Passive sentence reversal errors in autism: Replicating Ambridge, Bidgood, and Thomas (2020)

Samuel Jones
Lancaster University, UK
ESRC International Centre for Language and Communicative Development (LuCiD)

Madeline Dooley
University of Liverpool, UK

Ben Ambridge
University of Liverpool, UK
ESRC International Centre for Language and Communicative Development (LuCiD)

Abstract: Ambridge, Bidgood, and Thomas (2020) conducted an elicitation-production task in which children with and without autism described animations following priming with passive sentences. The authors reported that children with autism were more likely than IQ-matched children without autism to make reversal errors, for instance describing a scene in which the character Wendy surprised the character Bob by saying *Wendy was surprised by Bob*. We set out to test whether this effect replicated in a new sample of children with and without autism (*N* = 26) and present a cumulative analysis in which data from the original study and the replication are pooled (*N* = 56). The main effect reported by Ambridge et al. (2020) replicated: While children with and without autism produced a similar number of passive responses in general, the responses of children with autism were significantly more likely to include reversal errors. Despite age-appropriate knowledge of constituent order in passive syntax, thematic role assignment is impaired among some children with autism.

Keywords: autism; syntax; priming; passives; language disorder

Corresponding author(s): Correspondence concerning this article should be addressed to Sam Jones, Department of Psychology, Lancaster University, Lancaster, United Kingdom, LA1 4YF. Email: sam.jones@lancaster.ac.uk. Telephone: +44 (0) 1524 593698.

ORCID ID(s): Samuel Jones: https://orcid.org/0000-0002-8870-3223; Ben Ambridge: https://orcid.org/0000-0003-2389-8477

Citation: Jones, S. D., Dooley, M., & Ambridge, B. (2021). Passive sentence reversal errors in autism: Replicating Ambridge, Bidgood, and Thomas (2020). *Language Development Research*, 1(1), 99-122. DOI: 10.34842/g1zk-3715

Introduction

Language development in autism

Approximately 1% of English-learning children are affected by autism, defined as persistent deficits in social interaction and communication, and restricted and repetitive patterns of behaviour, interests, or activities (Baron-Cohen et al., 2009). The language abilities of children with autism vary widely. Some children have little or no language, while others have advanced language skills and may appear pedantic or verbose. Although, as a group, children with autism tend to use shorter and grammatically simpler sentences than children without autism (Eigsti, Bennetto, & Dadlani, 2007), the acquisition of morphosyntax and word order appear relatively standard among affected children (Tek, Mesite, Fein, & Naigles, 2014). Semantic-pragmatic and narrative development are, in contrast, key areas of difficulty for children with autism, who often show poor understanding of metaphorical and figurative language, poor inferencing skills, difficulty resolving semantic ambiguities (e.g. homographs), and pronoun reversals in which the speaker mistakenly uses *you* in self-reference and *I* to refer to the listener (Naigles & Tek, 2017; see Norbury, 2015, for review).

Target study: Ambridge, Bidgood, and Thomas (2020)

The aim of the current work was to replicate a study that attempted to separate out syntactic and semantic-pragmatic factors contributing to language deficits in a group of children with autism. It is important to be clear at the outset exactly what we mean by syntactic versus semantic-pragmatic factors. Here, we adopt the definition set out in the study that is the target of our replication (Ambridge et al, 2020: 185).

A widely held view in the literature is that, despite broader linguistic and communicative difficulties, ‘pure syntax’ is relatively spared in children with autism, i.e., spared relative to the broader cognitive deficits that accompany this condition. On this view, which might be summarized in the phrase ‘form is easy, meaning is hard’ (Naigles, 2002; Naigles & Tek, 2017), syntax itself is spared, and the communicative difficulties that are experienced by children with autism are caused by impairments in other areas of language, such as vocabulary, semantics, socio-pragmatics and narrative (e.g., Tager-Flusberg, Lord & Paul, 1997; Jordan, 1993). To be clear, there is evidence that even children without autism find certain semantic or pragmatic aspects of language more difficult than purely syntactic or structural ones (Naigles, 2002); the claim is, then, that this is even more true for children with autism.

Ambridge et al. (2020) investigated the ability of children aged 6-9 years, with and without autism, to accurately describe an animation using primed passive sentences such as *Bob was surprised/chased/pulled by Wendy*. These authors argue that the (English) passive is a particularly useful test case for separating out syntactic and semantic or pragmatic impairments, since it exemplifies standard syntactic representations and relations – e.g. [SUBJECT] [BE] [VERB] ([PP]) – yet is unusual in terms of its semantics and pragmatics, reversing the [AGENT][PATIENT] word order of actives, and treating the [PATIENT] rather than the [AGENT] as topical.

In the target study of Ambridge et al. (2020), one experimenter described an example video animation using a model passive sentence before a second experimenter provided the participant with a cue verb with which to describe a novel animation. For example, given the verb *surprise* and an animation in which the character Wendy surprised the character Bob, a successfully primed response of *Bob was surprised by Wendy* was coded as a correct passive. Correct actives, meanwhile, were coded when the child produced responses such as *Wendy surprised Bob*; that is, when there was little evidence of a priming effect and the child defaulted to the more frequent active form. Of central interest in the Ambridge et al. (2020) study – and in the current replication – was the rate of reversal errors children made, in which a passive response exhibited an error in thematic role assignment. For instance, in response to the animation in which Wendy surprised Bob – that is, Wendy is the [AGENT] and Bob is the [PATIENT] – the child produced the passive *Wendy was surprised by Bob*; mis-assigning Wendy as [PATIENT] and Bob as [AGENT]. Given that many children affected by autism have difficulty with the referential, inferential, and narrative building aspects of language, it was hypothesised that this group would produce a higher rate of passive reversal errors than IQ-matched children without autism.

