5.39)	-3.04	.004**	.85

Table 2. *Mean (SD) values for each variable, region and condition for ASD and TD groups, Experiment 1.*

	ASD				**TD**			
	Counterfactual		Factual		Counterfactual		Factual	
	Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent
Pre-critical region								
First-pass reading time (ms)	831 (514)	898 (610)	703 (417)	821 (664)	941 (573)	948 (606)	798 (520)	841 (568)
Regression path reading time (ms)	1009 (813)	1098 (920)	815 (547)	868 (671)	1003 (635)	1077 (698)	867 (549)	929 (619)
Total reading time (ms)	1350 (1011)	1559 (1178)	1051 (651)	1319 (1102)	1304 (935)	1571 (1193)	1056 (708)	1313 (1015)
Critical region								
First-pass reading time (ms)	243 (109)	248 (116)	226 (89)	260 (187)	251 (104)	241 (99)	246 (120)	237 (94)
Regression path reading time (ms)	579 (915)	625 (807)	450 (525)	579 (711)	494 (810)	471 (839)	358 (268)	384 (332)
Total reading time (ms)	357 (253)	423 (363)	350 (234)	455 (351)	327 (195)	345 (237)	301 (164)	367 (273)
Post-critical region								
First-pass reading time (ms)	396 (256)	390 (258)	366 (244)	420 (281)	429 (279)	399 (282)	418 (267)	452 (283)
Regression path reading time (ms)	1444 (1323)	1963 (1909)	1354 (1308)	2054 (2009)	1404 (1655)	1925 (2225)	1076 (1166)	1774 (1804)
Total reading time (ms)	558 (428)	616 (504)	556 (416)	688 (421)	557 (385)	578 (420)	509 (336)	647 (441)

Table 3. *Model Estimate, Standard Error (SE) and t-value for first pass reading time, regression path and total reading time in each region for Experiment 1 ('p < .10, *p < .05, **p < .01, ***p < .001).*

	First-pass reading times			**Regression path reading time**			**Total reading times**		
	Est.	*SE*	*t*	*Est.*	*SE*	*t*	*Est.*	*SE*	*t*
Pre-critical region									
Context	-0.078	0.013	-6.05 ***	-0.088	0.009	-9.31 ***	-0.084	0.010	-8.07 ***
Consistency	-0.013	0.011	-1.11	-0.016	0.009	-1.68	-0.065	0.011	-6.16 ***
Group	0.035	0.039	0.91	0.011	0.037	0.30	-0.011	0.044	-0.25
Imagination	0.003	0.002	1.42	0.005	0.002	2.36 *	0.005	0.003	1.94 '
Inhibitory Control	0.000	0.000	0.42	0.000	0.000	1.15	0.000	0.000	1.31
Context:Consistency	-0.013	0.020	-0.65	0.019	0.018	1.07	0.012	0.019	0.65
Context:Group	-0.003	0.020	-0.14	0.023	0.018	1.32	0.004	0.019	0.21
Consistency:Group	0.023	0.020	1.13	-0.005	0.018	-0.29	-0.016	0.019	-0.81
Context:Consistency:Group	0.019	0.039	0.48	-0.013	0.035	-0.37	0.004	0.037	0.10
Critical region									
Context	-0.011	0.010	-1.11	-0.020	0.021	-0.98	0.004	0.016	0.25
Consistency	-0.003	0.012	-0.26	-0.025	0.019	-1.29	-0.052	0.019	-2.78 **
Group	-0.008	0.020	-0.39	-0.113	0.047	-2.42 *	-0.061	0.032	-1.90 '
Imagination	0.003	0.001	2.54 *	0.007	0.003	2.7 *	0.003	0.002	1.66
Inhibitory Control	0.000	0.000	1.10	0.001	0.000	1.48	0.000	0.000	1.14
Context:Consistency	-0.020	0.020	-0.99	-0.031	0.033	-0.94	-0.039	0.026	-1.49
Context:Group	-0.004	0.019	-0.19	0.009	0.041	0.23	-0.008	0.028	-0.29
Consistency:Group	0.040	0.020	2.00 *	0.076	0.035	2.18 *	0.039	0.025	1.56
Context:Consistency:Group	0.032	0.041	0.77	-0.018	0.065	-0.28	-0.023	0.052	-0.45
Post-critical region									
Context	0.010	0.012	0.82	-0.032	0.020	-1.62	0.019	0.013	1.46
Consistency	-0.005	0.013	-0.39	-0.153	0.024	-6.35 ***	-0.064	0.016	-4.03 ***
Group	0.011	0.035	0.33	-0.106	0.083	-1.27	-0.037	0.048	-0.77
Imagination	0.001	0.002	0.65	0.006	0.005	1.13	0.003	0.003	0.95
Inhibitory Control	0.000	0.000	-1.16	0.000	0.001	-0.03	0.000	0.000	-0.15
Context:Consistency	-0.059	0.024	-2.46 *	-0.071	0.036	-1.98 *	-0.081	0.025	-3.28 **
Context:Group	0.035	0.024	1.45	-0.019	0.036	-0.53	-0.015	0.025	-0.61
Consistency:Group	0.038	0.024	1.56	-0.010	0.036	-0.28	0.028	0.027	1.02
Context:Consistency:Group	0.014	0.048	0.29	-0.015	0.072	-0.21	0.009	0.049	0.17

