

Appendix











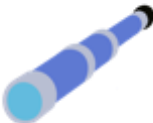



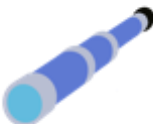

Appendix 1. Groups matching procedure

















Propensity score matching (PSM) was used to create comparable groups of autistic ($N = 73$) and neurotypical children. From an initial pool of $N = 145$ neurotypical children, we matched $N = 73$ individuals to our autistic participants based on relevant covariates, namely biological age, non-verbal IQ (i.e., measured with the Raven's-2) and receptive morphosyntax (i.e., assessed with the TROG). This method allows for more robust comparisons by reducing potential confounding factors and selection bias.









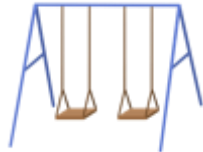
Matching was conducted automatically in R using R Studio (Version 2023.12.0 + 369) with the *MatchIt* package (Ho et al., 2011) using a one-to-one *nearest neighbor* method, in which for each participant of the ASD group, the algorithm searches for the participant with the most similar propensity score. The quality of the matching procedure was assessed using standardized mean differences (SMD) and statistical tests comparing group means. Good balance was achieved between autistic and neurotypical children in non-verbal IQ ($\text{SMD} < .1$, $t(143.88) = -0.58$, $p = .56$), receptive morphosyntax ($\text{SMD} < .15$, $W = 2539.5$, $p > 0.5$) and age ($\text{SMD} < .05$, $W = 2712.5$, $p > .05$). We failed to match groups on their socioeconomic status in addition to age, non-verbal IQ and receptive morphosyntax ($\text{SMD} = 0.42$, $W = 2179.5$, $p < .05$). Therefore, socioeconomic status was added in the statistical model as a covariate.





















Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42, 1–28.
<https://doi.org/10.18637/jss.v042.i08>

Appendix 2. List of items in all languages, with response options. Video examples for an iconic, deictic and conventional gesture in all languages are available following: <https://osf.io/vp764/>

	Sentence	Gesture	Correct response	Gestural/ Visual distractor	Semantic distractor	Oddball
Intro	C'est l'histoire de Louis. This is the story of Lewis. Das ist die Geschichte von Luis. Questa è la storia di Luigi.					
Practice 1	Louis sort de sa maison, et va donner à manger aux... Lewis likes to be outside and to watch animals with... Luis ist gerne draußen und beobachtet gerne die Tiere mit... Luigi ama stare fuori e guardare gli animali con...	CHICKEN				
Practice 1B	Regarde, ce sont des poules ! Look, these are chickens! Schau mal, das sind Hühner! Guarda, queste sono delle galline!	CHICKEN				
Practice 2	Louis aime être dehors et regarder les animaux avec... Lewis likes to be outside and to watch animals with... Luis ist gerne draußen und beobachtet gerne die Tiere mit... Luigi ama stare fuori e guardare gli animali con...	BINOCULARS				
Practice 2B	Regarde, ce sont des jumelles ! Look, they are binoculars! Guck, das ist ein Fernglas! Guarda, sono dei binocoli!	BINOCULARS				

DG 1	<p>Louis va dans la forêt pour ramasser ... [7words] Lewis goes into the forest to collect ... [7words] Luis geht in den Wald und sammelt ... [6 words] Luigi va nella foresta a raccogliere... [6 words]</p>	POINTING flowers				
IG 1	<p>Dans un arbre, il y a un animal. C'est un... [10words] In a tree, there is an animal. It is a... [10words] In einem Baum sitzt ein Tier. Es ist ein ...[9words] In un albero c'è un animale. È un...[8words]</p>	MONKEY				
CG 1	<p>Louis s'arrête. Devant lui, il y a deux route². Louis... [10words] Lewis stops. In front of him, there are two roads. Lewis...[11words] Luis hält an. Vor ihm sind zwei Wege. Luis ...[9words] Luigi si ferma. Davanti a lui ci sono due strade. Luigi... [11words]</p>	DON'T KNOW				
Filler 1	<p>Alors, il va à droite. Then he goes to the right. Dann geht er nach rechts. Quindi va a destra.</p>					
DG 2	<p>Dans le ciel il y a des...[7words] In the sky, there are...[5words] Im Himmel sind ...[3words] Nel cielo ci sono...[4words]</p>	POINTING cloud				