Ambridge et al. (2020) note that prior work testing passive sentence comprehension among children with autism reports mixed results. For instance, Tager-Flusberg (1981) reported that children with autism (aged M = 8;1) were no more likely than younger (M = 3;10) IQ-matched controls to mis-comprehend passive structures, as evidenced in an act-out task. In contrast, Paul, Fisher, and Cohen (1988), who used stimuli matched to those used by Tager-Flusberg (1981), reported evidence that children with autism do make more reversal errors than IQ-matched controls. Ambridge et al. (2020) was the first production study to look at reversal errors in children with autism, with prior production studies in this area excluding reversal errors from analyses (e.g. Allen, Haywood, Rajendran, & Branigan, 2011). Ambridge et al. (2020) report a modest though reliable pattern of higher reversal errors among children with autism relative to IQ-matched peers. These results were interpreted as further evidence that semantics, pragmatics, and narrative, rather than 'pure syntax', constitute key areas of language difficulty for children affected by autism (though other interpretations are possible; a point to which we return at length in the *Discussion*).

Why replicate?

The value of the Ambridge et al. (2020) study is that it investigates a specific grammatical structure – the passive. Work testing the processing and comprehension of specific grammatical structures among children with autism is lacking (Norbury, 2014), and this is unfortunate because such work can provide a basis for developing finely targeted games or activities to be used in programmes of language support. In the case of the passive, for instance, if children with autism do indeed have a good command of the syntax of this construction, but not of its semantic-pragmatic aspects, interventions based around this construction should focus on the latter, not the former. For example, a narrative-based intervention which emphasizes how the current discourse topic (e.g. *Have you heard the news about YouTube?...)* makes a natural passive PATIENT-SUBJECT (*...It got bought by Google*; Pullum, 2014: 64) is likely to be more useful than one focussed directly on syntax, such as a task encouraging children to produce passive sentences when describing pictures with no discourse context.

Nevertheless, one limitation of Ambridge et al. (2020) – and indeed all prior studies into passive sentence processing and comprehension in children with autism – is that as a hard-to-reach population, language development studies involving children with autism often have small sample sizes. For instance, Ambridge et al. (2020) tested 15 children with autism, while Tager-Flusberg (1981) tested 18 children, and Paul et al. (1988) tested just six children. For this reason, while the use of a paradigm sensitive enough to identify specific deficits in the processing of a defined linguistic structure among children with autism is welcomed, without further replication many readers may be unconvinced by this effect, especially given its small magnitude.

The current study

The purpose of the current study was, therefore, to test whether the findings of Ambridge et al. (2020) – i.e., higher rates of reversal error among children with autism than among IQ-matched peers – replicated in a new sample of children. In approaching this project, we faced similar resourcing constraints, and tested a similar number of children (N = 26, n = 13 with autism). However, re-using the original stimuli and procedure enabled us to produce – in addition to our own replication analysis – a cumulative analysis of the pooled data involving 28 participants per group (N = 56). Cumulative analysis should be distinguished from questionable research practices such as optional stopping or p-hacking, in which researchers covertly gather data up to the point at which their hypothesis is superficially confirmed, or add or remove specific data points in order to retrieve a p-value below the standard .05 alpha level. In contrast, we explicitly label ‘original’, ‘replication’, and ‘pooled’ data throughout this analysis, and all of our data and code is made publicly available via an online repository: https://osf.io/c2pjd/.

In both the present replication and the original study, we used syntactic (or ‘structural’) priming purely as a method for eliciting passives. The phenomenon of syntactic priming itself is not under investigation, and we remain agnostic with regard to the question of whether priming constitutes a particularly useful window into children's learning and representation of structural knowledge. Since the passive is a highly infrequent and marked construction, it is likely that most children would have produced very few, if any, passives, had we run the study as a simple elicited-production task with no priming element.

Method

Participants

Thirteen children aged 6 to 9 years (*M* = 8;0) with autism were recruited from specialist schools in North West England. Entry to these schools was based on a prior diagnosis of autism and an extensive battery of screening assessments, resembling that shown in Appendix A of Ambridge et al. (2020). In the current study, we took the additional precaution of screening children independently using the Lifetime version of the Social Communicative Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003). The SCQ comprises 40 items to which caregivers are required to provide yes or no responses. Responses are then tallied to determine the child's SCQ score. A child with an SCQ score of 15 or over is likely to be on the autistic spectrum. Children in the autism group of the current replication study had SCQ scores ranging from 19 to 29 (*M* = 22.85), providing independent validation of diagnosis and experimental group identity. Thirteen children without autism aged four to six (*M* = 5;3) were recruited from mainstream English pre-schools and schools. By-participant demographics and SCQ and IQ scores are presented in the Appendix.

Following Ambridge et al. (2020), children with and without autism were IQ-matched using the short version of the Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV; Wechsler, 2012). IQ scores were used to match the with-autism and without-autism groups, and for use as a control predictor in the statistical analyses, but were not used to define cut-offs for either group. The results of this administration are shown in Table 1, alongside corresponding administration results from Ambridge et al. (2020). Visual inspection of this data indicates reasonable similarity both across studies and between experimental and control groups within studies. Where there are discrepancies between groups, these are attributable to children with autism outperforming children without autism, meaning that matching may be considered conservative. For instance, in both the original study and in the replication, children with autism scored numerically higher on the object assembly subset of the WPPSI-IV, while in the replication a numerical advantage for picture memory was also recorded among children with autism.

It is important to note that, since the children with autism were considerably older than the IQ-matched children without autism (i.e., a mean age of 8;0, as opposed to 5;3), it would not be accurate to refer to the former group as 'children with autism but without intellectual disabilities' (previously termed 'high functioning autism'; though see Alvares, Bebbington, Clearly, Evans, Glasson, et al., 2020). However, since all of the children with autism were able to complete a relatively complex verbal task, any intellectual disabilities present for children in this group were relatively minor. It would also not, therefore, be appropriate to generalize the findings from the present study (or that of Ambridge et al, 2020, which was conducted with similar participant groups) to children with autism with greater intellectual disabilities.

