



Twelve- and Fourteen-Year-Old School Children Differentially Benefit from Sensorimotor- and Multisensory-Enriched Vocabulary Training

Brian Mathias^{1,2,3} · Christian Andrä^{4,5} · Anika Schwager⁶ ·
Manuela Macedonia^{3,7,8} · Katharina von Kriegstein^{2,3}

Published online: 11 March 2022
© The Author(s) 2022

Abstract

Both children and adults have been shown to benefit from the integration of multisensory and sensorimotor enrichment into pedagogy. For example, integrating pictures or gestures into foreign language (L2) vocabulary learning can improve learning outcomes relative to unisensory learning. However, whereas adults seem to benefit to a greater extent from sensorimotor enrichment such as the performance of gestures in contrast to multisensory enrichment with pictures, this is not the case in elementary school children. Here, we compared multisensory- and sensorimotor-enriched learning in an intermediate age group that falls between the age groups tested in previous studies (elementary school children and young adults), in an attempt to determine the developmental time point at which children's responses to enrichment mature from a child-like pattern into an adult-like pattern. Twelve-year-old and fourteen-year-old German children were trained over 5 consecutive days on auditorily presented, concrete and abstract, Spanish vocabulary. The vocabulary was learned under picture-enriched, gesture-enriched, and non-enriched (auditory-only) conditions. The children performed vocabulary recall and translation tests at 3 days, 2 months, and 6 months post-learning. Both picture and gesture enrichment interventions were found to benefit children's L2 learning relative to non-enriched learning up to 6 months post-training. Interestingly, gesture-enriched learning was even more beneficial than picture-enriched learning for the 14-year-olds, while the 12-year-olds benefitted equivalently from learning enriched with pictures and gestures. These findings provide evidence for opting to integrate gestures rather than pictures into L2 pedagogy starting at 14 years of age.

Keywords Multisensory learning · Sensorimotor learning · Gesture · Enrichment · Vocabulary learning · Foreign language education

Brian Mathias and Christian Andrä are joint first authors

Extended author information available on the last page of the article

Introduction

Multisensory and Sensorimotor Enrichment

Modern classrooms often make use of multisensory learning materials (Choo et al., 2012; Kiefer & Trumpp, 2012). One reason for doing so is that presence of complementary information across multiple sensory and motor modalities may speed up learning and make it more resistant to decay (Mahmoudi et al., 2012; Sadoski & Paivio, 2013; Shams & Seitz, 2008; von Kriegstein & Giraud, 2006). For example, children tend to benefit more from visual grapheme training when it is integrated with auditory phonological training (reviewed in Ehri et al., 2001). Writing letters by hand can also benefit children's learning above and beyond unisensory visual training (Zemlock et al., 2018). Congruent information presented across two or more sensory modalities during learning has been referred to as *multisensory enrichment* (Mayer et al., 2015), and the combination of body movements with information presented in one or more sensory modalities during learning has been referred to *sensorimotor enrichment* (reviewed in Macedonia, 2014).

Foreign language (L2) learning is one domain that stands to benefit from enriched classroom instruction. One of the most prevalent means of learning L2 vocabulary is students' use of written word lists (Oxford & Crookall, 1990; Schmitt & Schmitt, 2020). However, recent work has suggested that multisensory enrichment can boost L2 vocabulary acquisition. Silverman and Hines (2009), for example, found that the viewing of short video clips that supplemented teachers' regular instruction improved kindergartners' through second graders' acquisition of L2 vocabulary. The video clips were excerpts of documentaries, such as National Geographic's *Really Wild Animals* series (National Geographic Society, 2005), that contained target L2 words. Other studies have suggested benefits of flash cards (Li & Tong, 2019) and pictures paired with audio recordings (Andrä et al., 2020). Further work has provided evidence for benefits of sensorimotor enrichment on L2 vocabulary learning. In one study, children's performance of iconic gestures in tandem with physical exercise while listening to foreign language vocabulary increased recall compared to exercising without gestures (Mavilidi et al., 2015). Holding real objects associated with L2 words during learning has also been shown to benefit children's L2 memory even more than learning with pictures (Bara & Kaminski, 2019). Spanish 10-year-olds' comprehension of stories told in English improved if the instructor enacted gestures during the story-telling (Cabrera & Martínez, 2001), and German high school students' memory for Latin words benefitted from the integration of choral speech and meaningful gestures and movements into the memorization process (Hille et al., 2010). Finally, Macedonia et al., (2014; see also de Wit et al., 2018) demonstrated that 11-year-old children's L2 vocabulary learning outcomes were aided more by performing semantically related gestures themselves during learning than by viewing a pedagogical agent perform the gestures.

Cognitive and Neuroscientific Theories of Multisensory and Sensorimotor Enrichment

Benefits of multisensory and sensorimotor enrichment have been explained in terms of embodied memory for L2 words (reviewed in Atkinson, 2010), dual coding of L2 word representations (Engelkamp & Zimmer, 1985; Hommel et al., 2001; Paivio, 1991; Paivio & Csapo, 1969), mental imagery of multimodally represented L2 words (Jeannerod, 1995; Kosslyn et al., 2006; Saltz & Dixon, 1982), and predictive coding accounts of L2 representations (Mathias et al., 2021a; Mayer et al., 2017; von Kriegstein, 2012). Embodied accounts propose that grounding newly acquired words in sensorimotor experiences allows them to be mentally represented in terms of their perceptual and motor features (Barsalou, 2008; Kiefer & Trumpp, 2012). Dual coding accounts emphasize differences in the encoding of verbal stimuli, which are represented in an auditory code, and nonverbal stimuli, represented in visual or haptic codes. Associations between verbal and nonverbal codes are thought to aid memory retrieval. Mental imagery accounts propose that information encoded during learning is mentally reconstructed at test. Similarly, predictive coding accounts assume that incoming sensory information is processed using an internal generative model, which is capable of reconstructing multimodal representations (Friston & Kiebel, 2009). Listening to an L2 word that has been encoded both in terms of its auditory and visual features, for example, may trigger the reconstruction of its stored visual features, which aid in auditory perceptual recognition (reviewed in Mayer et al., 2015; von Kriegstein, 2012; Yildirim & Jacobs, 2012). A common thread of these accounts is that novel information can be mentally represented in terms of its perceptual and motor features, which may aid learning and memory.

At a neural level, the same sensory and motor brain regions that process visuo-motor enrichment information during learning are causally relevant for subsequent auditory L2 recognition (Mathias et al., 2021a, b; Mayer et al., 2015). The notion that brain regions that support the processing of enrichment also drive enrichment-based learning benefits has been referred to as multisensory learning theory (von Kriegstein, 2012). These studies show that benefits of enrichment on L2 learning are at least in part driven by specific motor and sensory representations that arise from the conditions under which L2 vocabulary was learned, as opposed to more general mechanisms such as enhanced attention or arousal.

Potential Limits of Enrichment Techniques

Benefits of enrichment in the domain of L2 vocabulary learning may be partially limited by the high dependence of semantics on linguistic context. Word meanings often depend on other words with which they co-occur (e.g., the word *bark* in *tree bark* versus *dog bark*; Bergen, 2015). Processes other than multimodal representation such as grammatical constraints on semantics and statistical learning may, like embodiment, shape how language is represented. Memory for abstract L2 words (e.g., *patience*) additionally poses a challenge for theories of enrichment because the

referents of abstract words cannot be perceived by the body's sensory systems (for a review, see Borghi et al., 2017). This is not the case for concrete words (e.g., *tent*). Abstract vocabulary learning typically lags behind concrete vocabulary learning during development (McFalls et al., 1996). However, previous work in children has shown that gesture and picture enrichment can benefit the learning of abstract words as well as concrete words (Andrä et al., 2020), suggesting that abstract semantics, like concrete semantics, may be grounded in perception and action (Harpaintner et al., 2018, 2020).

Comparing the Effectiveness of Multisensory and Sensorimotor Enrichment in Different Age Groups

A key question for the development of evidence-based teaching strategies is whether multisensory enrichment techniques are more (or less) effective than sensorimotor enrichment techniques. This question is of interest in light of growing support for the effectiveness of active learning techniques in educational settings, defined as instructional methods that engage students in the learning process (Drew & Mackie, 2011; Jensen et al., 2015; Michael, 2006; Prince, 2004; Sambanis, 2013). One recent study directly compared effects of a multisensory enrichment technique (learning with pictures) with sensorimotor-enriched learning (learning with gestures) in the context of L2 vocabulary learning (Andrä et al., 2020). In this study, both picture-enriched and gesture-enriched learning enhanced 8-year-olds' free recall and translation of L2 words compared to a unisensory learning baseline condition. Benefits of picture and gesture enrichment were approximately equivalent, even up to 6 months after the L2 vocabulary instruction had ended. This finding in children contrasts with findings in adults in laboratory environments. Adults' L2 vocabulary learning has been shown to benefit more from performing gestures during learning than viewing pictures (Mathias et al., 2021a; Mayer et al., 2015). This effect is particularly pronounced over the long-term (several months post-learning), suggesting that picture-enriched L2 words decay more quickly from memory than gesture-enriched L2 words.

The discrepancy between findings in children and adults with regard to enriched learning strategies suggests that teaching strategies derived from studies on adults may not directly translate into teaching strategies for children or vice versa. Some studies have revealed learning mechanisms that are highly similar across children and adults, such as auditory statistical learning, which remains relatively constant through the course of development (Raviv & Arnon, 2018; Saffran et al., 1999). However, children and adults are also known to differ with regard to several key learning mechanisms, such as their use of working memory (Luna et al., 2004) and deployment of visual and motor imagery (Frick et al., 2009; Funk et al., 2005; reviewed in Gabbard, 2009). Differences in enrichment effects for different age groups also have immediate implications for evidence-based teaching techniques, as gestures and other sensorimotor-based interventions may be more challenging for educators to integrate into pedagogy than multisensory-based interventions.

Aims and Hypotheses of the Current Study

In the present study, we compared multisensory- and sensorimotor-enriched learning in an age group that fell between the age groups tested in previous studies (elementary school children and young adults). The sample included both 12-year-olds (sixth graders) and 14-year-olds (eighth graders) who were all currently enrolled in their first semester of learning Spanish as a foreign language. We opted to include 12-year-olds and 14-year-olds because, in the German education system where the study was conducted, instruction in a second foreign language typically begins in grade six and instruction in a third foreign language typically begins in grade eight. This feature of the German school system allowed us to control for prior instruction in the selected L2 across children of different ages. Both age groups were therefore enrolled in their first semester of Spanish as a foreign language and had received no prior Spanish instruction.