Table 4. *Mean (SD) values for each variable, region and condition for ASD and TD groups, Experiment 2.*

	ASD				**TD**			
	Counterfactual		Factual		Counterfactual		Factual	
	Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent	Consistent	Inconsistent
Pre-critical region								
First-pass reading time (ms)	607 (394)	566 (353)	505 (292)	545 (360)	646 (330)	652 (346)	550 (293)	573 (351)
Regression path reading time (ms)	701 (459)	721 (588)	575 (352)	622 (433)	707 (346)	731 (441)	612 (339)	649 (392)
Total reading time (ms)	1013 (577)	1154 (860)	805 (548)	855 (567)	928 (502)	1164 (760)	740 (389)	942 (626)
Critical region								
First-pass reading time (ms)	247 (114)	241 (127)	252 (116)	266 (125)	218 (76)	248 (109)	250 (102)	261 (117)
Regression path reading time (ms)	414 (494)	483 (538)	490 (602)	490 (513)	408 (424)	454 (510)	398 (384)	481 (600)
Total reading time (ms)	383 (257)	398 (241)	405 (265)	468 (335)	327 (256)	397 (306)	333 (201)	412 (242)
Post-critical region								
First-pass reading time (ms)	377 (257)	405 (296)	439 (347)	451 (369)	407 (262)	407 (279)	405 (268)	425 (311)
Regression path reading time (ms)	1561 (1328)	1901 (1627)	1315 (1161)	1526 (1255)	1116 (904)	1796 (1737)	980 (839)	1531 (1482)
Total reading time (ms)	599 (422)	674 (482)	598 (462)	689 (516)	530 (309)	595 (419)	506 (304)	627 (497)

Table 5. *Model Estimate, Standard Error (SE) and t-value for first pass reading time, regression path and total reading time in each region for Experiment 2 ('p < .10, *p < .05, **p < .01, ***p < .001).*