Table 1. Mean (and standard deviation) scores for the Wechsler Preschool and Primary Scale of Intelligence, Fourth Edition (WPPSI-IV), across seven subsets. The unit-weighted composite mean is also shown. WA = Without Autism; ASC = autism spectrum condition.

	Ambridge, Bidgood, and Thomas (2020) (*N* = 30)		Jones, Dooley, and Ambridge (2020) (*N* = 26)	
	ASC (*n* = 15)	WA (*n* = 15)	ASC (*n* = 13)	WA (*n* = 13)
Receptive vocabulary	22.98 (3.36)	23 (2.7)	22.41 (3.04)	21.62 (1.82)
Block design	19.7 (7.7)	20.48 (2.69)	20.99 (3.22)	19.32 (2.2)
Picture memory	20.37 (5.33)	15.34 (4.28)	20.35 (2.73)	18.62 (2.53)
Information	19.19 (3.66)	21.02 (2.29)	20.63 (3.23)	19.64 (2.68)
Object assembly	28.38 (6.7)	22.45 (7.88)	25.22 (2.89)	19.74 (2.48)
Zoo locations	10.42 (2.53)	11.13 (0.99)	12.19 (1.47)	12.03 (2.02)
Picture naming	17.95 (2.71)	17.45 (2.5)	17.43 (2.33)	17.57 (1.5)
Unit-weighted composite mean	19.86 (2.97)	18.7 (2.18)	19.89 (1.92)	18.36 (1.24)

Scaled, unit-weighted composite means of WPPSI-IV scores were included as control variables in the hierarchical Bayesian models used throughout this study (referred to as 'IQ' in Ambridge et al., 2020). These composite means were calculated by summing the scaled scores for each subset for each child and then dividing by the number of subsets (i.e., seven). The value shown at the bottom of Table 1 was, in contrast, calculated by summing the raw (i.e., not scaled) mean scores across children and dividing by the number of children. Note that these mean scores in the replication align well with those in the original article (e.g., for the experimental group, $M = 19.89$ versus $M = 19.86$). While Ambridge et al. (2020) used independent t-tests to check for equivalence between groups – and reported no statistically significant differences on the basis of the data shown in Table 1 – we avoided this analysis given concerns regarding the use of inferential methods to test for so-called nuisance effects (Sassenhagen & Alday, 2016). Readers interested in formally testing for equivalence between groups may use our R code to do so.

Procedure and scoring

The procedure and scoring used in this study were identical to those used in Ambridge et al. (2020). The participant and two experimenters sat in a quiet room in front of a computer screen and played a bingo-style game designed to engage and sustain the participant's attention. One experimenter acted as adjudicator, and first passed a prime verb card (Table 2) to the second experimenter. The second experimenter then used the specified prime verb in a passive sentence to describe a short animation played on the computer screen. After this, a target verb card (Table 3) was passed to the participant, who was required to use the specified verb to describe a novel animation. After each trial, the adjudicator, who was not able to see the computer screen, looked into a tub and, if one was available, retrieved a bingo point card corresponding to the description made. The game was engineered to ensure that the participant always finished with more bingo points than the experimenter.

Inspection of Tables 2 and 3, which shows age-of-acquisition and (where available) familiarity ratings (Bird, Franklin & Howards, 2001; Kuperman, Stadthagen-Gonzalez & Brysbaert, 2012), suggests that the majority of prime and target verbs would be known by even the youngest children who participated in the present study. Note that in the original study target verbs (Table 3) were split into three semantic classes; agent-patient, experiencer-theme, and theme-experiencer. This manipulation was included to test whether children with or without autism found verbs of a particular semantic class easier to use in the task described. Prior research suggests, for instance, that children without autism have particular difficulty processing experiencer-theme verbs, such as *forget*, *love*, and *remember* (e.g. Ambridge, Bidgood, Pine, Rowland, & Freudenthal, 2016; see Ambridge et al., 2020, p. 4, for overview). Ambridge et al. (2020) report identifying this effect among children both with and without autism. However, due to focussed theoretical interest in the rate of reversal errors in the current replication, the verb type manipulation does not form part of the current analysis or write up, where we instead home in on the main effects of response type by group (though see R code for additional analyses).

Table 2. Twenty-four prime verbs, with available age-of-acquisition and imageability ratings from Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012) and Bird, Franklin, and Howards (2001).

Verb	AOA (Kuperman et al., 2012): Years	AOA (Bird et al., 2001): 1-7 scale*	Imageability (Bird et al., 2001): 1-7 scale
Avoid	8.50	4.22	3.40
Bite	3.58		
Call	4.74	2.54	4.21
Carry	5.16		
Chase	5.53	2.82	5.29
Cut	4.43		
Dress	4.05	2.31	
Drop	3.26	2.31	
Eat	2.78	1.67	
Follow	5.11	2.91	
Help	3.65	2.69	4.05
Hit	4.75	2.30	
Hold	4.67		
Hug	2.58	2.45	
Kick	4.06	2.43	
Kiss	3.61		
Lead	6.76		
Pat	5.07	2.42	
Pull	4.79		
Push	4.26	2.39	
Shake	5.26	2.84	
Squash	6.94		
Teach	4.67	3.04	
Wash	4.00	1.95	5.84

* 1 = 0-2 years; 2 = 3-4 years; 3 = 4-5years; 4 = 6-7 years; 5 = 9-10 years; 6 = 11-12years; 7 = 13 years or older.

Table 3. Thirty-six target verbs, with available age-of-acquisition and imageability ratings from Kuperman, Stadthagen-Gonzalez and Brysbaert (2012) and Bird, Franklin and Howards (2001).