Our aim was to test whether differences in effects of multisensory (picture) and sensorimotor (gesture) enrichment previously observed in adults (Mathias et al., 2021a; Mayer et al., 2015), but not in elementary school children, occur for this intermediate-aged group of high school children. We hypothesized that, if high school children are more similar to elementary school children (Andrä et al., 2020) in terms of their response to picture and gesture enrichment, then we would observe no differences between effects of the two learning conditions. However, if high school children are more similar to young adults (Mathias et al., 2021a; Mayer et al., 2015), then we would observe a greater benefit of gesture enrichment compared to picture enrichment. A third possibility was that the pattern of enrichment effects might diverge across age groups, i.e., 12-year-olds would show equivalent picture and gesture benefits, and 14-year-olds would show a greater gesture than picture benefit. None of these three possible outcomes was favored more or less than any of the others. Additionally, though the current study tested 12- and 14-year-olds due to German educational norms, our hypotheses could have been tested with children from any number of possible age groups.

Besides testing our main hypotheses outlined above, we expected three further effects that have already been shown in adults (Macedonia & Knösche, 2011; Mayer et al., 2015, 2017; Repetto et al., 2017) and elementary school children (Andrä et al., 2020). First, we expected that high school-aged children would demonstrate benefits of picture- and gesture-enriched learning compared to non-enriched (auditory-only) learning. Second, we expected the beneficial effects of picture and gesture enrichment to persist over long time scales (up to 6 months following learning; Andrä et al., 2020; Mayer et al., 2015). We therefore tested the high school children's knowledge of the enriched vocabulary at three different time points: 3 days, 2 months, and 6 months post-learning. Finally, we expected that both picture and gesture enrichment would benefit high school children's learning of both concrete (e.g., *tent*) and abstract words (e.g., *patience*) compared to non-enriched learning.

Methods

Participants

Participants were school children enrolled in Spanish foreign language courses at three public high schools located in the vicinity of Chemnitz, Germany. Forty-eight children were enrolled in grade 6 (12- to 13-year-olds) and 47 children were enrolled in grade 8 (14- to 15-year-olds). Regardless of their grade level (grade 6 or grade 8), all children were currently enrolled in their first course of Spanish as a foreign language and had not previously received any Spanish language training or lessons. Written informed consent was obtained from the legal guardians of all school children who participated. The investigators briefed the children and their teachers on the study procedures in an introductory session that took place prior to the experiment. Children who were absent from at least one training or test session were excluded from the analyses. Therefore, the analyses included 39 children in grade 6 (M age = 12.8 years, SD = 0.4 years, 20 females) and 36 children in grade 8 (M age = 14.8 years, SD = 0.4 years, 27 females). Based on the teachers' reports, none of the children possessed learning disabilities, and all of the children possessed normal or corrected-to-normal vision. Two of the children in grade 8 and none of the children in grade 6 spoke another language besides German and English. A power analysis based on an enrichment effect size of 0.33 (Andrä et al., 2020, Experiment 2), an alpha level of 0.05, and a desired power level of 0.8 suggested a minimum total sample size of $N=52$ participants. The study was reviewed and approved by the Education Department of the state of Saxony, Germany.

In the German education system, instruction in a second foreign language typically begins in grade 6 (approximately 12 years of age) and instruction in a third foreign language typically begins in grade 8 (approximately 14 years of age). However, a given L2 is typically not offered as both a second and third foreign language within the same school. In the case of the three public high schools included in the current study, Spanish was offered as a second foreign language at school 1 and as a third foreign language at schools 2 and 3. Following exclusions due to absences during training or test sessions, the analyses included 39 children (sixth graders) from school 1, 16 children (eighth graders) from school 2, and 20 children (eighth graders) from school 3. All three schools are located within 35 km of each other in towns with between 10,000 and 20,000 inhabitants of middle socioeconomic status.

Stimulus Materials

Spanish words used in the experiment were selected in consultation with the children's school teachers at each of the three high schools. Word selection was based on three factors: First, children had not yet encountered the words in lessons and the words were not anticipated to be included in the teaching curriculum for the 6-month duration of the investigation. Second, the words were considered by the teacher to be relevant for future use by the children. Third, words

were among the 90 words included in the “Vimmi” language corpus (Macedonia et al., 2010, 2011). The Vimmi corpus was created for experiments on L2 learning and contains videos of gestures designed to convey the meanings of words included in the corpus. This resulted in one set of 24 Spanish words for each of the three high schools, shown in Table 1.

Half of the words used at each school were concrete nouns and the other half were abstract nouns. Concreteness and imageability ratings (on a 0 to 10 scale) derived from a corpus of German lemmas (Köper & Schulte im Walde, 2016) are shown in Table 2. Imageability refers to the ease with which a word gives rise to a sensorimotor mental image (Paivio, 1971). *T* tests revealed significantly higher concreteness and imageability ratings for the concrete words compared to abstract words for each of the three schools (all *ps* < 0.001; Table 3). Abstract and concrete word frequencies in written German (<http://wortschatz.uni-leipzig.de/en>) did not significantly differ, as shown in Table 3.

The experiment made use of three stimulus types: audio recordings of Spanish words and their German translations, pictures depicting word meanings, and videos of an actress performing gestures that were semantically related to word meanings. Audio recordings of German words, as well as picture and video stimuli, were adopted from the Vimmi corpus (Macedonia et al., 2010, 2011; Mayer et al., 2015).

The German word recordings featured a female bilingual Italian-German speaker (age 44). Recordings of Spanish translations featured a female native speaker of European Spanish (age 25). Recordings were made using a RØDE NT55 microphone (RØDE Microphones, Silverwater, Australia) in a sound-dampened room.

The pictures consisted of black-and-white line drawings created by a professional cartoon artist (<https://www.klaus-pitter.com/>). The drawings iconically communicated word meanings by depicting objects, humans, or scenes. Abstract nouns were conveyed using scenes. Pictures representing one of the concrete nouns and one of the abstract nouns are shown in Fig. 1. The complexity of line drawings was not matched for concrete and abstract nouns, as differences in complexity are also expected to occur in naturalistic teaching settings.

Videos were recorded using a Canon Legria HF S10 camcorder (Canon Inc., Tokyo, Japan). Each video was 4-s long and shot in color. The actress shown in the videos began and ended each video by standing motionless with her arms by her sides. During the videos, she used head movements, movements of one or both arms or legs, fingers, or combinations of these body parts to convey the meaning of the foreign language word through the movement. For example, the word *tent* was conveyed by moving the arms and fingers together to form an upside-down “V” shape, and the word *patience* was conveyed by lifting up the arms and subsequently slowly moving them outward from the body and downward (Fig. 1). The actress always maintained a neutral facial expression. Gestures selected for abstract nouns were previously agreed upon by three independent raters (Macedonia et al., 2011; Mayer et al., 2015).

Table 1 Vocabulary used at each of the three high schools included in the experiment. Twenty-four German and Spanish words were used at each high school. English translations are also shown. Assignment of words to the gesture enrichment, picture enrichment, and no enrichment conditions was counterbalanced across participants at each school, ensuring that each German and Spanish word was represented equally in each of the learning conditions

High School 1			
Concrete nouns		Abstract nouns	
German	Spanish	German	Spanish
Ampel	Semáforo	Befehl	Orden
Briefkasten	Buzón	Bitte	Ruego
Denkmal	Monumento	Empfehlung	Recomendación
Fernbedienung	Telemando	Gedanke	Pensamiento
Flugzeug	Avión	Korrektur	Corrección
Handtasche	Bolso de mano	Mentalität	Mentalidad
Kabel	Cable	Mut	Valor
Koffer	Maleta	Talent	Talento
Schlüssel	Llave	Teilnahme	Participación
Straßenbahn	Tranvía	Tradition	Tradición
Tageszeitung	Diario	Veränderung	Cambio
Zelt	Tienda	Warnung	Advertencia
High School 2			
Concrete nouns		Abstract nouns	
German	Spanish	German	Spanish
Anhänger	Colgante	Absage	Negativa
Bildschirm	Pantalla	Anforderung	Exigencia
Briefkasten	Buzón	Aufwand	Esfuerzo
Faden	Hilo	Besitz	Propiedad
Kabel	Cable	Bestimmung	Destinación
Katalog	Catálogo	Geduld	Paciencia
			English
			Cancellation
			Requirement
			Effort
			Property
			Destination
			Patience

Table 1 (continued)

Maske	Máscara	Mask	Gleichgültigkeit	Indiferencia	Indifference
Reifen	Neumático	Tire	Rücksicht	Respeto	Consideration
Ring	Anillo	Ring	Tatsache	Hecho	Fact
Sammlung	Colección	Collection	Vorwand	Pretexto	Excuse
Verband	Vendaje	Bandage	Wohltat	Beneficio	Benefaction
Zelt	Tienda	Tent	Zulassung	Admisión	Admission
High School 3					
Concrete nouns					
German	Spanish	English	Abstract nouns	Spanish	English
Ampel	Semáforo	Traffic light	German	Negativa	Cancellation
Anhänger	Colgante	Trailer	Absage	Exigencia	Requirement
Balkon	Balcón	Balcony	Anforderung	Esfuerzo	Effort
Bildschirm	Pantalla	Monitor	Aufwand	Destinación	Destination
Faden	Hilo	Thread	Bestimmung	Paciencia	Patience
Kabel	Cable	Cable	Geduld	Indiferencia	Indifference
Katalog	Catálogo	Catalog	Gleichgültigkeit	Valor	Bravery
Maske	Máscara	Mask	Mut	Respeto	Consideration
Reifen	Neumático	Tire	Rücksicht	Sensibilidad	Sympathy
Schubblade	Cajón	Drawer	Verständnis	Pretexto	Excuse
Verband	Vendaje	Bandage	Vorwand	Beneficio	Benefaction
Zigarette	Cigarrillo	Cigarette	Wohltat	Admisión	Admission
			Zulassung		

Table 2 Concreteness and imageability ratings of the words used in the experiment (derived from Köper & Schulte im Walde, 2016)