IG 2	<p>Maintenant Louis veut aller à la maison pour jouer... [9words]</p> <p>Now Lewis wants to go home to play...[7words]</p> <p>Jetzt will Luis nach Hause gehen und ... spielen. [8words]</p> <p>Ora Luigi vuole andare a casa a sonare...[8words]</p>	GUITARE			<p>In the Italian version:</p> 	
CG 2	<p>Mais sur le chemin, Louis voit sa copine Lisa. Il...[10words]</p> <p>But on the way, Lewis sees his friend Lea. He...[10words]</p> <p>Aber auf dem Weg sieht Luis seine Freundin Lea. Er ...[10words]</p> <p>Ma sulla strada, Luigi vede la sua amica Lisa. Lui...[10words]</p>	HELLO				
DG 3	<p>Il y a un terrain de jeux. Louis et Lisa jouent sur... [12words]</p> <p>There is a playground. Lewis and Lea play on the... [10words]</p> <p>Es gibt einen Spielplatz. Luis und Lea spielen auf ...[9words]</p> <p>C'è un parco giochi. Luigi e Lisa stanno giocando su...[10words]</p>	POINTING seesaw				
Filler 2	<p>C'est chouette ! Ils s'amuse beaucoup.</p> <p>It's great! They have a lot of fun.</p> <p>Es ist toll! Sie haben viel Spaß.</p> <p>È fantastico! Si stanno divertendo molto.</p>					

IG 3	<p>Lisa s'assoit. Elle prend¹ sa bouteille et elle...[8words]</p> <p>Lea sits down. She takes her bottle and she...[9words]</p> <p>Lisa setzt sich hin. Sie nimmt ihre Flasche und sie ...[10words]</p> <p>Lisa si siede. Prende la sua bottiglia e...[8words]</p>	DRINKING				
CG 3	<p>Il fait bientôt nuit. Louis... [5words]</p> <p>It is almost dark. Lewis... [5words]</p> <p>Es wird bald dunkel. Luis ...[5words]</p> <p>È quasi buio. Luigi... [4words]</p>	COLD				
DG 4	<p>Il veut vite aller... [4words]</p> <p>He wants to go quickly...[5words]</p> <p>Er will schnell gehen.[4words]</p> <p>Vuole andarsene in fretta...[4words]</p>	POINTING home				
IG 4	<p>Alors Louis se dépêche, il...[5words]</p> <p>So Lewis hurries, he...[4words]</p> <p>Luis beeilt sich und...[4words]</p> <p>Così Luigi si sbriga, lui... [5words]</p>	RUN				
CG 4	<p>Louis retrouve sa maman et il dit tout ce qu'il a fait aujourd'hui. La maman de Louis écoute et elle... [7words]</p> <p>Lewis finds his mom and tells her everything he did today.</p> <p>Lewis ' mom listens and she...[6words]</p> <p>Luis sieht seine Mama und erzählt ihr alles, was er heute gemacht hat. Luis' Mutter hört zu und sie... [6words]</p> <p>Luigi ritrova sua madre e le racconta tutto quello che ha fatto oggi. La mamma di Luigi ascolta e... [6words]</p>	CLAP				

Appendix 3. Information on the statistical approach: contrast coding, model checks and assessment.

The group variable was recoded using a scaled sum contrast method: a value of -0.5 was assigned to the typically developing (TD) children group, while a value of 0.5 was assigned to the group of children with autism spectrum disorder (ASD). Consequently, a negative coefficient associated with this variable implied lower performance exhibited by the group of children with ASD.

A scaled sum contrast was also implemented to recode the variable "language of administration" reflecting the language in which the participant completed the task. This approach facilitates the comparison of each language group against the mean overall performance in all languages.