Verb	AOA (Kuperman et al., 2012): Years	AOA (Bird et al., 2001): 1-7 scale*	Imageability (Bird et al., 2001): 1-7 scale
Amaze	7.50	3.83	4.92
Annoy	7.22	3.11	4.57
Bite	3.58		
Bother	6.50	3.36	3.52
Carry	5.16		
Chase	5.53	2.82	5.29
Dress	4.05	2.31	
Forget	4.78	3.25	3.36
Frighten	8.83	2.86	
Hate	5.53	3.33	3.95
Hear	3.80	2.53	
Hit	4.75	2.30	
Hug	2.58	2.45	
Ignore	6.74	4.30	
Impress	10.17		
Kick	4.06	2.43	
Know	4.50	2.75	
Like	3.69	2.49	3.32
Love	5.17	2.51	5.03
Pat	5.07	2.42	
Please	3.48		
Pull	4.79		
Push	4.26	2.39	
Remember	5.63	3.27	3.91
Scare	4.22		
See	3.06	2.39	
Shock	7.53	4.13	4.13
Smell	4.22	2.41	
Squash	6.94		

Table 3 continued. Thirty-six target verbs, with available age-of-acquisition and imageability ratings from Kuperman, Stadthagen-Gonzalez and Brysbaert (2012) and Bird, Franklin and Howards (2001).

Verb	AOA (Kuperman et al., 2012): Years	AOA (Bird et al., 2001): 1-7 scale*	Imageability (Bird et al., 2001): 1-7 scale
Surprise	5.47		
Tease	5.11		
Understand	6.17	3.94	3.40
Upset	5.26	3.29	4.16
Wash	4.00	1.95	5.84
Watch	4.33		
Worry	6.61	4.15	4.76

* 1 = 0-2 years; 2 = 3-4 years; 3 = 4-5years; 4 = 6-7 years; 5 = 9-10 years; 6 = 11-12years; 7 = 13 years or older.

Children's responses were coded using the regime described in Ambridge et al. (2020; pp. 6–7), and touched on in the introduction to the current study. Given an animation in which Wendy scared Bob, for instance, a response of *Bob was scared by Wendy* was coded as a 'correct passive'; a response of *Wendy was scared by Bob* was coded as an 'incorrect passive' (i.e., a reversal error – the response of primary interest); a response of *Wendy scared Bob* was coded as a 'correct active'; and responses such as *scared Bob* were coded as 'other use of target verb'. Responses outside of these four categories were excluded from the analysis. Given only two incorrect active responses among participants, this category was excluded from all statistical analyses, as it was excluded in the original study.

Statistical analysis

A series of maximal Bayesian hierarchical models were fitted using the brms package in R (Bürkner, 2018; R Core Team, 2016). In each model, response type (i.e., correct active, correct passive, incorrect passive, and other verb) was predicted by group (i.e., non-autism, autism) and WPPSI score, with target sentence (i.e., verb) and participant as grouping variables. In brms syntax:

Model = brm(formula = Response ~ Group + WPPSI +
(1 + Group + WPSSI | *Target sentence) +*
(1 | Participant)

One additional model was fitted with identical fixed and random effects and total passive responses (i.e., correct plus incorrect passives) as the dependent variable. The purpose of this model was to determine whether groups produce similar rates of passive response overall. Following Ambridge et al. (2020), we set a conservative prior of 2.77 on beta (β; see R code for detailed model specification, and p. 7 of the original study for the justification of prior). These models were fitted not only to our replication data ($N = 26$) but also to the original data ($N = 30$) and to the pooled data ($N = 56$; i.e., the original and replication data combined). Each model fitted well, as indicated by rhat values uniformly at one and credible posterior predictive visualisation checks (see brms package documentation for details of diagnostics; Bürkner, 2018). We believe the model specifications used here and indeed in the original target study to be well justified. However, researchers keen to test different configurations of, for instance, prior or random or fixed effects are invited to do so using our data and code. Note that we switched coding of non-autism and autism groups relative to the original study.

Across the analyses presented in this paper, then, children without autism form the baseline group, rather than children with autism. This allows readers to see more clearly the associations between a diagnosis of autism and the likelihood of giving a response of a certain type. Note also that we do not follow Ambridge et al.'s (2020) approach of calculating *pMCMC* values, or 'Bayesian *p*-values', but rather report a combination of proportional odds and 90% highest density intervals (HDIs), i.e., the most credible 90% span of the posterior distribution. This broadly follows the approach outlined in McElreath (2016; though McElreath uses narrower 89% HDIs – the choice is arbitrary), which we believe to provide an intuitive method of communicating results and propagating uncertainty in the data. Readers who disagree are welcome to calculate *pMCMC* values or conduct alternative analyses (e.g. using 89% or 95% HDIs) using our data and R code.

The results that follow can be interpreted in the following way. A HDI bound above zero (e.g. 0.2 to 0.5) suggests a positive association between variables (e.g. a diagnosis of autism and higher rates of reversal error). A HDI bound below zero (e.g. -0.8 to -0.3) suggests a negative association between variables (e.g. a diagnosis of autism and lower rates of accurate passive responses). And a HDI spanning zero (e.g. -0.3 to 0.4) suggests no linear relationship between predictors is plausible (i.e., no difference between children with and without autism with respect to a particular response).

Results

A by-participant summary of the results can be found in the Appendix. As this table shows, at least one reversed passive was produced by 6/13 children with autism and 2/13 children without autism. Descriptive statistics of task performance are presented for reference in Table 4. Importantly, the correct passive and incorrect passive columns of Table 4 provide evidence of a priming effect. Every animation in the task could have been described accurately using the cue verb in an active sentence. However, children both with and without autism appeared to be primed to some extent by the experimenter's example sentence and used passive syntax to describe animations in 31.45% of trials overall (i.e., 256 passives out of 814 total responses), despite passive syntax being low frequency in everyday speech.