Concrete nouns			Abstract nouns		
Word	Concreteness	Imageability	Word	Concreteness	Imageability
Airplane	7.7	7.6	Admission	3.2	4.2
Balcony	7.0	7.0	Benefaction	2.1	3.5
Bandage	5.2	5.2	Bravery	3.0	4.4
Cable	6.5	6.6	Cancellation	3.4	3.6
Catalog	5.0	6.0	Change	2.9	2.9
Cigarette	7.9	7.4	Command	4.0	3.8
Collection	4.1	4.5	Consideration	2.4	2.8
Newspaper	6.3	6.9	Correction	4.4	3.4
Key	6.2	6.0	Destination	2.5	2.9
Letter box	6.9	6.1	Effort	2.2	2.5
Mask	6.4	6.6	Excuse	3.4	3.8
Memorial	5.8	5.9	Fact	2.0	1.9
Monitor	6.7	6.5	Indifference	2.1	3.5
Purse	7.7	7.1	Mentality	1.9	2.6
Remote control	6.2	5.7	Participation	3.7	3.8
Ring	7.1	7.1	Patience	2.1	3.5
Screen	6.1	5.5	Plea	4.6	4.3
Suitcase	7.5	7.9	Property	3.1	4.1
Tent	7.6	7.9	Recommendation	3.3	3.0
Thread	6.2	6.0	Requirement	1.8	2.3
Tire	6.8	6.0	Sympathy	1.8	2.5
Traffic light	6.7	6.5	Talent	3.4	3.7
Trailer	5.7	5.5	Thought	2.9	3.8
Tram	6.6	7.2	Tradition	2.3	2.7
			Warning	3.4	3.7

Design

The experiment had a $3 \times 3 \times 1 \times 2$ mixed design with within-participants factors learning condition (gesture enrichment, picture enrichment, no enrichment), testing time point (3 days, 2 months, and 6 months post-learning), and word type (concrete, abstract), and between-participants factor grade level (grade 6, grade 8). All factors other than the participants' grade level were within- rather than between-participant factors in order to allow for within-participants comparisons of performance in the enriched learning conditions relative to the unisensory (auditory-only) learning condition. If the test type factor, for example, were a between-participants factor, individual differences between participants in vocabulary learning outcomes could mask differences in outcomes between learning conditions.

Table 3 Concreteness ratings, imageability ratings, and frequencies for the concrete and abstract German words used in the experiment at each of the three high schools. $df=22$ for all t tests. *** $p < .001$ **High School 1**

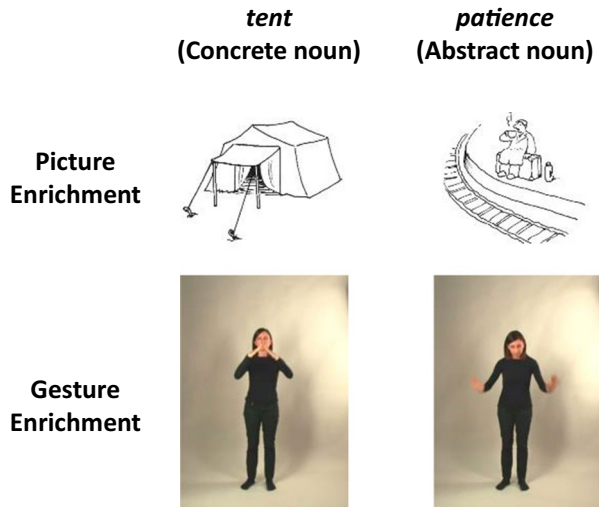
	Concrete nouns		Abstract nouns				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Concreteness	6.2	1.0	2.5	.6	10.90	< .001***	4.5
Imageability	6.2	.9	3.2	.7	8.92	< .001***	3.7
Frequency	11.0	1.2	11.2	1.2	.34	.74	.2

High School 2

	Concrete nouns		Abstract nouns				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Concreteness	6.4	.8	2.5	.6	13.42	< .001***	5.6
Imageability	6.2	.7	3.3	.7	10.78	< .001***	4.3
Frequency	11.1	1.1	11.2	1.2	.18	.86	.1

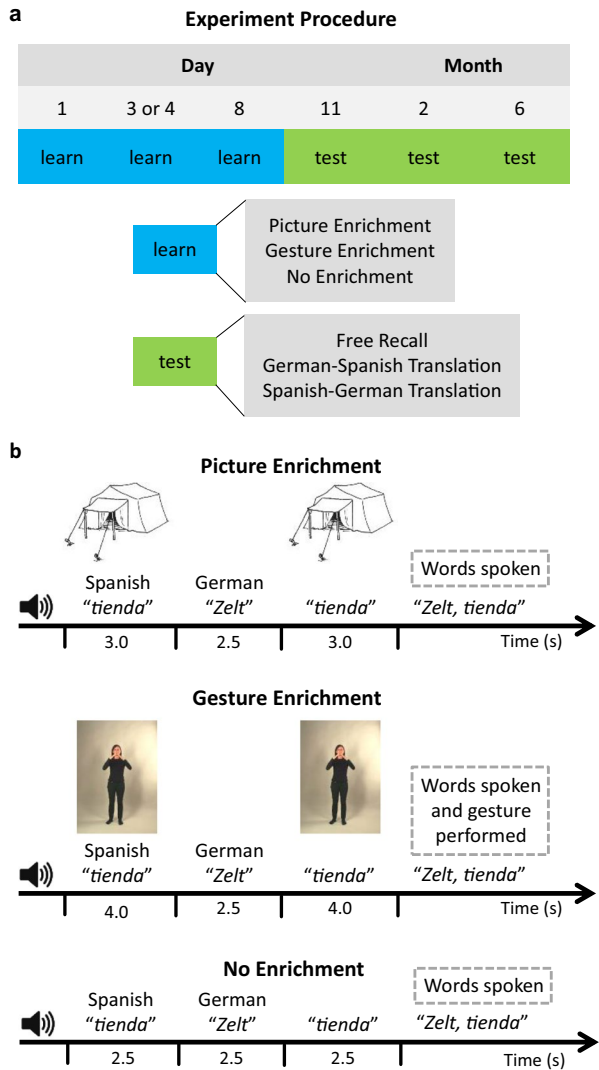
High School 3

	Concrete nouns		Abstract nouns				
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Concreteness	6.8	.7	3.3	.8	11.68	< .001***	5.1
Imageability	6.8	.8	3.5	.6	11.63	< .001***	4.9
Frequency	11.7	1.0	11.2	.9	1.27	.22	.5

Fig. 1 Picture and gesture stimuli. Top: Pictures used in the picture enrichment condition for one of the concrete nouns (*tent*) and one of the abstract nouns (*patience*). Bottom: Screen captures from the corresponding videos of the actress performing gestures, which were used in the gesture enrichment condition**Procedure**

Learning Phase Children completed L2 vocabulary training that took place over a period of 8 days (Fig. 2a). Training was integrated within children's regular Spanish course meetings, and therefore took place on day 1 for 90 min, day 3 or 4 for 45 min,

Fig. 2 Experimental procedure and design. **a** The learning phase of each experiment occurred over 8 days (“learn”). Free recall and translation tests (“test”) were administered 3 days, 2 months, and 6 months following the end of the learning phase. High school children learned foreign language words in picture, gesture, and no enrichment conditions. **b** In each learning trial, auditorily presented Spanish words were accompanied either by a picture (picture enrichment), a video of an actress performing a gesture (gesture enrichment), or no complementary stimulus (no enrichment). Spanish words were followed by the auditorily presented German translation and a repetition of the Spanish word accompanied again by the enrichment stimulus. The children then spoke the foreign and native words following their teacher. In the gesture enrichment condition, the children performed gestures with their teacher while speaking. The children’s task was to learn the correct association between the Spanish words and their German translations



and day 8 for 90 min. The second training session occurred on either day 3 or day 4 because of differences in Spanish course scheduling between schools.

During the training sessions, L2 words and their L1 translations were presented in picture-enriched, gesture-enriched trials, and non-enriched trials (Fig. 2b). In all trial types, children first heard a Spanish word, which was followed by its auditorily-presented German translation and then by a repetition of the Spanish word. The children’s teacher then cued the children to recite the German and Spanish words aloud with the word *juntos*, which means *all together*. The teacher stood at the front of the classroom during the entire training period. In the picture enrichment condition, recorded Spanish words were accompanied by iconic line drawings, which

were presented on a screen at the front of the classroom. Pictures were presented for 3 s. In the gesture enrichment condition, recorded Spanish words were accompanied by videos of an actress performing an iconic gesture, which lasted 4 s. At the end of the trial, children performed the gesture along with the teacher. The time interval between the onset of the German word's presentation and the onset of the Spanish word's repetition was 2.5 s in all three learning conditions. In order to equate the time interval between the offset of the pictures or videos and the subsequent German word onset, and to allow for comparison with previous experiments (Andrä et al., 2020), the time interval between Spanish and German word onsets in the non-enriched learning condition was set to 2.5 s. Children's locations in the classroom were randomly assigned for each training block. Children sat at desks during the non-enriched and picture-enriched trials and stood next to desks during the gesture-enriched trials. One of the investigators monitored the testing equipment and initiated each trial as soon as the children were ready.

Learning phase trials were blocked by learning condition. Each block contained 8 trials (4 concrete word trials and 4 abstract word trials) and lasted approximately 4 min. On day 1, the children completed 2 picture-enriched blocks, 2 gesture-enriched blocks, and 2 non-enriched blocks. Each German word and its Spanish translation were therefore presented in two trials on day 1. On day 3 or day 4, the children completed 1 picture-enriched block, 1 gesture-enriched block, and 1 non-enriched block. Fewer blocks were administered on day 3 or 4 compared to other days due to the shorter Spanish course meeting time on day 3 or 4. On day 8, the children completed 2 picture-enriched blocks, 2 gesture-enriched blocks, and 2 non-enriched blocks. Children rested in a separate room between every two blocks for approximately 10 min, during which time they played simple riddle games with one of the experimenters.

Children were equally divided into groups of up to 9 students in order to counterbalance the assignment of word stimuli to the three learning conditions. This ensured that each stimulus item was learned by students in each of the three learning conditions, and that stimuli did not vary systematically between learning conditions. Additionally, word orders within each block and orders of enrichment condition blocks were counterbalanced across learning days.