A sliding contrast with weighted coefficients enabled to recode the 'gesture types' categorical variable so as to allow the assessment of relative performance differences: this method compared each gesture type directly, with assigned weights (-0.67, 0.33, and 0.67 for deictic, iconic, and conventional types, respectively).

Missing data on the receptive vocabulary measure (PPVT) and working memory span (FMT) were not imputed, given the robustness of generalized mixed-effects models against missing data (Vasishth et al., 2022). All continuous variables were centered around the mean. None of the variables including age, sex assigned at birth, non-verbal IQ, receptive morphosyntactic score, socioeconomic status, language, and gesture comprehension scores had any missing data: we included in this study only the participants who could complete the tasks mentioned.

Prior to model execution, no significant collinearity among variables was detected, as indicated by variance inflation factors below 5.

A first model with the most complex random structure was fitted, including random intercepts and random slopes for gesture type by *participant* and for *group* by item. As the model did not converge, correlations between random effects were not estimated. As it did not converge either, the random slope for group was removed. Eventually, the final model included random intercepts for both participant and items only. Fixed-effects included: group, gesture type and their interaction, group in interaction with age, as well as sex assigned at birth, socioeconomic status, receptive morphosyntactic skills score, non-verbal IQ, receptive vocabulary score, and language of test. The model demonstrated considerable overall explanatory power (conditional $R^2 = 0.60$), with the fixed effects contributing substantially (marginal $R^2 = 0.42$).

Vasishth, S., Schad, D., Bürki, A., & Kliegl, R. (2022). *Linear mixed models in linguistics and psychology: A comprehensive introduction*. https://vasishth.github.io/Freq_CogSci/

Appendix 4. Fixed and random effects of the final model. *Gtype* holds for Gesture type. *DG*: deictic gesture, *IG*: iconic gesture, *CG*: conventional gesture.

	Performance					
	Odds Ratios	CI	β	Std. Err	Z value	<i>p</i>
(Intercept)	87.78	30.12 – 255.81	4.47	0.55	8.20	< .001***
Group	0.2	0.05 – 0.76	-1.61	0.68	-2.37	0.02 *
Gtype_DG-IG	0.58	0.24 – 1.39	-0.54	0.44	-1.22	0.22
Gtype_IG-CG	1.99	0.81 – 4.93	0.69	0.46	1.49	0.43
Age	2.88	1.56 – 5.32	1.06	0.31	3.37	< .001***
Sex [Reference F]	1.01	0.37 – 2.78	0.01	0.51	0.02	0.98
SES	0.89	0.59 – 1.34	-0.12	0.21	-0.56	0.58
Non-verbal_IQ	0.77	0.47 – 1.25	-0.26	0.25	-1.06	0.29
TROG_score	1.2	0.83 – 1.74	0.19	0.19	0.99	0.32
PPVT_score	1.47	1.07 – 2.01	0.38	0.16	2.39	0.02 *
Lang_English	0.37	0.01 – 15.08	-0.98	1.89	-0.52	0.60
Lang_French	1	0.06 – 16.40	0.00	1.43	0.00	0.99
Lang_German	0.95	0.04 – 22.72	-0.05	1.62	-0.03	0.98
WM_span	1.4	0.84 – 2.34	0.34	0.26	1.29	0.20
Group*Gtype_DG-IG	0.98	0.26 – 3.74	-0.02	0.68	-0.03	0.97
Group*Gtype_IG-CG	0.61	0.14 – 2.61	-0.49	0.74	-0.67	0.50
Group*Age	0.95	0.32 – 2.79	-0.05	0.55	-0.09	0.93

Random Effects

σ^2	3.29
τ_{00} participant	1.38
τ_{00} item	0.15
ICC	0.32
$N_{\text{participant}}$	119
N_{item}	12
Observations	1428
Marginal R^2 / Conditional R^2	0.42 / 0.60

Appendix 5. Complementary mediation analyses

We conducted a causal mediation analysis using R *mediation* package (Tingley et al., 2014) to examine the extent to which vocabulary knowledge (as measured with the PPVT) impacted (i.e., mediated) the ASD and TD group difference in performance.