Table 4. Performance (mean, with standard deviation in brackets) in the original study (Ambridge, Bidgood, & Thomas, 2020; ABT) and current replication study (Jones, Dooley, & Ambridge, 2020; JDA) by group (WA= Without Autism; ASC = autism spectrum condition).

Study	Group	Correct active	Incorrect active	Correct passive	Incorrect passive	Other verb
ABT	WA	3.13 (1.67)	0.09 (0.29)	0.91 (1.2)	0.18 (0.44)	0.33 (0.56)
ABT	ASC	2.2 (1.82)	0.07 (0.25)	0.91 (1.33)	0.77 (1.29)	0.55 (0.79)
JDA	WA	3.58 (1.89)	0.0 (0.0)	2.1 (1.97)	0.02 (0.16)	0.08 (0.27)
JDA	ASC	2.62 (1.57)	0.05 (0.32)	1.13 (1.54)	0.33 (0.62)	0.38 (0.63)

Prior to our main analysis, we tested whether groups were similarly likely to produce passive sentences overall, i.e., correct passives and reversal errors combined. The results of this analysis are presented in Table 5, which shows estimates and 90% HDIs for the original (N = 30), replication (N = 26), and pooled (N = 56) data. In the original study it was reported on the basis of descriptive statistics (i.e., no model was fitted) that children with autism were more likely to produce passive sentences than children without autism. While the Bayesian analysis of the original data implies this effect (estimate = 0.11), we note that the 90% HDI for this estimate crosses zero (HDI = -0.05, 0.25), indicating that the true effect may be practically null. In the replication data, the estimate suggests children with autism were in contrast less likely to produce passive sentences than children without autism (estimate = -0.13), however the 90% HDI for this estimate again suggests that the effect may not be substantial (HDI = -0.29, 0.01). Analysis of the pooled data indicates the absence of any group effect on the production of passive sentences (estimate = 0.02, HDI = -0.08, 0.13). Overall, then, children with and without autism were equally likely to respond using passive syntax. Children with autism produced passives in 131 out of 374 responses (i.e., 35.03%), while children without autism produced passives in 134 out of 440 responses (i.e., 30.45%).

Table 5. Estimates and 90% highest density intervals (HDI) for the association between a diagnosis of autism and a passive response.

Data	Estimate	90% HDI
Original	0.11	-0.05, 0.25
Replication	-0.13	-0.29, 0.01
Pooled	0.02	-0.08, 0.13

We then looked at rates of reversal error. Analyses of the original ($N = 30$), replication ($N = 26$), and pooled ($N = 56$) data indicate that children with autism were more likely to make reversal errors than children without autism (Table 6; pooled HDI = 1.06, 4.21). Overall, 47 out of 374 responses made by children with autism contained reversal errors (i.e., 12.57%), while just 9 out of 440 responses made by children without autism contained reversal errors (i.e., 2.05%). In the pooled analysis, the beta coefficient for the association between a diagnosis of autism and the production of a reversal error was = 2.59. Exponentiating this estimate shows that, while groups produced a comparable number of passives in general (Table 5), the proportional odds of a child with autism mis-assigning thematic roles and producing a reversal error were approximately thirteen times (13.33) higher than the odds of a child without autism doing likewise.

Table 6. Estimates and 90% highest density intervals (HDI) for the association between a diagnosis of autism and reversal errors.

Study	Estimate	90% HDI
Original	2.11	-0.52, 4.50
Replication	2.67	-0.39, 6.03
Pooled	2.59	1.06, 4.21

Next, we looked at whether children with autism were more or less likely than children without autism to respond using correct actives (Table 7). Analysis re-confirmed that in the original study children with autism were less likely than children without autism to produce correct actives (HDI = -1.99, -0.12). However, replication and data pooling indicate a density interval spanning zero (pooled HDI = -1.20, 0.01). Overall, then, it is not clear that children without autism produced substantially more correct active responses than children with autism. The number of correct active responses made by children in each group was high. Overall, 199 out of 374 responses made by children with autism were correct actives (i.e., 53.21%), while 284 out of 440 responses made by children without autism were correct actives (i.e., 64.55%).

Table 7. Estimates and 90% highest density intervals (HDI) for the association between a diagnosis of autism and correct actives.

Study	Estimate	90% HDI
Original	-1.05	-1.99, -0.12
Replication	0.12	-0.75, 0.98
Pooled	-0.58	-1.20, 0.01

Finally, we looked at whether children with autism were more or less likely than children without autism to respond with an alternative use of the target verb. For instance, responding *Wendy pulling Bob* where the target passive sentence was *Bob was pulled by Wendy*[1]. The results of these analyses are shown in Table 8. Estimates and HDIs indicate that children with autism were consistently more likely than children without autism to use the target verb in a response other than the correct active or a passive.

Table 8. Estimates and 90% highest density intervals (HDI) for the association between a diagnosis of autism and other uses of the target verb.

Study	Estimate	90% HDI
Original	1.60	-0.11, 3.37
Replication	2.08	-1.11, 5.07
Pooled	2.47	0.88, 4.06

In the pooled data, 39 out of 374 responses made by children with autism involved an alternative use of the target verb (i.e., 10.43%), while 18 out of 440 responses made by children without autism involved an alternative use of the target verb (i.e., 4.09%). We note that many of these responses were reasonable. For instance, the response of *Homer was annoying Marge* instead of the expected target *Marge was annoyed by Homer*; the response of *Wendy was letting Bob pat her* instead of *Wendy was patted by Bob*; and the response of *Marge is carrying Homer* instead of *Homer was carried by Marge*.