Test Phase Children completed vocabulary tests at three time points: 3 days, 2 months, and 6 months following the completion of the learning phase. Free recall, German-Spanish, and Spanish-German translation tests were conducted at each time point. Tests were conducted entirely verbally, since the children did not yet possess adequate writing skills in Spanish as a foreign language. Free recall tasks tend to be more difficult for children than cued memory tasks (Karpicke et al., 2016), but have nevertheless been used for measuring children's memory capabilities (e.g., Jack et al., 2014; Lehmann & Hasselhorn, 2010; Mavilidi et al., 2015). Despite low recall rates reported in previous studies (e.g., 0.98 words on average following four 15-min training sessions spread over two weeks; Mavilidi et al., 2015), recall performance has been shown to capture L2 enrichment effects (Andrä et al., 2020; Mavilidi et al., 2015).

Native German-speaking examiners conducted the test sessions individually at the same school where the learning phase took place. The examiners were university students enrolled in teaching certification programs at the University of Leipzig, Germany. Examiners were blind with respect to which words had been learned in which enrichment condition. Further, they had no knowledge of the gestures or pictures that were paired with individual words in the experiment.

During each test session, one of the school children sat at a desk opposite one of the examiners. In the free recall test, children were asked to verbalize as many German-Spanish or Spanish-German translations, individual German words, or individual Spanish words as they could remember from the training. A time limit of 5 min was imposed; children were not instructed about this time limit, and no child's responses in any experiment exceeded 5 min. Following the free recall test, the children completed the two translation tests. The free recall test was always administered prior to the translation tests to eliminate influences of memory cues present in the translation tests.

During the German-Spanish translation test, the examiner spoke the German words one at a time, and the children were asked each time to speak the correct Spanish translation. During the Spanish-German translation test, the examiner presented audio recordings of the Spanish words one at a time, and the children were asked each time to speak the correct German translation. The German-Spanish translation test was always administered prior to the Spanish-German test, as translation from one's native to a foreign language has been shown to be a more difficult task than the translation from a foreign language into one's native language and in order to avoid cueing the Spanish words on the final test (Kroll & Stewart, 1994). Children were given 5 s to state their answers before moving to the next word. Test word orders in the two translation tests were randomized for each testing time point (3 days, 2 months, and 6 months post-learning).

Examiners recorded test sessions as an audio file for subsequent analysis using a recording device such as mobile phones or laptops. The children did not receive any feedback regarding the correctness of their answers. Children were instructed not to discuss the tests with their classmates and to not think about or rehearse the vocabulary outside of the training sessions. Children were also not rewarded for test performance at any point during the study to avoid encouraging rehearsal of the L2 words outside of the training sessions. Each test session lasted approximately 10–15 min.

No participants dropped out of the study between the day 3 testing time point and the month 2 time point. Between the month 2 time point and the month 6 time point, five 14-year-old (grade 8) participants dropped out and one 12-year-old (grade 6) dropped out. All other dropouts occurred during the learning phase. All of the children remained in Germany during the 6-month duration of the study and therefore remained immersed in a German-speaking environment throughout the study.

Data Analysis

Test Scoring Audio files from individual test sessions were independently scored for accuracy by two raters. The raters were native German speakers who were both currently enrolled in the Spanish language teaching certification program at the University of Leipzig. The two raters had not conducted any of the test sessions and were also blind with respect to which words had been learned in each enrichment condition. The two raters were in agreement for 94.2% of free recall test responses, 93.0% of L1-L2 translation test responses, and 99.1% of L2-L1 translation test responses. In cases of disagreement, a third independent rater was employed and the majority decision was adopted. The third rater was also a native German speaker currently enrolled in the Spanish language teaching certification program at the University of Leipzig.

One point was given for each correct translation provided during the free recall test. No points were given for a German word that was missing a corresponding Spanish translation or vice versa. One point was also given for each correct translation provided in the German-Spanish translation test and the Spanish-German translation test. Thus, a maximum of 24 points could be achieved on each of the three tests (4 points for each combination of the learning condition and word type factors).

Scores across the three vocabulary tests (free recall, German-Spanish translation, and Spanish-German translation) were summed for each participant, yielding combined test scores for each experimental condition. Effects of enrichment were evaluated based on performance across all vocabulary tests rather than performance on the individual tests for two main reasons. First, we did not hypothesize differential effects of learning enrichment across the three test types, as previous studies have demonstrated effects of enrichment on both recall and translation performance (Andrä et al., 2020; Macedonia & Knösche, 2011; Mathias et al., 2021a; Zimmer et al., 2000). Second, although the three tests may capture different aspects of foreign language learning, effective learning interventions should improve performance across a variety of measures. Analyses conducted at the level of the individual tests can be found in the [supplementary material](#).

Linear Mixed Effects Modeling of Composite Test Scores Linear mixed effects models were used to evaluate effects of learning condition, grade level, time point, and word type on summed test scores. A mixed effects modeling approach was used as mixed models are better able to accommodate unbalanced designs compared to traditional analyses of variance (ANOVAs). Mixed effects models are regression models which contain both random and fixed effects, whereas fixed effects are assumed to be related to independent variables, random effects are assumed to account for sources of variation due to random variables. Fixed effect coefficients in a mixed effects model are interpreted in the same way as in classical regression models. We refer the interested reader to Winter (2018) for an introduction to mixed effects modeling.

Models were generated in R version 1.2.1335 using the “lme4” package (Bates et al., 2015; Kuznetsova et al., 2017). All mixed effects models included fixed effects of learning condition (gesture, picture, none), grade level (6, 8), time point

(3 days post-learning, 2 months post-learning, 6 months post-learning), and word type (concrete, abstract). To select the random effects structure, we performed backwards model selection, beginning with a random intercept by participant and random slopes by participant for each of the four independent factors (learning condition, grade level, time point, and word type). We removed random effects terms that accounted for the least variance one by one until the fitted mixed model was no longer singular, i.e., until variances of one or more linear combinations of random effects were no longer (close to) zero. The final mixed model included two random effects terms: a random intercept by participant and a random slope by participant for the word type factor. The inclusion of the random effects term for the intercept of individual participants is equivalent to assuming that each participant may have a different baseline level of test performance. The inclusion of the random slope by participant for the word type factor is equivalent to assuming that each participant may differ in how they are influenced by the experimental manipulation of word type.

Model contrasts were coded using simple coding, i.e., ANOVA-style coding, such that the model coefficient represented the size of the contrast from a given predictor level to the (grand) mean (represented by the intercept). The complete set of mixed effects model coefficient estimates for all fixed and random effects is shown in supplementary Table S1. Following the procedure outlined by Alday et al. (2017), we summarize the model results and test for significance of fixed effects and interactions using a type II Wald χ^2 test implemented in the “car” package in R (Fox & Weisberg, 2011). A Wald χ^2 test was used instead of an F test in order to avoid issues with estimating denominator degrees of freedom in unbalanced designs (Alday et al., 2017; Bates et al., 2015; Liu, 2016). Post hoc Tukey tests were conducted using the “emmeans” package (Lenth et al., 2019). Cramer’s v and Cohen’s d were used as measures of effect size. The significance threshold was set to $\alpha = 0.05$ (Greenland et al., 2016).

Control Analysis on Influence of School Attended and Stimulus List Since the 12- and 14-year-olds included in the study attended three different high schools and learned partially overlapping lists of L2 words, we addressed in a control analysis whether enrichment effects were influenced by the school attended and corresponding stimulus list. We compared performance of the 14-year-olds enrolled at school 2 with 14-year-olds enrolled at school 3. If the school attended and/or stimulus list had no influence on enrichment effects, then we would expect to observe no interactions between the school factor and any other experimental factors. The control analysis yielded no effect of school and no interactions of experimental factors with the school factor (see [supplementary material](#), “Control analysis on influences of school and stimulus list” and Table S2 for a summary of the results). Children in the same age group at two different schools showed the same enrichment effects despite differences in stimulus lists. We therefore in the following pool together the 14-year-old participants who attended schools 2 and 3.

We would expect the lack of school-driven and stimulus-driven differences on enrichment effects between 14-year-old students who attended schools 2 and 3 to extend to 12-year-old students who attended school 1. Schools 1, 2, and 3 did not differ in terms of demographics, and beneficial effects of enrichment on the learning of L2 vocabulary

have been previously been found using a variety of stimulus items including also several word classes beyond those tested here such as verbs, adverbs, adjectives, and prepositions (Andrä et al., 2020; Macedonia & Klimesch, 2014; Macedonia & Knösche, 2011; Mayer et al., 2015, 2017; Repetto et al., 2017; Saltz & Donnenwerth-Nolan, 1981).

Linear Mixed Effects Modeling of Individual Test Scores In addition to analyzing children’s scores summed across the three test types, we performed follow-up analyses to evaluate children’s performance at the level of the individual tests (free recall, L1-L2 translation, and L2-L1 translation). Mixed effects models were fitted to data from each test separately. Mean test scores on the three individual tests are shown in Table S3, and type II Wald χ^2 tests summarizing the model results for each test type are shown in Tables S4, S5, and S6. Please see “[Linear mixed effects modeling of individual test scores](#)” in the [supplementary material](#) for further details.

Results

Children’s mean composite vocabulary test scores at 3 days, 2 months, and 6 months post-learning by condition are shown in Table 4.

Do Younger and Older Children Benefit from Gesture and Picture Enrichment Similarly or Differently?

We first addressed our main aim of the paper by testing whether 12-year-olds (sixth graders) and 14-year-olds (eighth graders) benefitted similarly or differently from gesture and picture enrichment. Mixed effects modeling of children’s vocabulary test scores revealed an interaction between the children’s grade level and the learning condition ($\chi^2(2, N=75)=5.70, p=0.04, v=0.19$; the complete set of model coefficients for all fixed and random effects is shown in supplementary Table S1, and significance testing of model effects is shown in Table 5). Tukey’s HSD post hoc tests showed that children in both grade levels benefitted from gesture enrichment relative to non-enriched learning (grade 6, $\beta=1.56, t=6.05, p<0.001, d=1.39$; grade 8, $\beta=1.87, t=6.99, p<0.001, d=1.60$), shown in Fig. 3. This was also the case for the picture enrichment condition (grade 6, $\beta=1.47, t=5.70, p<0.001, d=1.36$; grade 8, $\beta=0.92, t=3.42, p=0.008, d=0.82$). However, gesture enrichment enhanced learning outcomes even more than picture enrichment for the eighth graders ($\beta=0.95, t=3.56, p=0.005, d=0.85$), which was not the case for the sixth graders ($\beta=0.09, t=0.35, p=0.99, d=0.08$). In sum, both groups of children benefitted from both types of enrichment, and gesture enrichment was even more beneficial than picture enrichment for the older children than the younger children. This result is likely triggered primarily by performance on the L2-L1 translation test (see Table S6).