Two models were fitted on a dataset exempt of missing data: (1) a *mediator* model, examining the impact of group on PPVT scores, controlling for sex assigned at birth, socioeconomic status and the language version of the task, and (2) an *outcome* model, corresponding to the model used in this study (see the Results section of the manuscript, and Appendix 3 and 4 for its specification and output).

Different estimations were then computed: (a) the Average Direct Effect (ADE), capturing the direct effect of group on the performance and excluding the mediation effect; (b) the Average Causal Mediation Effect (ACME), which measures the indirect effect of group on the performance through the mediator (i.e., vocabulary knowledge) and (c) the Total Effect, reflecting the overall relationship between group and performance, entailing both the direct effect and the indirect effect mediated through vocabulary knowledge. Finally, a quasi-Bayesian approximation method with 1000 simulations enabled to estimate confidence intervals and p-values for the 3 different effects (i.e., ACME, ADE, Total Effects).

Results (see Table Appendix 5) indicated that there were both significant direct and mediated effects: On the one hand, vocabulary knowledge significantly mediated the relationship between group and performance: the Average Causal Mediation Effect (ACME) was statistically significant for both children with TD (ACME = 0.0334) and ASD (ACME = 0.0142). On the other hand, the Average Direct Effects (ADE) was significant (ADE = 0.1176 for TD children; ADE = 0.0983 for children with ASD). The Total Effect was significant (Estimate = 0.1318), with approximately 25.96% and 10.81% of the effect that could be attributed to the mediation in the groups of TD children and children with ASD, respectively. On average, approximately 18.39% of the total effect was mediated by vocabulary skills (Estimate = 0.1839).

Table Appendix 5. Causal mediation for vocabulary knowledge (PPVT)

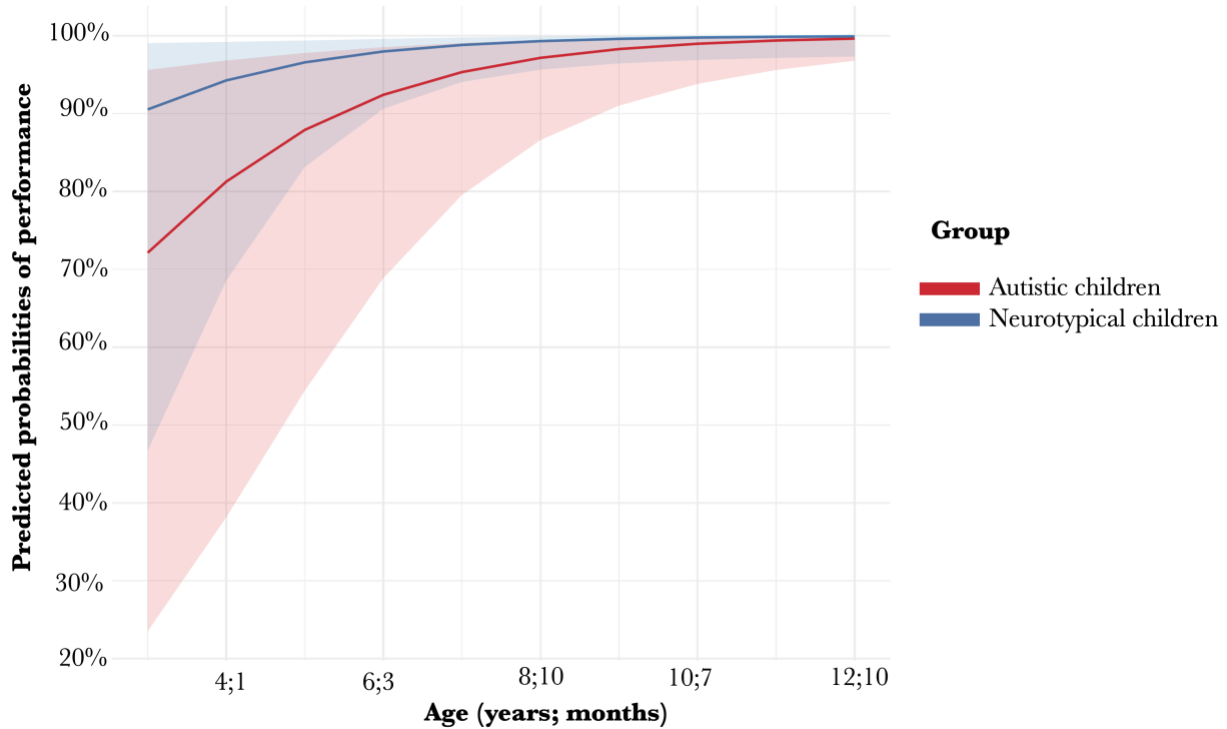
	Group	Estimate	95% CI Lower	95% CI Upper	p value
ACME	(TD)	0.0334	0.0162	0.0600	< .001***
ACME	(ASD)	0.0142	0.0068	0.0200	< .001***
ADE	(TD)	0.1176	0.0445	0.2200	< .001***
ADE	(ASD)	0.0983	0.0366	0.1800	< .001***
Total Effect		0.1318	0.0608	0.2300	< .001***
Prop. Mediated	(TD)	0.2596	0.1465	0.4200	< .001***
Prop. Mediated	(ASD)	0.1081	0.0441	0.2800	< .001***
ACME	(average)	0.0238	0.0126	0.0400	< .001***
ADE	(average)	0.1080	0.0412	0.2000	< .001***
Proportion Mediated	(average)	0.1839	0.0973	0.3500	< .001***