	First-pass reading times			**Regression path reading time**			**Total reading times**		
	Est.	*SE*	*t*	*Est.*	*SE*	*t*	*Est.*	*SE*	*t*
Pre-critical region									
Context	-0.074	0.013	-5.48 ***	-0.083	0.012	-7.19 ***	-0.131	0.014	-9.49 ***
Consistency	-0.001	0.012	-0.10	-0.009	0.011	-0.86	-0.059	0.012	-5.12 ***
Group	0.026	0.038	0.67	0.021	0.037	0.57	-0.015	0.040	-0.37
Imagination	0.004	0.002	1.83 '	0.004	0.002	1.94 '	0.005	0.003	2.07 *
Inhibitory Control	0.000	0.000	0.11	0.000	0.000	1.33	0.000	0.000	1.33
Context:Consistency	-0.022	0.021	-1.06	-0.012	0.020	-0.61	0.007	0.020	0.33
Context:Group	-0.024	0.021	-1.17	-0.001	0.020	-0.08	0.034	0.020	1.66 '
Consistency:Group	-0.023	0.023	-1.00	-0.011	0.021	-0.53	-0.052	0.023	-2.24 *
Context:Consistency:Group	0.031	0.041	0.75	-0.003	0.039	-0.09	-0.025	0.041	-0.60
Critical region									
Context	0.028	0.012	2.38 *	0.025	0.019	1.32	0.037	0.012	3.08 **
Consistency	-0.019	0.011	-1.77 '	-0.043	0.018	-2.38 *	-0.061	0.015	-4.08 ***
Group	-0.032	0.019	-1.73 '	-0.058	0.041	-1.43	-0.073	0.031	-2.37 *
Imagination	0.004	0.001	3.52 ***	0.007	0.003	2.77 **	0.004	0.002	2.26 *
Inhibitory Control	0.000	0.000	0.94	0.000	0.000	1.12	0.000	0.000	0.915
Context:Consistency	-0.005	0.017	-0.26	-0.008	0.030	-0.25	-0.031	0.024	-1.263
Context:Group	0.012	0.019	0.63	-0.005	0.033	-0.17	-0.007	0.024	-0.294
Consistency:Group	-0.022	0.020	-1.10	-0.004	0.034	-0.13	-0.039	0.028	-1.389
Context:Consistency:Group	0.079	0.035	2.27 *	0.006	0.061	0.10	0.029	0.049	0.601
Post-critical region									
Context	0.026	0.013	2.01 *	-0.064	0.017	-3.73 ***	-0.001	0.013	-0.07
Consistency	-0.003	0.017	-0.18	-0.127	0.028	-4.61 ***	-0.046	0.020	-2.37 *
Group	-0.015	0.035	-0.42	-0.119	0.073	-1.64	-0.066	0.048	-1.37
Imagination	0.002	0.002	0.81	0.006	0.005	1.37	0.004	0.003	1.27
Inhibitory Control	0.000	0.000	0.04	0.000	0.001	0.82	0.000	0.000	0.31
Context:Consistency	0.002	0.025	0.08	0.009	0.034	0.28	-0.014	0.025	-0.54
Context:Group	-0.020	0.025	-0.80	0.039	0.034	1.15	0.017	0.026	0.64
Consistency:Group	0.013	0.026	0.49	-0.077	0.046	-1.69	-0.006	0.028	-0.20
Context:Consistency:Group	0.008	0.051	0.16	0.035	0.069	0.52	-0.005	0.050	-0.10

Appendix

Experimental Items, Experiment 1. Note that for each of the items below, conditions are listed in the order: counterfactual-consistent, counterfactual-inconsistent, factual-consistent, factual-inconsistent.

1
If Joanne had remembered her umbrella, her hair would have been dry when she arrived home.
If Joanne had remembered her umbrella, her hair would have been wet when she arrived home.
Because Joanne had remembered her umbrella, her hair had been dry when she arrived home.
Because Joanne had remembered her umbrella, her hair had been wet when she arrived home.

2
If the racing driver had won the championship, he would have gone to the pub to celebrate with his team.
If the racing driver had won the championship, he would have gone to the pub to commiserate with his team.
Because the racing driver had won the championship, he had gone to the pub to celebrate with his team.
Because the racing driver had won the championship, he had gone to the pub to commiserate with his team.

3
If David had been wearing his glasses, he would have found that the words were clear on the poster.
If David had been wearing his glasses, he would have found that the words were blurry on the poster.
Because David had been wearing his glasses, he had found that the words were clear on the poster.
Because David had been wearing his glasses, he had found that the words were blurry on the poster.