Table 4 Children's composite scores on free recall and translation tests. A maximum of 12 points could be achieved for each combination of the learning condition and word type factors at each time point. *M* = mean composite test score, *SE* = standard error

		3 days post-learning		2 months post-learning		6 months post-learning	
Learning condition		Grade 6 <i>M (SE)</i>	Grade 8 <i>M (SE)</i>	Grade 6 <i>M (SE)</i>	Grade 8 <i>M (SE)</i>	Grade 6 <i>M (SE)</i>	Grade 8 <i>M (SE)</i>
Abstract words	Picture	9.18 (.59)	8.00 (.52)	8.95 (.73)	7.22 (.58)	8.38 (.61)	6.14 (.46)
	Gesture	9.28 (.52)	8.58 (.63)	8.72 (.50)	7.78 (.53)	8.85 (.56)	7.31 (.45)
	None	7.69 (.56)	6.92 (.69)	7.64 (.55)	5.42 (.56)	6.59 (.57)	5.28 (.62)
Concrete words	Picture	8.18 (.60)	9.86 (.54)	9.00 (.70)	9.08 (.52)	8.13 (.83)	8.92 (.49)
	Gesture	9.38 (.61)	10.97 (.57)	7.90 (.60)	10.39 (.63)	8.23 (.61)	9.89 (.60)
	None	7.77 (.64)	9.53 (.60)	7.00 (.70)	8.50 (.61)	6.33 (.69)	8.06 (.55)

Table 5 Type II Wald χ^2 test of mixed effects model effects of learning condition, grade level, word type, and time point. *df*=degrees of freedom. * $p < .05$, ** $p < .01$, *** $p < .001$

	χ^2	<i>df</i>	<i>p</i>	Cramer's <i>v</i>
Learning	91.82	2	< .001***	.78
Grade	2.24	1	.13	.17
Word	44.45	1	< .001***	.77
Time	25.57	2	< .001***	.41
Learning \times grade	5.70	2	.04*	.19
Learning \times word	1.08	2	.58	.08
Grade \times word	90.30	1	< .001***	1.10
Learning \times time	4.29	4	.37	.12
Grade \times time	13.15	2	.001**	.30
Word \times time	.53	2	.77	.06
Learning \times grade \times word	.55	2	.76	.06
Learning \times grade \times time	3.41	4	.49	.11
Learning \times word \times time	3.02	4	.55	.10
Grade \times word \times time	.24	2	.89	.04
Learning \times grade \times word \times time	1.25	4	.87	.06

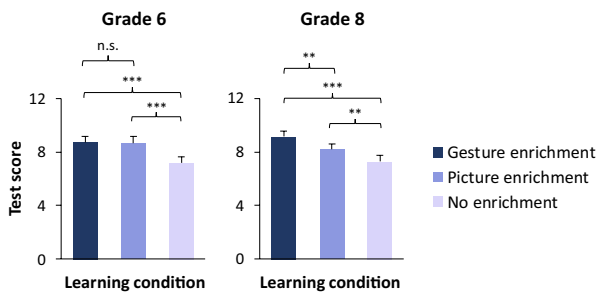


Fig. 3 Test scores by learning condition and children's grade level. Children in grades 6 (12-year-olds; left) and grade 8 (14-year-olds; right) demonstrated higher overall test scores following gesture-enriched learning and picture-enriched learning compared to non-enriched learning. Eighth graders benefitted significantly more from gesture enrichment than picture enrichment, while sixth graders demonstrated equivalent learning outcomes for both gesture- and picture-enriched words. This difference was significant; i.e., there was an interaction between the learning condition and grade level factors. A maximum of 12 points per learning condition could be achieved, as scores were averaged across word types. ** $p < 0.01$, *** $p < .001$. n.s. = not significant

Do Younger and Older Children Benefit from Gesture and Picture Enrichment Relative to Non-enriched Learning?

We next tested whether gesture enrichment and picture enrichment would benefit children's test scores compared to non-enriched learning, irrespective of grade level, as expected from previous studies in elementary school children and adults (Andrä et al., 2020; Mayer et al., 2015). The mixed effects model indicated significantly higher scores for words learned with gesture and picture enrichment compared to words learned with no enrichment (gesture condition, $\beta = 1.71$, $t = -9.23$, $p < 0.001$,

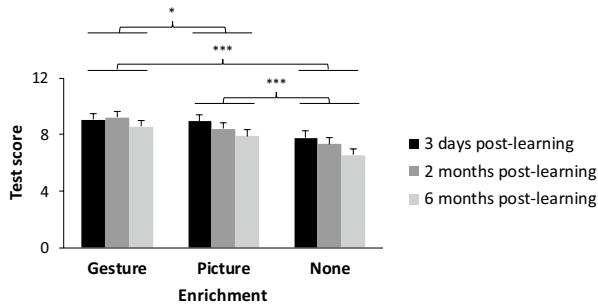


Fig. 4 Test scores by learning condition and time point. Children demonstrated higher overall test scores following gesture-enriched learning and picture-enriched learning compared to non-enriched learning, and enrichment benefits did not significantly differ across time points. A maximum of 12 points per learning condition could be achieved at each time point, as scores were averaged across word types. * $p < 0.05$, *** $p < .001$

$d = 1.52$; picture condition, $\beta = 1.19$, $t = 6.42$, $p < 0.001$, $d = 1.06$). The model also revealed that, overall, scores for gesture-enriched words were significantly higher than scores for picture-enriched words ($\beta = 0.52$, $t = 2.81$, $p = 0.014$, $d = 0.46$).

Do Enrichment Benefits Persist Over Long Time Scales?

We also expected the beneficial effects of picture and gesture enrichment on children's learning to persist over long time scales (up to 6 months following learning; Andrä et al., 2020; Mayer et al., 2015). This was found to be the case; there was no significant interaction between learning condition and time point (χ^2 (4, $N = 75$) = 4.29, $p = 0.37$, $v = 0.12$). Both gesture- and picture-enriched learning benefitted children's L2 vocabulary learning outcomes compared with non-enriched learning, irrespective of testing time point and children's grade level, shown in Fig. 4.

Does Enrichment Benefit the Learning of Both Concrete and Abstract Words?

In agreement with previous studies in elementary school children and adults (Andrä et al., 2020; Macedonia & Knösche, 2011; Mayer et al., 2017), picture and gesture enrichment benefitted high school children's learning of both concrete and abstract word types compared to non-enriched learning: The mixed effects model indicated no significant interaction between learning condition and word type variables (χ^2 (2, $N = 75$) = 1.08, $p = 0.58$, $v = 0.08$). Picture-enriched learning yielded significantly higher test scores than non-enriched learning for both concrete words ($\beta = 1.00$, $t = 3.80$, $p = 0.002$, $d = 0.62$) and abstract words ($\beta = 1.39$, $t = 5.28$, $p < 0.001$, $d = 0.87$), shown in Fig. 5. Gesture-enriched learning also yielded significantly higher test scores than non-enriched learning for both concrete words ($\beta = 1.60$, $t = 6.08$, $p < 0.001$, $d = 1.00$) and abstract words ($\beta = 1.83$, $t = 6.97$, $p < 0.001$, $d = 1.15$).

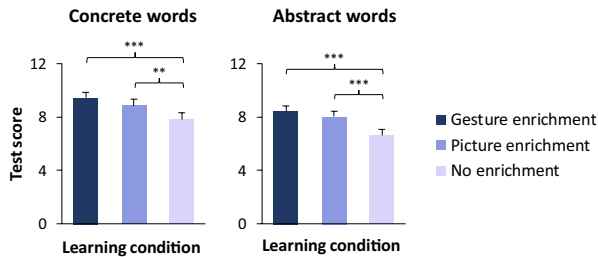


Fig. 5 Test scores by learning condition and word type. Children demonstrated higher overall test scores following gesture-enriched learning and picture-enriched learning compared to non-enriched learning for both concrete words (left) and abstract words (right). A maximum of 12 points per combination of the learning condition and word type factors could be achieved. ** $p < 0.01$, *** $p < .001$

The mixed modeling of children's test scores revealed several additional significant effects, which we report here for completeness. Test scores for concrete words were, overall, significantly higher than scores for abstract words, a main effect of word type ($\chi^2 (1, N=75)=44.45, p<0.001, \nu=0.77$). There was also a significant main effect of time point ($\chi^2 (2, N=75)=25.57, p<0.001, \nu=0.41$). These main effects were expected based on previous reports of children's greater performance for concrete than abstract nouns (Schwanenflugel, 1991), and reports of memory decay over time (Caramelli et al., 2004; Howe & Brainerd, 1989). The model also revealed significant grade \times word type and grade \times time point interactions, shown in Table 4. The grade \times word type interaction was driven by significantly greater performance for concrete words compared to abstract words for the eighth graders ($\beta=2.51, t=11.32, p<0.001, d=2.71$), which was not the case for the sixth graders ($\beta=0.37, t=1.75, p=0.30, d=0.40$). The grade level \times time point interaction was driven by a significant reduction in performance at 2 months post-learning compared to 3 days post-learning for the eighth graders ($\beta=1.38, t=5.14, p<0.001, d=1.23$), which did not occur for the sixth graders ($\beta=0.45, t=1.74, p=0.50, d=0.40$).