Note. ACME : Average Causal Mediation Effect; ADE: Average Direct Effect; TD: typically developing group, ASD: Autism Spectrum Disorder; CI : Confidence Interval.

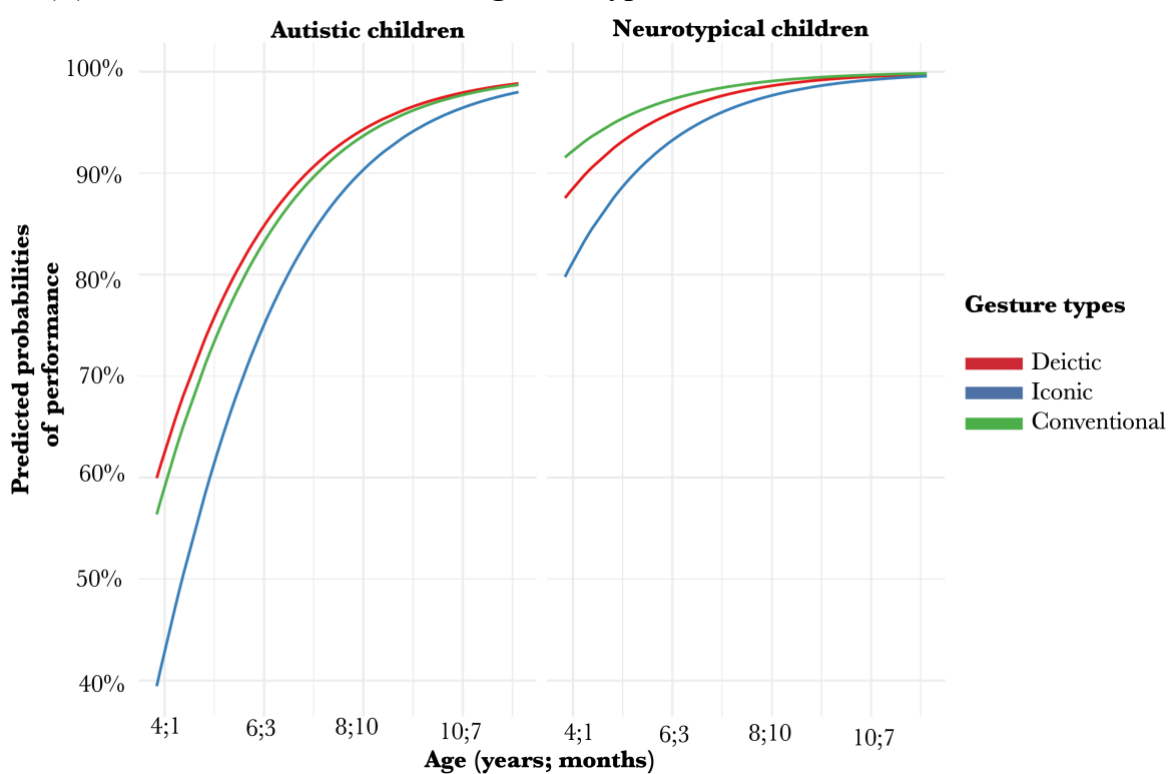
Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). Mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5).
<https://doi.org/10.18637/jss.v059.i05>