4
If Alice had not got stuck in a traffic jam driving home, she would have arrived home on time that evening.
If Alice had not got stuck in a traffic jam driving home, she would have arrived home late that evening.
Because Alice had got stuck in a traffic jam driving home, she had arrived home late that evening.
Because Alice had got stuck in a traffic jam driving home, she had arrived home on time that evening.

5
If Catherine had worn a coat in the rain, her dress would have been dry when she arrived at work.
If Catherine had worn a coat in the rain, her dress would have been soaked when she arrived at work.

Because Catherine had worn a coat in the rain, her dress had been dry when she arrived at work.
Because Catherine had worn a coat in the rain, her dress had been soaked when she arrived at work.

6
If David had eaten a large lunch, his stomach would have been full that afternoon.
If David had eaten a large lunch, his stomach would have been empty that afternoon.
Because David had eaten a large lunch, his stomach had been full that afternoon.
Because David had eaten a large lunch, his stomach had been empty that afternoon.

7
If Sylvia had worn thick socks in the snow, her feet would have been warm inside her boots.
If Sylvia had worn thick socks in the snow, her feet would have been cold inside her boots.
Because Sylvia had worn thick socks in the snow, her feet had been warm inside her boots.
Because Sylvia had worn thick socks in the snow, her feet had been cold inside her boots.

8
If Hannah had fixed her freezer, when she checked the ice trays the water would have been frozen that evening.
If Hannah had fixed her freezer, when she checked the ice trays the water would have been liquid that evening.
Because Hannah had fixed her freezer, when she checked the ice trays the water had been frozen that evening.
Because Hannah had fixed her freezer, when she checked the ice trays the water had been liquid that evening.

9
If Jacob had felt confident with his looks, whenever he looked in the mirror he would have thought he looked good for his age.
If Jacob had felt confident with his looks, whenever he looked in the mirror he would have thought he looked bad for his age.
Because Jacob had felt confident with his looks, whenever he looked in the mirror, he had thought he looked good for his age.
Because Jacob had felt confident with his looks, whenever he looked in the mirror, he had thought he looked bad for his age.

10
If Fred had seen the looming dark clouds, he would have prepared himself for rain as he got dressed.
If Fred had seen the looming dark clouds, he would have prepared himself for sunshine as he got dressed.
Because Fred had seen the looming dark clouds, he had prepared himself for rain as he got dressed.
Because Fred had seen the looming dark clouds, he had prepared himself for sunshine as he got dressed.

11
If Tim's team had correctly answered the most questions in the pub quiz, they would have won the top prize.

If Tim's team had correctly answered the most questions in the pub quiz, they would have lost the top prize.
Because Tim's team had correctly answered the most questions in the pub quiz, they had won the top prize.
Because Tim's team had correctly answered the most questions in the pub quiz, they had lost the top prize.

12
If John's mother had cooked him an enormous roast dinner, he would have felt very full as he watched TV.
If John's mother had cooked him an enormous roast dinner, he would have felt very hungry as he watched TV.
Because John's mother had cooked him an enormous roast dinner, he had felt very full as he watched TV.
Because John's mother had cooked him an enormous roast dinner, he had felt very hungry as he watched TV.

13
If Andy had understood everything in the statistics lecture, he would have found his homework very easy to complete correctly.
If Andy had understood everything in the statistics lecture, he would have found his homework very difficult to complete correctly.
Because Andy had understood everything in the statistics lecture, he had found his homework very easy to complete correctly.
Because Andy had understood everything in the statistics lecture, he had found his homework very difficult to complete correctly.

14
If Jill had not put her ice lolly back in the freezer before she answered the door, it would have been melted when she returned to it.
If Jill had not put her ice lolly back in the freezer before she answered the door, it would have been frozen when she returned to it.
Because Jill had put her ice lolly back in the freezer before she answered the door, it had been frozen when she returned to it.
Because Jill had put her ice lolly back in the freezer before she answered the door, it had been melted when she returned to it.

15
If Guy’s debate team had made a successful argument, the compere would have announced them as the winners of the big debate.
If Guy’s debate team had made a successful argument, the compere would have announced them as the losers of the big debate.
Because Guy's debate team had made a successful argument, the compere had announced them as the winners of the big debate.
Because Guy's debate team had made a successful argument, the compere had announced them as losers of the big debate.