Discussion

The present study was motivated by previous findings that adults'—but not elementary school children's—L2 vocabulary learning benefits to a greater extent from sensorimotor (gesture) than from multisensory (picture) enrichment (Andrä et al., 2020; Mathias et al., 2021a; Mayer et al., 2015). We addressed the question of whether intermediate age groups would display enrichment benefits that are more comparable to those displayed by adults (i.e., gesture enrichment facilitating learning more than picture enrichment) or to those displayed by elementary school children (i.e., similar learning outcomes for gesture and picture enrichment). We found that both picture and gesture enrichment interventions were beneficial relative to non-enriched (auditory-only) learning for 12-year-olds (sixth graders) and 14-year-olds

(eighth graders). Interestingly, however, gesture-enriched learning was even more beneficial than picture-enriched learning for the eighth graders, while the sixth graders benefitted equivalently from learning enriched with pictures and gestures. This finding suggests that the effectiveness of gesture and picture enrichment techniques differs between younger and older L2 learners. While the pattern of enrichment effects for eighth graders qualitatively resembles that observed previously for young adults (Mathias et al., 2021a; Mayer et al., 2015), the pattern of effects observed for sixth graders resembles that observed previously for elementary school children (Andrä et al., 2020). As was the case in previous studies on L2 enrichment, picture and gesture enrichment benefitted the learning of both concrete nouns (e.g., *tent*) and abstract nouns (e.g., *patience*), and effects of enrichment persisted over a long time scale (up to 6 months post-learning). Taken together, the findings suggest that congruent information presented in visual and motor modalities during auditory word learning may be differentially weighted by learners of different ages.

Gesture Enrichment Benefitted Learning More than Picture Enrichment in Fourteen-Year-Old Children but not Twelve-Year-Old Children

Children of both age groups were able to make use of enrichment information in a way that supported vocabulary knowledge. Across all time points and word types, performing gestures during L2 learning enhanced subsequent learning outcomes relative to auditory-only learning by about 22% in sixth graders and 25% in eighth graders. Viewing pictures enhanced learning outcomes by about 20% in sixth graders and 12% in eighth graders. These benefits are substantial when considering that children received minimal L2 exposure: Each L2 word was presented a total of only five times across three learning days. The children also never viewed the written words and thus relied only on spoken words to form representations of the L2 tokens. Effects of enrichment were robust enough to support L2 translation for up to 6 months following learning, despite the omission of L2 stimuli from Spanish lessons by the children's teachers for the 6 months following the learning phase.

Beneficial effects of gesture enrichment on L2 vocabulary learning are consistent with a variety of psychological accounts. From an embodied perspective (reviewed in Atkinson, 2010; Barsalou, 2008; Meteyard et al., 2012; Wellsby & Pexman, 2014), gesture enrichment could have improved children's L2 memory by grounding the meanings of novel L2 words in sensorimotor experiences. This occurred for both concrete and abstract nouns, in support of the notion that abstract concepts, like concrete concepts, may be grounded in perception and action (Harpaintner et al., 2018, 2020). Dual coding accounts (Engelkamp & Zimmer, 1985; Hommel et al., 2001; Paivio, 1991; Paivio & Csapo, 1969) would suggest that both the L1 and L2 words were likely encoded verbally, while the viewing of pictures generated non-verbal visual encoding, and the performance of gestures generated both visual and haptic encodings. Enhanced retention of gesture- and picture-enriched words may be attributable to their more complex codes. Hearing an L1 or L2 word at test may also have triggered reconstructions of picture or gesture enrichment material, which

could have offered a greater number of routes to retrieving the correct translation, consistent with imagery accounts (Jeannerod, 1995; Kosslyn et al., 2006; Saltz & Dixon, 1982) and predictive coding accounts (Mathias et al., 2021a; Mayer et al., 2017; von Kriegstein, 2012).

The roughly equivalent benefits of gesture and picture enrichment for the sixth-grade children is consistent with the pattern of gesture and picture benefits recently shown in 8-year-old school children (Andrä et al., 2020). The superior effects of gesture enrichment relative to picture enrichment for the eighth-grade children is consistent with the pattern of gesture and picture benefits recently shown in adults (Mathias et al., 2021a; Mayer et al., 2015; Repetto et al., 2017). Differences in enrichment benefits between age groups cannot be attributed to differences in gesture or picture stimuli, L2 perceptual characteristics, or training procedures, as these did not differ across age groups. Differences can also not be attributed to testing environments or to the translation from lab-based experiments in adults to a school setting, as both age groups in the current study were tested in similar school environments. The use of the same design and number of stimuli for two different age groups also overcomes the difficulty in comparing across studies in children and adults that vary in terms of the number of stimuli tested (Andrä et al., 2020; Macedonia & Knösche, 2011; Macedonia et al., 2011; Mavilidi et al., 2015; Mayer et al., 2015; Repetto et al., 2017).

We offer two speculative explanations for the differences in effects of gesture and picture enrichment between sixth- and eighth-graders. The first explanation relates to potential advances in literacy in eighth graders compared to sixth graders. Children in the initial stages of reading skill acquisition may rely to a greater extent on visual context for L1 word learning relative to older children and adults (Nicholas & Lightbown, 2008). During the emergence of literacy, pictures and picture books serve as critical tools for language comprehension and vocabulary acquisition as they illustrate the meaning of spoken text (Ann Evans & Saint-Aubin, 2005; Feathers & Arya, 2012). Children are generally able to understand the referential nature of pictures—the idea that pictured contents can represent objects and concepts in the real world—by the age of two (Allen Preissler & Carey, 2004; Ganea et al., 2009). While chapter books tend to include illustrations for children up to about 12 years, books intended for older children and adolescents rarely do so, and picture books tend not to be used as learning materials in older children's classrooms (Beckett, 2013). Instead, the majority of L1 vocabulary learning in adolescents and adults is thought to occur incidentally during the reading of written text (Webb, 2008); this is potentially also the case for L2 (Brown et al., 2008; Grabe, 2009; Huckin & Coady, 1999). Thus, pictures are likely to play a greater role in aiding the learning of L2 vocabulary in younger children who may still be in the process of acquiring L1 competencies compared to older children (Spichtig et al., 2017).

The second explanation relates to differences in the degree to which children of different ages may rely on procedural and declarative memory systems for remembering L2 words. Theories of memory distinguish between procedural (implicit) and declarative (explicit) memory systems (Cohen et al., 1997; Squire & Zola-Morgan, 1991; Tulving & Madigan, 1970). Vocabulary learning is typically situated theoretically in the domain of declarative memory (Cabeza & Moscovitch, 2013), whereas other

types of language learning such as grammar learning have become associated with the procedural memory system (Hamrick, 2015; Ullman, 2004). It has been suggested that gesture enrichment may engage the procedural memory system to a greater extent than audiovisual learning in adults (Macedonia & Mueller, 2016; Mathias et al., 2021a), consistent with proposals that declarative and procedural memory systems in adults are interactive rather than distinct (Davis & Gaskell, 2009). Though declarative memory functions are not yet fully developed in younger children (Schneider, 2008), several studies have observed no differences between young children and adults in terms of procedural memory abilities (Finn et al., 2016; Karatekin et al., 2007; Meulemans et al., 1998). It could be the case that, while children of both age groups made use of procedural memory systems for the learning of gesture-enriched vocabulary, picture enrichment recruited procedural memory systems only in the younger children. This would result in equivalent gesture- and picture-enriched learning outcomes in the younger children, and a reduction in picture enrichment benefits in the older children.

These potential explanations are currently speculative. Future studies may investigate how benefits of picture and gesture enrichment in children of different ages relate to the concurrent acquisition of reading and other academic skills, as well as the maturation of procedural and declarative memory. In terms of enrichment strategies that would be recommended for evidence-based L2 teaching, two open questions remain. First, does the current set of results extend to more commonly used forms of L2 instruction in which vocabulary acquisition is integrated into other L2 learning activities that are not focused explicitly on acquiring new vocabulary? Second, would the combination of gestures and pictures provide even larger enrichment benefits or would it create a dual attentional load resulting in inferior memory outcomes?

Neuroscience Evidence for Contributions of Sensory and Motor Representations to Enrichment Benefits

At present, the majority of neuroscience studies investigating learning enrichment have been conducted in adults. These studies suggest that beneficial effects of sensorimotor and multisensory enrichment derive, at least in part, from L2 representations stored in sensory and motor areas of the cortex. For example, listening to gesture-enriched L2 vocabulary elicits responses within regions associated with viewing and performing movements (Macedonia et al., 2011; Mayer et al., 2015), and these areas were found using a non-invasive neurostimulation method to causally facilitate the translation of L2 vocabulary (Mathias et al., 2021a, b). These findings are comparable to neuroimaging studies in children, which have demonstrated preschoolers' greater motor (Kersey & James, 2013) and visual (James, 2010) cortical responses while viewing letters that they have previously been taught to write, compared to letters that they have been taught to recognize visually. Thus, the reactivation of neural sensory and motor structures at test that are involved in processing enrichment material during learning may drive enrichment benefits (multisensory learning theory, Shams & Seitz, 2008; von Kriegstein, 2012; von Kriegstein & Giraud, 2006).

Findings that sensory and motor brain areas directly contribute to the translation of sensorimotor-enriched L2 vocabulary undermine some alternative explanations for the effectiveness of enrichment such as increased arousal or attention relative to unisensory learning (Kelly et al., 2009; Krönke et al., 2013). In line with this evidence, 12-year-olds in the current study showed equivalent picture and gesture benefits, which would not be expected based on relative levels of sensorimotor arousal during picture- and gesture-enriched learning or based on the relative novelty of picture and gesture enrichment as classroom learning strategies. Unlike arousal-based learning interventions such as exercise (Hötting et al., 2016) or the manipulation of emotion (Storbeck & Maswood, 2016), enrichment learning binds together congruent stimuli presented in two or more modalities (reviewed in Markant et al., 2016).

Role of L2 Test Type

Both the 12-year-olds and 14-year-olds showed benefits of gesture and picture enrichment also at the level of each of the individual vocabulary tests (free recall, L1-L2 translation, L2-L1 translation), with the exception of the 12-year-olds' performance on the L1-L2 translation test, for which the gesture benefit was not significant. Results conducted at the level of the individual tests suggest that the interaction between grade level and learning condition factors for composite test scores was triggered by primarily performance on the L2-L1 translation test. The lower free recall test scores relative to translation test scores are in line with previous findings showing that free recall tasks tend to be more difficult than cued memory tasks for both primary school children (Karpicke et al., 2016) and adults (for review see Cleary, 2018). The overall magnitude of free recall scores is consistent with previously reported scores in L2 free recall tasks (Andrä et al., 2020; Mavilidi et al., 2015). Low test scores in the current and previous studies are likely attributable to the short timeframe of L2 training. We would expect beneficial effects of enrichment to scale up as the timeframe of training increases. However, it remains unknown whether enrichment would generate even stronger effects if integrated into coursework over a longer period.