Appendix 6. Visualization of the predicted probabilities of performance as a function of age. These visualizations are only indicative and should be observed with caution, as age effects should be investigated more closely with appropriate tools to draw robust conclusions (e.g., with Generalized Additive Models; GAMs).

(A) Overall performance



(B) Performance in the different gesture types



Appendix 7. Exploration of the impact of the multilingual status on performance: GLMM outputs

■ Model A (multilingual status as a fixed effect)

Formula: performance ~ 1 + group * gesture type + group * age + Multilingual status + sex + SES + non-verbal IQ + TROG + PPVT + testing language + Working Memory + (1 | participant) + (1 | item)

	β	Std. Err	Z value	<i>p value</i>
(Intercept)	4.49	0.55	8.19	< .001***
Group	-1.60	0.68	-2.36	0.02*
Gtype_DG-IG	-0.54	0.44	-1.22	0.22
Gtype_IG-CG	0.69	0.46	1.49	0.14
Age	1.05	0.31	3.36	< .001***
Multilingual_status	-0.10	0.45	-0.21	0.83
Sex [Reference F]	-0.01	0.52	-0.02	0.99
SES	-0.12	0.21	-0.59	0.56
Non-verbal_IQ	-0.26	0.25	-1.06	0.29
TROG_score	0.18	0.19	0.95	0.34
PPVT_score	0.39	0.16	2.40	0.02*
Lang_English	-1.00	1.88	-0.53	0.60
Lang_French	-0.02	1.43	-0.01	0.99
Lang_German	-0.12	1.65	-0.07	0.94
Working_memory	0.34	0.26	1.30	0.19
Group: Gtype_DG-IG	-0.02	0.68	-0.04	0.97
Group : Gtype_IG-CG	-0.49	0.74	-0.67	0.51
Group : Age	-0.07	0.55	-0.12	0.90

■ Model B (multilingual status in interaction with *group*)

Formula: performance ~ 1 + *group* * gesture type + group * age + multilingual status * group + sex + SES + non-verbal IQ + TROG + PPVT + testing language + Working Memory + (1 | participant) + (1 | item)

	β	Std. Err	Z value	<i>p</i> value
(Intercept)	4.40	0.54	8.19	< .001***
Group	-1.69	0.70	-2.43	0.02*
Gtype_DG-IG	-0.54	0.44	-1.22	0.22
Gtype_IG-CG	0.69	0.46	1.50	0.13
Age	1.03	0.31	3.31	< .001***
Multilingual_status	0.11	0.49	0.23	0.82
Sex [Reference F]	-0.03	0.51	-0.05	0.96
SES	-0.11	0.21	-0.53	0.60
Non-verbal_IQ	-0.20	0.25	-0.79	0.43
TROG_score	0.12	0.19	0.60	0.55
PPVT_score	0.39	0.16	2.45	0.01**
Lang_English	-0.87	1.86	-0.47	0.64
Lang_French	0.22	1.43	0.15	0.88
Lang_German	0.27	1.68	0.16	0.87
Working_memory	0.30	0.26	1.15	0.25
Group: Gtype_DG-IG	-0.02	0.69	-0.03	0.98
Group: Gtype_IG-CG	-0.50	0.74	-0.68	0.50
Group : Age	-0.14	0.55	-0.26	0.79
Group : Multilingual_status	-1.17	0.99	-1.19	0.24

Appendix 8. Visualisation of the proportions of errors produced by the two groups.

The group of neurotypical children produced a total of 31 errors, while the group of autistic children a total of 106 errors.