---

1 As these examples show, this response category includes both grammatical and ungrammatical uses of the target verb. Of the 17 responses in this category, ten were fully grammatical, two (both produced by children without autism) included a past-tense overgeneralization error (*bited,* in both cases), and five were unclear. These were all cases such as *Marge remembering Homer* which is ungrammatical as a standalone sentence, but which could be acceptable as a response to an implicit question such as *What can you see in this video?*.

Discussion

The language of children with autism varies dramatically, from children who have little or no language to children who have advanced language skills and may appear pedantic or verbose (Norbury, 2014). While as a group children with autism often use shorter and grammatically simpler sentences than children without autism (Eigsti et al., 2007), it has been argued that the main areas of language difficulty for children with autism are semantics, pragmatics, and narrative, rather than 'pure syntax' (Naigles & Tek, 2017). The current study aimed to tease apart these effects through a replication of work by Ambridge et al. (2020). These authors asked 30 children aged 6-9 years, with and without autism, to describe a series of animations using a cue verb, primed by the experimenter to use passive syntax. The response of primary interest was the rate of reversal errors, in which passive syntax is used accurately but thematic roles are mis-assigned (e.g. the child describes an animation in which Wendy [AGENT] surprises Bob [PATIENT] with the phrase *Wendy* [PATIENT] *was surprised by Bob* [AGENT]). Ambridge et al. (2020) report a higher rate of reversal errors among children with autism than among children without autism.

We set out to test whether this effect replicated in a new sample of children with and without autism (N = 26) and presented a cumulative analysis in which data from the original study and the replication were pooled (N = 56). Analysis indicated that the main effect reported by Ambridge et al. (2020) replicated in this new sample of children. Table 5 of the current study shows that children with autism were in general as likely as children without autism to produce passive sentences. However, the groups differed substantially in the rate of reversal errors they made, with children with autism approximately thirteen times more likely than children without autism to make an error in thematic role assignment, for instance describing a scene in which Wendy surprised Bob using the phrase *Wendy was surprised by Bob* (Table 6). Results corroborate Ambridge et al.'s (2020) conclusion that despite age-appropriate knowledge of (at some level) constituent order in passive syntax, the ability of certain children with autism to map syntax to thematic roles is impaired.

Embedding the cue verb in an accurate passive sentence was clearly challenging for children both with and without autism, due to their young age and the high complexity and low frequency of this syntactic structure. This was reflected in the high rate of 'default' active responses made by children with and without autism (i.e., 53.21% and 64.55% respectively; see Table 7), and the relatively high rate of alternative responses made by children with autism (i.e., 10.43%; see Table 8). The real challenge, of course, is to explain why children with autism produced inaccurate passives in 12.57% of trials (versus 2.05% of trials among children without autism), instead of defaulting to active syntax or responding with an alternative verb usage if task demands were high. Ambridge et al. (2020, pp 15–17) discuss two possibilities. The first is that children with autism struggle to understand the discourse-pragmatic conditions under which typical AGENT-PATIENT order is reversed (e.g., when the PATIENT is topical; *Have you heard the news about YouTube? It got bought by Google*; Pullum, 2014: 64). The second and related possibility is that reversal errors are part and parcel of the same narrative deficit that sometimes causes children with autism to mention characters or events in the wrong order. Both of these possibilities are consistent with the replication and cumulative datasets presented here, which converge on a very similar pattern of results. Rather than re-describe these possibilities, then, we here present an alternative account that nevertheless remains compatible with those summarised in Ambridge et al. (2020).

Under construction-based accounts of language acquisition (e.g., Tomasello, 2003; Dabrowksa, 2004; Goldberg, 2019), children build constructions – including the passive – by analogizing across input utterances that exemplify these constructions. This is true even for those accounts that explicitly retain the original exemplars (e.g., Abbot-Smith & Tomasello, 2006; Ambridge, 2020). For example, suppose that a child without autism hears sentences such as *Chloe was hit by Danny, James was kicked by Billy* and *Sarah was dressed by her Dad*. The assumption is that, on the basis of such utterances, the child forms a construction schema of the form *[PATIENT] [BE] [ACTION] by [AGENT]* (even if only very approximately; Ambridge, 2020). This construction will allow her to produce an appropriate passive sentence such as *Bob was pushed by Wendy* (a target utterance in the present study). Suppose, now, that a child with autism hears sentences such as *Chloe was hit by Danny, James was kicked by Billy* and *Sarah was dressed by her Dad,* but instead forms a construction schema of the form *[PERSON] [BE] [ACTION] by [PERSON]*. This more general construction will allow her to produce both appropriate passive sentences such as *Bob was pushed by Wendy* and (as a description of the same event) incorrect reversed passive sentences such as *Wendy was pushed by Bob*.

This account, as it is presented above, would seem to predict – incorrectly – that children with autism will produce correct and reversed passives at rates of around 50/50. In fact, however, the notion of a child forming either a *[PATIENT] [BE] [ACTION] by [AGENT]* or a *[PERSON] [BE] [ACTION] by [PERSON]* construction is a gross oversimplification. In reality, 'constructions' are probabilistic and multi-facetted: The first slot is neither PERSON nor PATIENT but a probabilistic cluster of all the properties of all of the different entities that have appeared in this position in input utterances (see Ambridge, 2020, for a detailed discussion of how re-representing exemplar utterances at an increasingly abstract level in a computational model results in abstractions that *approximate* – but never map on to entirely – linguistic constructions at various levels of abstraction).