Potential Effect of Stimulus Complexity and Timing

The learning conditions in the current study differed not only in terms of the sensory modalities in which the enrichment was presented (i.e., sensorimotor versus sensory). Videos of gestures are inherently dynamic while pictures are static, which may make gestural stimuli more visually complex than picture stimuli. Although the chunking of linguistic units may aid language learning (McCauley & Christiansen, 2017), no previous studies have compared chunking processes across gesture, picture, and auditory stimuli. One could speculate that auditory-only learning involves fewer chunks or discontinuities than gesture-enriched or picture-enriched learning. However, speakers are known to produce gestures more often when cognitive load is high, suggesting that gestures function to reduce speech memory demands

(Melinger & Kita, 2007; reviewed in Risko & Gilbert, 2016). Gesture-, picture-, and non-enriched trials also varied in terms of stimulus presentation duration to allow for qualitative comparison of the current results with those of Andrä et al. (2020), who used the same stimulus timings. Gestures were presented for 4.0 s, pictures for 3.5 s, and spoken words in the auditory-only condition for 2.5 s. A shorter time interval was used for the presentation of the spoken Spanish words in the auditory-only condition, compared to videos in the gesture condition and pictures in the picture condition, in order to avoid introducing long time intervals during which participants would have waited between consecutively-presented stimuli. Long time periods during which no sensory information is presented could have the effect of decreasing attention, motivation, or stimulus-driven arousal in the non-enriched task compared to the other tasks. Mayer and colleagues (2015) presented young adults with the same learning conditions (gesture-, picture-, and non-enriched learning) and maintained a constant trial length across conditions. Even when trial lengths were identical across conditions, participants showed enhanced L2 vocabulary learning for the enriched conditions compared to the non-enriched condition.

Study Limitations

We have focused here on verbal L2 learning and recall, as have most studies examining effects of enrichment on L2 learning (e.g., Krönke et al., 2013; Macedonia & Klimesch, 2014; Macedonia & Knösche, 2011; Mathias et al., 2021a, b; Mayer et al., 2015). An open question is whether the learning of written L2 words can also be enhanced by multisensory and sensorimotor enrichment. Similarly, whether complementary information presented in other sensory modalities such as haptic input can benefit L2 learning remains untested. It is likely that memories for gestures in the current study involved both sensory and motor components, as the children viewed the gestures while performing them. Since motoric enrichment techniques are consistently accompanied by sensory feedback, and perceptual and motor learning generally occur together (reviewed in Ostry & Gribble, 2016), we characterize gesture enrichment as “sensorimotor”-enriched learning rather than “motor”-enriched learning.

We assume that gestures provide a more unusual tool for learning than pictures in classroom contexts. It would be interesting to additionally compare gesture enrichment with other forms of enrichment that are similarly unusual for students. If the unusualness of gestures contributes to beneficial learning effects, then we would expect it to similarly modulate effects of gesture enrichment in both age groups investigated here.

The current study focused on effects of enrichment on children’s learning of concrete and abstract nouns. Although studies in adults have tested enrichment effects on several other word classes including verbs, adverbs, adjectives, and prepositions (Andrä et al., 2020; Macedonia & Klimesch, 2014; Macedonia & Knösche, 2011; Mayer et al., 2015, 2017; Repetto et al., 2017; Saltz & Donnenwerth-Nolan, 1981), effects of enrichment on these other word classes in children remain unknown. Concrete and abstract nouns provide useful test cases as they differ in terms of how easily

they can be represented in pictures and gestures (Borghini et al., 2017). The relationship between the overall number of words that children learn and the strength of enrichment benefits also remains unclear. Andr  et al. (2020) found that reducing the number of L2 words learned by eight-year-old school children from 40 to 24 words over a week-long training period did not increase the benefits of enrichment on recall or translation accuracy. Future work may examine the influence of these variables on enrichment effects by testing children’s learning of other word types and by manipulating the number of words that are learned over a fixed training period.

Conclusion and Practical Implications

We identified a dissociation in the effects of multisensory (picture) and sensorimotor (gesture) enrichment on L2 learning across 12- and 14-year-old school children. Whereas 14-year-old children benefitted more from learning with gestures than with pictures, 12-year-old children showed equivalent learning benefits following gesture- and picture-enriched learning. Gesture and picture enrichment strategies were tested systematically using large sample sizes of children in naturalistic school environments. We conclude that visual and motor enrichment information may be weighted differently by children of different ages and that sensorimotor forms of enrichment may be more beneficial to older children for L2 vocabulary learning than audiovisual enrichment.

The differences in effects of enrichment strategies between age groups observed here suggest that strategies derived from studies on one age group may not directly translate into teaching strategies to be used in another age group. Our findings provide evidence-based grounds for opting to include gestures rather than pictures in L2 vocabulary teaching for school children starting at fourteen years of age. Gestures and other sensorimotor-based interventions may be more challenging for educators to integrate into pedagogy than picture-based interventions. The finding that picture-based interventions are just as helpful as gesture-based interventions in the context of L2 learning for younger children therefore has immediate implications for evidence-based teaching techniques in younger age groups. For older students, the use of gestures as a method for enhancing L2 learning provides another tool in students’ and educators’ active learning toolboxes.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10648-021-09648-z>.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was funded by the German Research Foundation (grant KR 3735/3-1) and a Schulbezogene Forschung grant from the Saxony Zentrum f r Lehrerbildung und Schulforschung (ZLS).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alday, P. M., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2017). Electrophysiology reveals the neural dynamics of naturalistic auditory language processing: Event-related potentials reflect continuous model updates. *Eneuro*, 4(6).
- Allen Preissler, M., & Carey, S. (2004). Do both pictures and words function as symbols for 18-and 24-month-old children? *Journal of Cognition and Development*, 5(2), 185–212.
- Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein, K. (2020). Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children. *Educational Psychology Review*, 32, 815–850.
- Ann Evans, M., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological Science*, 16(11), 913–920.
- Atkinson, D. (2010). Extended, embodied cognition and second language acquisition. *Applied Linguistics*, 31(5), 599–622.
- Bara, F., & Kaminski, G. (2019). Holding a real object during encoding helps the learning of foreign vocabulary. *Acta Psychologica*, 196, 26–32.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Bates, D., Maechler, M., Bolker, B. S., & Walker, D. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Beckett, S. L. (2013). *Crossover picturebooks: A genre for all ages*. Routledge.
- Bergen, B. (2015). Embodiment, simulation and meaning. In N. Riemer (Ed.), *The Routledge handbook of semantics* (pp. 158–173). New York: Routledge.
- Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, 143(3), 263.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20, 136–163.
- Cabeza, R., & Moscovitch, M. (2013). Memory systems, processing modes, and components: Functional neuroimaging evidence. *Perspectives on Psychological Science*, 8(1), 49–55.
- Cabrera, M. P., & Martínez, P. B. (2001). The effects of repetition, comprehension checks, and gestures, on primary school children in an EFL situation. *ELT Journal*, 55(3), 281–288.
- Caramelli, N., Setti, A., & Maurizzi, D. D. (2004). Concrete and abstract concepts in school age children. *Psychology of Language and Communication*, 8(2), 19–34.
- Choo, L. B., Lin, D. T. A., & Pandian, A. (2012). Language learning approaches: A review of research on explicit and implicit learning in vocabulary acquisition. *Procedia-Social and Behavioral Sciences*, 55, 852–860.
- Cleary, A.M. (2018). From free recall to recognition. In H. Otani & B. L. Schwartz (Eds.), *Handbook of research methods in human memory*. New York: Routledge.
- Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory*, 5(1–2), 131–178.
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London b: Biological Sciences*, 364(1536), 3773–3800.
- de Wit, J., Schodde, T., Willemsen, B., Bergmann, K., de Haas, M., Kopp, S., Krahmer, E., & Vogt, P. (2018, February). The effect of a robot's gestures and adaptive tutoring on children's acquisition of second language vocabularies. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction* (pp. 50–58). ACM.
- Drew, V., & Mackie, L. (2011). Extending the constructs of active learning: Implications for teachers' pedagogy and practice. *Curriculum Journal*, 22(4), 451–467.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, D. M., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis. *Reading Research Quarterly*, 36, 250–287.
- Engelkamp, J., & Zimmer, H. D. (1985). Motor programs and their relation to semantic memory. *German Journal of Psychology*, 28, 239–254.
- Feathers, K. M., & Arya, P. (2012). The role of illustrations during children's reading. *Journal of Children's Literature*, 38(1), 36.
- Finn, A. S., Kalra, P. B., Goetz, C., Leonard, J. A., Sheridan, M. A., & Gabrieli, J. D. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology*, 142, 212–220.

- Frick, A., Daum, M. M., Wilson, M., & Wilkening, F. (2009). Effects of action on children's and adults' mental imagery. *Journal of Experimental Child Psychology*, 104(1), 34–51.
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society B: Biological Sciences*, 364(1521), 1211–1221.
- Funk, M., Brugger, P., & Wilkening, F. (2005). Motor processes in children's imagery: The case of mental rotation of hands. *Developmental Science*, 8(5), 402–408.
- Gabbard, C. (2009). Studying action representation in children via motor imagery. *Brain and Cognition*, 71(3), 234–239.
- Ganea, P. A., Allen, M. L., Butler, L., Carey, S., & DeLoache, J. S. (2009). Toddlers' referential understanding of pictures. *Journal of Experimental Child Psychology*, 104(3), 283–295.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, 31(4), 337–350.
- Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, 44, 9–15.
- Harpaintner, M., Sim, E. J., Trumpp, N. M., Ulrich, M., & Kiefer, M. (2020). The grounding of abstract concepts in the motor and visual system: An fMRI study. *Cortex*, 124, 1–22.
- Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2018). The semantic content of abstract concepts: A property listing study of 296 abstract words. *Frontiers in Psychology*, 9, 1748.
- Hille, K., Vogt, K., Fritz, M., & Sambanis, M. (2010). Szenisches Lernen im Fremdsprachenunterricht: die Evaluation eines Schulversuchs. Diskurs Kindheits- und Jugendforschung/Discourse. *Journal of Childhood and Adolescence Research*, 5(3), 337–350.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). Codes and their vicissitudes. *Behavioral and Brain Sciences*, 24(5), 910–926.
- Hötting, K., Schickert, N., Kaiser, J., Röder, B., & Schmidt-Kassow, M. (2016). The effects of acute physical exercise on memory, peripheral BDNF, and cortisol in young adults. *Neural Plasticity*, 1–12.
- Howe, M. L., & Brainerd, C. J. (1989). Development of children's long-term retention. *Developmental Review*, 9(4), 301–340.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition*, 21(2), 181–193.
- Jack, F., Leov, J., & Zajac, R. (2014). Age-related differences in the free-recall accounts of child, adolescent, and adult witnesses. *Applied Cognitive Psychology*, 28(1), 30–38.
- James, K. H. (2010). Sensori-motor experience leads to changes in visual processing in the developing brain. *Developmental Science*, 13(2), 279–288.
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, 33(11), 1419–1432.
- Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE—Life Sciences Education*, 14(1), ar5.
- Karatekin, C., Marcus, D. J., & White, T. (2007). Oculomotor and manual indexes of incidental and intentional spatial sequence learning during middle childhood and adolescence. *Journal of Experimental Child Psychology*, 96, 107–130.
- Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). Retrieval-based learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology*, 7, 350.
- Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes*, 24(2), 313–334.
- Kersey, A. J., & James, K. H. (2013). Brain activation patterns resulting from learning letter forms through active self-production and passive observation in young children. *Frontiers in Psychology*, 4, 567.
- Kiefer, M., & Trumpp, N. M. (2012). Embodiment theory and education: The foundations of cognition in perception and action. *Trends in Neuroscience and Education*, 1(1), 15–20.
- Köper, M., & Im Walde, S. S. (2016). Automatically generated affective norms of abstractness, arousal, imageability and valence for 350 000 german lemmas. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 2595–2598). Portorož, Slovenia: European Language Resources Association (ELRA).
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. New York: Oxford University Press.

- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174.
- Krönke, K. M., Mueller, K., Friederici, A. D., & Obrig, H. (2013). Learning by doing? The effect of gestures on implicit retrieval of newly acquired words. *Cortex*, 49(9), 2553–2568.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13).
- Lehmann, M., & Hasselhorn, M. (2010). The dynamics of free recall and their relation to rehearsal between 8 and 10 years of age. *Child Development*, 81(3), 1006–1020.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). Package “emmeans”: Estimated Marginal Means, aka Least-Squares Means. *The Comprehensive R Archive Network*, 1–67.
- Li, J. T., & Tong, F. (2019). Multimedia-assisted self-learning materials: The benefits of E-flashcards for vocabulary learning in Chinese as a foreign language. *Reading and Writing*, 32(5), 1175–1195.
- Liu, X. (2016). Linear mixed-effects models. In X. Liu (Ed.), *Methods and Applications of Longitudinal Data Analysis* (pp. 61–94). Heidelberg: Academic Press.
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, 75(5), 1357–1372.
- Macedonia, M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in Psychology*, 5.
- Macedonia, M., Bergmann, K., & Roithmayr, F. (2014). Imitation of a pedagogical agent’s gestures enhances memory for words in second language. *Science Journal of Education*, 2(5), 162–169.
- Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education*, 8(2), 74–88.
- Macedonia, M., & Knösche, T. R. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education*, 5(4), 196–211.
- Macedonia, M., & Mueller, K. (2016). Exploring the neural representation of novel words learned through enactment in a word recognition task. *Frontiers in Psychology*, 7, 953.
- Macedonia, M., Müller, K., & Friederici, A. D. (2010). Neural correlates of high performance in foreign language vocabulary learning. *Mind, Brain, and Education*, 4(3), 125–134.
- Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, 32(6), 982–998.
- Mahmoudi, S., Jafari, E., Nasrabadi, H. A., & Liaghatdar, M. J. (2012). Holistic Education: An Approach for 21 Century. *International Education Studies*, 5(2), 178–186.
- Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced memory as a common effect of active learning. *Mind, Brain, and Education*, 10(3), 142–152.
- Mathias, B., Sureth, L., Hartwigsen, G., Macedonia, M., Mayer, K. M., & von Kriegstein, K. (2021a). Visual sensory cortices causally contribute to auditory word recognition following sensorimotor-enriched vocabulary training. *Cerebral Cortex*, 31(1), 513–528.
- Mathias, B., Waibel, A., Hartwigsen, G., Sureth, L., Macedonia, M., Mayer, K. M., & von Kriegstein, K. (2021b). Motor cortex causally contributes to vocabulary translation following sensorimotor-enriched training. *Journal of Neuroscience*, 41(41), 8618–8631.
- Mavilidi, M. F., Okely, A. D., Chandler, P., Cliff, D. P., & Paas, F. (2015). Effects of integrated physical exercises and gestures on preschool children’s foreign language vocabulary learning. *Educational Psychology Review*, 27(3), 413–426.
- Mayer, K. M., Macedonia, M., & von Kriegstein, K. (2017). Recently learned foreign abstract and concrete nouns are represented in distinct cortical networks similar to the native language. *Human Brain Mapping*, 38(9), 4398–4412.
- Mayer, K. M., Yildiz, I. B., Macedonia, M., & von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology*, 25(4), 530–535.
- McCauley, S. M., & Christiansen, M. H. (2017). Computational investigations of multiword chunks in language learning. *Topics in Cognitive Science*, 9(3), 637–652.
- McFalls, E. L., Schwanenflugel, P. J., & Stahl, S. A. (1996). Influence of word meaning on the acquisition of a reading vocabulary in second-grade children. *Reading and Writing*, 8(3), 235–250.
- Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, 22(4), 473–500.
- Meteyard, L., Cuadrado, S. R., Bahrami, B., & Vigliocco, G. (2012). Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex*, 48(7), 788–804.

- Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, 69(3), 199–221.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4), 159–167.
- National Geographic Society. (2005). *Really wild animals: Totally tropical rainforests, deep sea dive, and swinging safari [DVDs]*. National Geographic School Publishing.
- Nicholas, H., & Lightbown, P. M. (2008). Defining child second language acquisition, defining roles for L2 instruction. In J. Philp, R. Oliver, & A. Mackey (Eds.), *Second language acquisition and the younger learner: Child's play?* (pp. 27–51). Amsterdam, Philadelphia: John Benjamins.
- Ostry, D. J., & Gribble, P. L. (2016). Sensory plasticity in human motor learning. *Trends in Neurosciences*, 39(2), 114–123.
- Oxford, R., & Crookall, D. (1990). Vocabulary learning: A critical analysis of techniques. *TESL Canada*, 7, 9–30.
- Paivio, A. (1971). Imagery and language. *Imagery: Current cognitive approaches* (pp. 7–32). New York, London: Academic Press.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45, 255–287.
- Paivio, A., & Csapo, K. (1969). Concrete image and verbal memory codes. *Journal of Experimental Psychology*, 80(2p1), 279.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231.
- Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Developmental Science*, 21(4), e12593.
- Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment effects of gestures and pictures on abstract words in a second language. *Frontiers in Psychology*, 8, 2136.
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688.
- Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing*. Routledge.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52.
- Saltz, E., & Dixon, D. (1982). Let's pretend: The role of motoric imagery in memory for sentences and words. *Journal of Experimental Child Psychology*, 34(1), 77–92.
- Saltz, E., & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 322–332.
- Sambanis, M. (2013). *Fremdsprachenunterricht und Neurowissenschaften*. Narr Francke Attempto Verlag.
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching*. Cambridge University Press.
- Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain, and Education*, 2(3), 114–121.
- Schwanenflugel, P. J. (1991). Why are abstract concepts hard to understand? In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 223–250). LEA.
- Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11), 411–417.
- Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology*, 101(2), 305.
- Spichtig, A., Pascoe, J., Ferrara, J., & Vorstius, C. (2017). A comparison of eye movement measures across reading efficiency quartile groups in elementary, middle, and high school students in the US. *Journal of Eye Movement Research*, 10(4), 5.
- Squire, L. R., & Zola-Morgan, M. (1991). The brain and memory: A comprehensive reference. *Cold Spring Harbor Perspectives in Biology*, 7(3), a021667.
- Storbeck, J., & Maswood, R. (2016). Happiness increases verbal and spatial working memory capacity where sadness does not: Emotion, working memory and executive control. *Cognition and Emotion*, 30(5), 925–938.
- Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, 21(1), 437–484.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1–2), 231–270.
- von Kriegstein, K., & Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biology*, 4(10), e326.

- von Kriegstein, K. (2012). A multisensory perspective on human auditory communication. In M. M. Murray & M. T. Wallace (Eds.), *The neural bases of multisensory processes* (pp. 683–700). CRC Press/Taylor & Francis.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232–245.
- Wellsby, M., & Pexman, P. M. (2014). Developing embodied cognition: Insights from children's concepts and language processing. *Frontiers in Psychology*, 5, 506.
- Winter, B. (2018). A very basic tutorial for performing linear mixed effects analyses: Tutorial 2. *QUBES Educational Resources*. <https://doi.org/10.25334/Q4W716>
- Yildirim, I., & Jacobs, R. A. (2012). A rational analysis of the acquisition of multisensory representations. *Cognitive Science*, 36(2), 305–332.
- Zemlock, D., Vinci-Booher, S., & James, K. H. (2018). Visual–motor symbol production facilitates letter recognition in young children. *Reading and Writing*, 31(6), 1255–1271.
- Zimmer, H. D., Helstrup, T., & Engelkamp, J. (2000). Pop-out into memory: A retrieval mechanism that is enhanced with the recall of subject-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(3), 658.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Brian Mathias^{1,2,3}  · **Christian Andrä**^{4,5} · **Anika Schwager**⁶ · **Manuela Macedonia**^{3,7,8}  · **Katharina von Kriegstein**^{2,3} 

✉ Brian Mathias
brian.mathias@abdn.ac.uk

¹ School of Psychology, University of Aberdeen, Aberdeen, Scotland, United Kingdom

² Chair of Cognitive and Clinical Neuroscience, Faculty of Psychology, Technical University Dresden, Dresden, Germany

³ Max Planck Research Group Neural Mechanisms of Human Communication, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

⁴ Department of Educational Sciences, Faculty of Human Sciences, University of Potsdam, Potsdam, Germany

⁵ University of Applied Sciences of Sport and Management Potsdam, Potsdam, Germany

⁶ Didactics of Physical Education, Institute of Primary and Pre-Primary Education, Faculty of Education, University of Leipzig, Leipzig, Germany

⁷ Institute for Information Engineering, Johannes Kepler University Linz, Linz, Austria

⁸ Center for Business Languages and Intercultural Communication, Johannes Kepler University Linz, Linz, Austria