16
If Liz had stuck to the weight watchers plan, she would have noticed that her clothes were feeling loose after two weeks.

If Liz had stuck to the weight watchers plan, she would have noticed that her clothes were feeling tight after two weeks.
Because Liz had stuck to the weight watchers plan, she had noticed that her clothes were feeling loose after two weeks.
Because Liz had stuck to the weight watchers plan, she had noticed that her clothes were feeling tight after two weeks.

17
If the central heating system had not broken down, the housemates would have found the house was warm in the morning.
If the central heating system had not broken down, the housemates would have found the house was cold in the morning.
Because the central heating system had broken down, the housemates had found the house was cold in the morning.
Because the central heating system had broken down, the housemates had found the house was warm in the morning.

18
If Lee had worked out at the gym after work and felt tired, he would have decided to go to bed early that evening.
If Lee had worked out at the gym after work and felt tired, he would have decided to go to bed late that evening.
Because Lee had worked out at the gym after work and felt tired, he had decided to go to bed early that evening.
Because Lee had worked out at the gym after work and felt tired, he had decided to go to bed late that evening.

19
If Maria had not lost her case, she would have left the courtroom feeling happy about the outcome.
If Maria had not lost her case, she would have left the courtroom feeling upset about the outcome.
Because Maria had lost her case, she had left the courtroom feeling upset about the outcome.
Because Maria had lost her case, she had left the courtroom feeling happy about the outcome.

20
If the comedian had thoroughly prepared his act, his audience would have left feeling happy at the end.
If the comedian had thoroughly prepared his act, his audience would have left feeling disappointed at the end.
Because the comedian had thoroughly prepared his act, his audience had left feeling happy at the end.
Because the comedian had thoroughly prepared his act, his audience had left feeling disappointed at the end.

21
If Alison had not got very drunk at the party, she would have woken up feeling refreshed the next day.
If Alison had not got very drunk at the party, she would have woken up feeling hungover the next day.

Because Alison had got very drunk at the party, she had woken up feeling hungover the next day.
Because Alison had got very drunk at the party, she had woken up feeling refreshed the next day.

22
If Inez' grandchildren had not knocked over the antique vase, it would have been intact when her guests arrived.
If Inez' grandchildren had not knocked over the antique vase, it would have been broken when her guests arrived.
Because Inez' grandchildren had knocked over the antique vase, it had been broken when her guests arrived.
Because Inez' grandchildren had knocked over the antique vase, it had been intact when her guests arrived.

23
If Charlie and his wife had not been arguing, they would have eaten dinner together happily at the table.
If Charlie and his wife had not been arguing, they would have eaten dinner together in silence at the table.
Because Charlie and his wife had been arguing, they had eaten together in silence at the table.
Because Charlie and his wife had been arguing, they had eaten together happily at the table.

24
If Stephen had not done well in his first year at University, he would have been disappointed with the results.
If Stephen had not done well in his first year at University, he would have been happy with the results.
Because Stephen had done well in his first year at University, he had been happy with his results.
Because Stephen had done well in his first year at University, he had been disappointed with his results.

25
If Jessica had not enjoyed science, the day she spent visiting the science museum would have felt too long for her.
If Jessica had not enjoyed science, the day she spent visiting the science museum would have felt too short for her.
Because Jessica enjoyed science, the day she spent visiting the science museum had felt too short for her.
Because Jessica enjoyed science, the day she spent visiting the science museum had felt too long for her.

26
If Elaine's car had not passed its MOT, it would have been deemed dangerous to drive on the roads.
If Elaine's car had not passed its MOT, it would have been deemed safe to drive on the roads.
Because Elaine's car had passed its MOT, it had been deemed safe to be driven on the roads.
Because Elaine's car had passed its MOT, it had been deemed dangerous to be driven on the roads.