An advantage of this account is that it can potentially also explain the finding of Paul, Fisher, and Cohen (1988) that children with autism make more reversal errors of this type than do IQ-matched controls, when assessed using comprehension methods (though see Tager-Flusberg for a null finding using a similar methodology). But is there any reason to believe that children with autism are more likely than children without autism to form (probabilistically) these overly general constructions? We are not aware of any directly-relevant research evidence, but the possibility is generally consistent with the empathizing-systemizing view of Baron-Cohen and colleagues (e.g., Baron-Cohen, 2009), under which people with autism lie at the more systemizing end of the continuum. Classifying verb arguments as AGENT, PATIENT, EXPERIENCER or THEME might require a degree of empathising, of understanding others' perspectives and emotions. Classifying verb arguments as PERSON does not, and is a more systematic approach, in that it posits a higher level of generalization; that is, of systematicity.

Of course, this possibility is highly speculative at present but could potentially be investigated in future research, for example by investigating whether children with autism make similar errors for other constructions that require human participants to be classified into fine-grained psychological categories like RECIPIENT (e.g., dative/ditransitive constructions). Another potentially illuminating direction for future research would be to replicate the priming task described in this study using animations depicting a mixture of human interactions (e.g., Wendy surprising Bob) and systematic physical processes (e.g. a cam rotating and making a lever move). People with autism and Asperger syndrome are reported to show better understanding of physical systems than people without autism, despite apparent deficits in interpreting human intentions among this population (Lawson, Baron-Cohen, & Wheelwright, 2004). It would be interesting to test, therefore, whether among children with autism the rate of reversal errors would be lower for passive sentences describing systems (e.g., *the cam was moved by the lever*) than for sentences describing human interactions (e.g., *Wendy was surprised by Bob*).

In the pooled analysis presented in this study, the odds of a child with autism producing a reversal error were approximately thirteen times higher than the odds of a child without autism doing likewise. Nevertheless, we noted that children with autism produced reversal errors on only 12.57% of their total responses. Despite substantial proportional odds, then, it may be argued that this modest magnitude on an absolute scale makes the passive reversal effect trivial, particularly considering how rarely passive syntax occurs in natural speech. That is, passive sentences may occur so rarely in natural speech that apparently mild deficits in mapping thematic roles among some children with autism may not cause significant problems in language use. It is important to note, however, that the current study looked at a sample of children with relatively low scores on the SCQ measure of autism (some only a few points above the cut-off of 15). It may well be, therefore, that children with higher scores would produce more reversal errors (or even a different pattern of responses entirely). Determining how patterns of performance in the current paradigm link to specific cognitive profiles will enable us to determine whether the results reported here may guide the fine-tuning of programmes of language support for children with autism. It is likely that the task will need to be modified for use with participants showing different symptomologies.

Conclusion

The current study presented a replication of Ambridge et al. (2020). While children with and without autism produced a similar number of passive responses in general, the responses of children with autism were significantly more likely to include errors in thematic role assignment. Despite age-appropriate knowledge of (at some level) constituent order in passive syntax, the ability of certain children with autism to use word order to appropriately mark thematic roles is impaired.

References

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review, 23,* 275–290. https://doi.org/10.1515/TLR.2006.011

Ambridge, B., Bidgood, A., Pine, J. M., Rowland, C. F., & Freudenthal, D. (2016). Is passive syntax semantically constrained? Evidence from adult grammaticality judgment and comprehension studies. *Cognitive Science*, *40*(6), 1435–1459. https://doi.org/10.1111/cogs.12277

Allen, M. L., Haywood, S., Rajendran, G., & Branigan, H. (2011). Evidence for syntactic alignment in children with autism. *Developmental Science*, *14*(3), 540–548. https://doi.org/10.1111/j.1467-7687.2010.01001.x

Alvares, G. A., Bebbington, K., Cleary, D., Evans, K., Glasson, E. J., Maybery, M. T., ... & Whitehouse, A. J. (2020). The misnomer of 'high functioning autism': Intelligence is an imprecise predictor of functional abilities at diagnosis. *Autism*, *24*(1), 221-232. https://doi.org/10.1177/1362361319852831

Ambridge, B. (2019). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, 014272371986973. https://doi.org/10.1177/0142723719869731

Ambridge, B. (2020). Abstractions made of exemplars or 'You're all right, and I've changed my mind': Response to commentators. *First Language*, *40*(5-6), 640-659. https://doi.org/ 10.1177/0142723720949723

Ambridge, B., Bidgood, A., & Thomas, A. (2020). Disentangling syntactic, semantic and pragmatic impairments in ASD: Elicited production of passives. *Journal of Child Language,* 1–18. https://doi.org/10.1017/S0305000920000215

Baron-Cohen, S. Autism: The empathizing–systemizing (E-S) theory. *Ann. N. Y. Acad. Sci.* 1156, 68–80 (2009). https://doi.org/10.1111/j.1749-6632.2009.04467.x

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *British Journal of Psychiatry, 194*(6), 500–509. https://doi.org/10.1192/bjp.bp.108.059345

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33,* 73–79. https://doi.org/10.3758/BF03195349

Bürkner, P.-C. (2018). Bayesian Regression Models using "Stan." *Journal of Statistical Software*. CRAN repository. https://doi.org/10.18637/jss.v080.i01>

Dąbrowska, E. (2004). *Language, mind and brain: Some Psychological and neurological constraints on theories of grammar*. Edinburgh University Press. http://www.jstor.org/stable/10.3366/j.ctvxcrgdw

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978-990. https://doi.org/10.3758/s13428-012-0210-4

Eigsti, I.-M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders, 37*(6), 1007–1023. https://doi.org/10.1007/s10803-006-0239-2

Goldberg, A. E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton University Press.