27
If John had not hurt his back, when he woke up in the morning, it would have felt fine all over.
If John had not hurt his back, when he woke up in the morning, it would have felt stiff all over.
Because John had hurt his back, when he woke up in the morning, it had felt stiff all over.
Because John had hurt his back, when he woke up in the morning, it had felt fine all over.

28
If the earthquake had not caused a radioactive isotope to leak from the nuclear plant, the air would have been clean the next day.
If the earthquake had not caused a radioactive isotope to leak from the nuclear plant, the air would have been contaminated the next day.
Because the earthquake had caused a radioactive isotope to leak from the nuclear plant, the air had been contaminated the next day.
Because the earthquake had caused a radioactive isotope to leak from the nuclear plant, the air had been clean the next day.

29
If Greg had not remembered his asthma inhaler at the football game, his breathing would have been erratic at half time.
If Greg had not remembered his asthma inhaler at the football game, his breathing would have been fine at half time.
Because Greg had remembered his asthma inhaler at the football game, his breathing had been fine at half time.
Because Greg had remembered his asthma inhaler at the football game, his breathing had been erratic at half time.

30
If the gymnast had not warmed up sufficiently and not performed well, her coach would have criticised her effort.
If the gymnast had not warmed up sufficiently and not performed well, her coach would have praised her effort.
Because the gymnast had warmed up sufficiently and performed well, her coach had praised her effort.
Because the gymnast had warmed up sufficiently and performed well, her coach had criticised her effort.

31
If Sally had not had a skiing accident, she would have boarded the flight home with no problems after a week.
If Sally had not had a skiing accident, she would have boarded the flight home with crutches after a week.
Because Sally had had a skiing accident, she had boarded the flight home with crutches after a week.
Because Sally had had a skiing accident, she had boarded the flight home with no problems after a week.

32

If Jamie had not put sugar in her coffee, she would have tasted that it was bitter and very strong.
If Jamie had not put sugar in her coffee, she would have tasted that it was sweet and very strong.
Because Jamie had put sugar in her coffee, she had tasted that it was sweet and very strong.
Because Jamie had put sugar in her coffee, she had tasted that it was bitter and very strong.

Experimental items, Experiment 2. Note that for each of the items below, conditions are listed in the order: counterfactual-consistent, counterfactual-inconsistent, factual-consistent, factual-inconsistent.

1
If Spain were not a member of the European Union, they would pay for things using pesetas in shops today.
If Spain were not a member of the European Union, they would pay for things using Euros in shops today.
Because Spain is a member of the European Union, they pay for things using euros in shops today.
Because Spain is a member of the European Union, they pay for things using pesetas in shops today.

2
If people did not celebrate Christmas, the 25th December would be a normal day in the UK.
If people did not celebrate Christmas, the 25th December would be a holiday in the UK.
Because people celebrate Christmas, the 25th December is a holiday in the UK.
Because people celebrate Christmas, the 25th December is a normal day in the UK.

3
If the UK had not won the contest to host the 2012 Olympics Games, they would have taken place in Paris that year.
If the UK had not won the contest to host the 2012 Olympics Games, they would have taken place in London that year.
Because the UK won the contest to host the 2012 Olympic Games, they took place in London that year.
Because the UK won the contest to host the 2012 Olympic Games, they took place in Paris that year.

4
If the internet had not been created, we would tend to write letters to communicate.
If the internet had not been created, we would tend to write emails to communicate.
Because the internet was created, we tend to write emails to communicate.
Because the internet was created, we tend to write letters to communicate.

5
If Labour had not lost the May 2015 election, the Prime Minister would be Ed Miliband for five years.
If Labour had not lost the May 2015 election, the Prime Minister would be David Cameron for five years.

Because Labour lost the May 2015 election, the Prime Minister is David Cameron for five years.
Because Labour lost the May 2015 election, the Prime Minister is Ed Miliband for five years.

6
If electricity had not been discovered, we would light our homes with candles at night.
If electricity had not been discovered, we would light our homes with lightbulbs at night.
Because electricity was discovered, we light our homes with lightbulbs at night.
Because electricity was discovered, we light our homes with candles at night.

7
If the Titanic had not hit an iceberg, it would have survived along with all the passengers.
If the Titanic had not hit an iceberg, it would have sunk along with all the passengers.
Because the Titanic hit an iceberg, it sunk along with all the passengers.
Because the Titanic hit an iceberg, it survived along with all the passengers.

8
If Kate Middleton had not married Prince William, today she would be unknown throughout the world.
If Kate Middleton had not married Prince William, today she would be famous throughout the world.
Because Kate Middleton married Prince William, today she is famous throughout the world.
Because Kate Middleton married Prince William, today she is unknown throughout the world.

9
If tuition fees had not increased in 2012, Bachelor's degrees would cost three thousand pounds per year.
If tuition fees had not increased in 2012, Bachelor's degrees would cost nine thousand pounds per year.
Because tuition fees have increased in 2012, Bachelor's degrees cost nine thousand pounds per year.
Because tuition fees have increased in 2012, Bachelor's degrees cost three thousand pounds per year.

10
If the car driving Princess Diana had not crashed, today she would be alive, along with Dodi.
If the car driving Princess Diana had not crashed, today she would be dead, along with Dodi.
Because the car driving Princess Diana did crash, today she is dead, along with Dodi.
Because the car driving Princess Diana did crash, today she is alive, along with Dodi.

11
If Coca-Cola had not changed the colour of Santa Claus' suit, it would be green and white.
If Coca-Cola had not changed the colour of Santa Claus' suit, it would be red and white.
Because Coca-Cola changed the colour of Santa Claus' suit, it is red and white.
Because Coca-Cola changed the colour of Santa Claus' suit, it is green and white.

12
If there weren't any laws against murder, the majority of murderers would be free in this country.

If there weren't any laws against murder, the majority of murderers would be imprisoned in this country.
Because there are laws against murder, the majority of murderers are imprisoned in this country.
Because there are laws against murder, the majority of murderers are free in this country.

13
If mobile phones weren't invented, meeting up with people would be harder when out and about.
If mobile phones weren't invented, meeting up with people would be easier when out and about.
Because mobile phones were invented, meeting up with people is easier when out and about.
Because mobile phones were invented, meeting up with people is harder when out and about.

14
If IKEA furniture did not have instructions, assembling it would be difficult for most people.
If IKEA furniture did not have instructions, assembling it would be easy for most people.
Because IKEA furniture comes with instructions, assembling it is easy for most people.
Because IKEA furniture comes with instructions, assembling it is difficult for most people.

15
If we did not have the NHS, prescription medication would be relatively expensive in most cases.
If we did not have the NHS, prescription medication would be relatively cheap in most cases.
Because we have the NHS, prescription medication is relatively cheap in most cases.
Because we have the NHS, prescription medication is relatively expensive in most cases.

16
If we did not have washing machines, we would wash our clothes by hand in our homes.
If we did not have washing machines, we would wash our clothes by machine in our homes.
Because we have washing machines, we wash our clothes by machine in our homes.
Because we have washing machines, we wash our clothes by hand in our homes.

17
If waste produced from nuclear energy did not emit radiation, it would be safe for us.
If waste produced from nuclear energy did not emit radiation, it would be dangerous for us.
Because waste produced from nuclear energy emits radiation, it is dangerous for us.
Because waste produced from nuclear energy emits radiation, it is safe for us.

18
If cars had not been invented, we would get to places a lot slower in most cases.
If cars had not been invented, we would get to places a lot faster in most cases.
Because cars were invented, we get to places a lot faster in most cases.
Because cars were invented, we get to places a lot slower in most cases.

19
If the mp3 format had not been invented, CD sales would be higher in retail outlets.
If the mp3 format had not been invented, CD sales would be lower in retail outlets.
Because the mp3 format was invented, CD sales are lower in retail outlets.
Because the mp3 format was invented, CD sales are higher in retail outlets.