Lawson, J., Baron-Cohen, S., & Wheelwright, S. (2004). Empathising and systemising in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 34*(3), 301–310. https://doi.org/10.1023/B:JADD.0000029552.42724.1b

McElreath, R. (2016). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. London: Taylor and Francis. https://doi.org/10.3102/1076998616659752

Naigles, L. R., & Tek, S. (2017). 'Form is easy, meaning is hard' revisited: (Re) characterizing the strengths and weaknesses of language in children with autism spectrum disorder. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(4), e1438. https://doi.org/10.1002/wcs.1438

Norbury, C. F. (2014). Autism spectrum disorders and communication. In L. Cummings (Ed.), *The Cambridge Handbook of Communication Disorders* (pp. 141–157). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139108683.011

Paul, R., Fisher, M., & Cohen, D. (1988). Brief report: Sentence comprehension strategies in children with autism and specific language disorders. *Journal of Autism and Developmental Disorders, 18*, 669–679. https://doi.org/10.1007/BF02211884

Pullum, G. K. (2014). Fear and loathing of the English passive. *Language & Communication, 37*, 60-74. http://dx.doi.org/10.1016/j.langcom.2013.08.009

R Core Team. (2016). R. *R Core Team*. Retrieved from https://www.r-project.org/

Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire: Manual*. Los Angeles, CA: Western Psychological Services.

Sassenhagen, J., & Alday, P. M. (2016). A common misapplication of statistical inference: Nuisance control with null-hypothesis significance tests. *Brain and Language, 162*, 42–45. https://doi.org/10.1016/j.bandl.2016.08.001

Tager-Flusberg, H. (1981). On the nature of linguistic functioning in early infantile autism. *Journal of Autism and Developmental Disorders, 11*(1), 45–56. https://doi.org/10.1007/BF01531340

Tek, S., Mesite, L., Fein, D., & Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(1), 75–89. https://doi.org/10.1007/s10803-013-1853-4

Tomasello, M. (2003). *Constructing a language*. Harvard University Press.
Wechsler, D. (2012).

Wechsler, D. (2012). *WPPSI-IV: Wechsler preschool and primary scale of intelligence* (4th ed.). Bloomington, MN: Pearson, Psychological Corporation.

Data, code and materials availability statement

All data, code, and materials are available at the Open Science Framework repository accompanying this article: https://osf.io/c2pjd/.

Ethics statement

Ethics approval was obtained from the ethics committee of the University of Liverpool.

Authorship and Contributorship Statement

Samuel Jones analyzed the data, wrote the first draft of the manuscript, and revised the manuscript during peer review. Madeline Dooley collected the data. Ben Ambridge conceived of and designed the study, oversaw data analysis, and revised the manuscript prior to and during peer review. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

Samuel Jones and Ben Ambridge are members of the ESRC International Centre for Language and Communicative Development (LuCiD). The support of the Economic and Social Research Council [ES/S007113/1 and ES/L008955/1] is gratefully acknowledged.

Appendix

By-participant SCQ, seven-subset WPPSI-IV, and task scores. WA = without autism; ASC = autism spectrum condition.

Participant	Group	SCQ	Vocab	Blocks	Pictures	Information	Assembly	Zoo	Naming	Age	Correct Passive	Incorrect Passive	Correct Active	Other Verb	Irrelevant
1	WA	NA	20	16	22	13	20	6	21	6.3	0.22	0	0.67	0.06	0
2	WA	NA	20	18	20	18	17	11	17	4.44	0.39	0	0.61	0	0
3	WA	NA	22	18	20	20	22	12	17	5.2	0.67	0.06	0.28	0	0
4	WA	NA	19	18	16	18	17	11	16	4.96	0.06	0.06	0.89	0.06	0
5	WA	NA	24	22	22	20	17	14	17	5.45	0.33	0	0.67	0	0
6	WA	NA	24	20	18	20	20	13	18	5.47	0	0	1	0	0
7	WA	NA	22	18	16	20	22	11	19	4.5	0.22	0	0.78	0	0
8	WA	NA	20	18	15	21	16	13	16	5.47	0.56	0	0.44	0	0
9	WA	NA	22	24	21	20	19	14	17	4.03	0.39	0	0.61	0	0
10	WA	NA	20	18	22	19	22	14	17	4.33	0.56	0	0.44	0	0
11	WA	NA	24	22	17	23	25	12	18	5.47	0.5	0	0.44	0	0.06
12	WA	NA	24	21	17	25	20	13	20	5.34	0.61	0	0.39	0	0
13	WA	NA	20	18	16	18	20	12	16	6.51	0.17	0	0.72	0.06	0.06
14	ASC	23	21	20	16	19	24	11	14	9.43	0	0.06	0.39	0	0.61
15	ASC	24	23	18	20	22	28	12	18	8.45	0.11	0.17	0.28	0.17	0.28
16	ASC	21	23	24	20	18	28	13	18	8.96	0.06	0	0.56	0.11	0.28
17	ASC	23	22	20	18	23	30	12	14	7.79	0.11	0	0.39	0	0.5
18	ASC	21	21	16	18	14	24	13	17	9.07	0.11	0	0.28	0.11	0.5
19	ASC	26	15	20	18	16	24	11	18	8.35	0.33	0.11	0.28	0.11	0.17

Appendix continued

By-participant SCQ, seven-subset WPPSI-IV, and task scores. WA = without autism; ASC = autism spectrum condition.

Participant	Group	SCQ	Vocab	Blocks	Pictures	Information	Assembly	Zoo	Naming	Age	Correct Passive	Incorrect Passive	Correct Active	Other Verb	Irrelevant
20	ASC	29	28	24	26	25	28	14	22	6.63	0.56	0.06	0.33	0	0.06
21	ASC	25	23	18	22	23	19	9	17	7.63	0.11	0.28	0.39	0.06	0.17
22	ASC	22	21	20	23	19	24	14	19	6.65	0.17	0.06	0.5	0	0.28
23	ASC	20	24	26	23	25	28	12	17	8.27	0.06	0	0.61	0.11	0.22
24	ASC	23	21	16	21	18	22	11	13	6.58	0.17	0	0.67	0	0.17
25	ASC	19	25	22	18	23	25	14	19	9.01	0.39	0	0.5	0.06	0.06