20
If there wasn't a limit to how much alcohol you could drink before driving, travelling on the roads would be riskier on average.
If there wasn't a limit to how much alcohol you could drink before driving, travelling on the roads would be safer on average.
Because there is a limit to how much alcohol you can drink before driving, travelling on the roads is safer on average.
Because there is a limit to how much alcohol you can drink before driving, travelling on the roads is riskier on average.

21
If glasses had not been invented, many people would have vision that is worse as a result.
If glasses had not been invented, many people would have vision that is better as a result.
Because glasses were invented, many people have vision that is better as a result.
Because glasses were invented, many people have vision that is worse as a result.

22
If there were no anti-smoking laws, smoking would be permitted in most workplaces.
If there were no anti-smoking laws, smoking would be forbidden in most workplaces.
Because there are anti-smoking laws, smoking is forbidden in most workplaces.
Because there are anti-smoking laws, smoking is permitted in most workplaces.

23
If dogs had never been tamed, they would be considered wild nowadays.
If dogs had never been tamed, they would be considered domesticated nowadays.
Because dogs have been tamed, they are considered domesticated nowadays.
Because dogs have been tamed, they are considered wild nowadays.

24
If the Channel Tunnel had not been built, the journey between London and Paris would be longer for travellers.
If the Channel Tunnel had not been built, the journey between London and Paris would be shorter for travellers.
Because the Channel Tunnel was built, the journey between London and Paris is shorter for travellers.
Because the Channel Tunnel was built, the journey between London and Paris is longer for travellers.

25
If the Berlin wall had not been pulled down, Germany would be a divided country in Europe.
If the Berlin wall had not been pulled down, Germany would be a united country in Europe.
Because the Berlin wall was pulled down, Germany is a united country in Europe.
Because the Berlin wall was pulled down, Germany is a divided country in Europe.

26
If AL 'Qaeda had not carried out the attacks on September 11th, the twin towers would be standing in New York City.
If AL 'Qaeda had not carried out the attacks on September 11th, the twin towers would be destroyed in New York City.

Because Al 'Qaeda carried out the attacks on September 11th, the twin towers are destroyed in New York City.
Because Al 'Qaeda carried out the attacks on September 11th, the twin towers are standing in New York City.

27
If it weren't forbidden for Muslims to drink alcohol, they would be able to drink wine among other things.
If it weren't forbidden for Muslims to drink alcohol, they would be able to drink water among other things.
Because it is forbidden for Muslims to drink alcohol, they can drink water among other things.
Because it is forbidden for Muslims to drink alcohol, they can drink wine among other things.

28
If the financial crisis had not affected the UK, there would be a lot of money across the country.
If the financial crisis had not affected the UK, there would be a lot of poverty across the country.
Because the financial crisis affected the UK, there is a lot of poverty across the country.
Because the financial crisis affected the UK, there is a lot of money across the country.

29
If we had not discovered fire, we would eat our food when it is raw most of the time.
If we have not discovered fire, we would eat our food when it is cooked most of the time.
Because we discovered fire, we eat our food when it is cooked most of the time.
Because we discovered fire, we eat our food when it is raw most of the time.

30
If broadband had not been invented, internet connections would be much slower these days.
If broadband had not been invented, internet connections would be much faster these days.
Because broadband was invented, internet connections are much faster these days.
Because broadband was invented, internet connections are much slower these days.

31
If Susan Boyle hadn't taken part in Britain's Got Talent, today she would be working as a cashier to earn money.
If Susan Boyle hadn't taken part in Britain's Got Talent, today she would be working as a singer to earn money.
Because Susan Boyle took part in Britain's Got Talent, today she is working as a singer to earn money.
Because Susan Boyle took part in Britain's Got Talent, today she is working as a cashier to earn money.

32
If Bill Gates had not created Microsoft, his fortune would be less than the average person.
If Bill Gates had not created Microsoft, his fortune would be more than the average person.
Because Bill Gates created Microsoft, his fortune is more than the average person.
Because Bill Gates created Microsoft, his fortune is less than the average person.