Investigating Object Orientation Effects Across 18 Languages

Sau-Chin Chen[1], Erin Buchanan[2], Zoltan Kekecs[3,4], Jeremy K. Miller[5], Anna Szabelska[6], Balazs Aczel[3], Pablo Bernabeu[7], Patrick Forscher[8,9], Attila Szuts[3], Zahir Vally[10], Ali H. Al-Hoorie[11], Mai Helmy[12,13], Caio Santos Alves da Silva[14], Luana Oliveira da Silva[14], Yago Luksevicius de Moraes[14], Rafael Ming C. S. Hsu[14], Anthonieta Looman Mafra[14], Jaroslava V. Valentova[14], Marco Antonio Correa Varella[14], Barnaby Dixon[15], Kim Peters[15], Nik Steffens[15], Omid Ghaesmi[16], Andrew Roberts[16], Robert M. Ross[16], Ian D. Stephen[16,17], Marina Milyavskaya[18], Kelly Wang[18], Kaitlyn M. Werner[18], Dawn L. Holford[19], Miroslav Sirota[19], Thomas Rhys Evans[20], Dermot Lynott[7], Bethany M. Lane[21], Danny Riis[21], Glenn P. Williams[22], Chrystalle B. Y. Tan[23], Alicia Foo[24], Steve M. J. Janssen[24], Nwadiogo Chisom Arinze[25], Izuchukwu Lawrence Gabriel Ndukaihe[25], David Moreau[26], Brianna Jurosic[27], Brynna Leach[27], Savannah Lewis[27], Peter R. Mallik[27], Kathleen Schmidt[28], William J. Chopik[29], Leigh Ann Vaughn[30], Manyu Li[31], Carmel A. Levitan[32], Daniel Storage[33], Carlota Batres[34], Janina Enachescu[35], Jerome Olsen[35], Martin Voracek[35], Claus Lamm[36], Ekaterina Pronizius[36], Tilli Ripp[37], Jan Philipp Röer[37], Roxane Schnepper[37], Marietta Papadatou-Pastou[38], Aviv Mokady[39], Niv Reggev[39], Priyanka Chandel[40], Pratibha Kujur[40], Babita Pande[40], Arti Parganiha[40], Noorshama Parveen[40], Sraddha Pradhan[40], Margaret Messiah Singh[40], Max Korbmacher[41], Jonas R. Kunst[42], Christian K. Tamnes[42], Frederike S. Woelfert[42], Kristoffer Klevjer[43], Sarah E. Martiny[43], Gerit Pfuhl[43], Sylwia Adamus[44], Krystian Barzykowski[44], Katarzyna Filip[44], Patrícia Arriaga[45], Vasilije Gvozdenović[46], Vanja Kovic[46], Tao-tao Gan[47], Chuan-Peng Hu[48], Qing-Lan Liu[47], Zhong Chen[49], Fei Gao[49], Lisa

Li[49], Jozef Bavolár[50], Monika Hricová[50], Pavol Kacmár[50], Matúš Adamkovic[51,52], Peter Babincák[51], Gabriel Baník[51,52], Ivan Ropovik[52,53], Danilo Zambrano Ricaurte[54], Sara Álvarez Solas[55], Harry Manley[56], Panita Suavansri[56], Chun-Chia Kung[57], Belemir Çoktok[58], Asil Ali Özdogru[58], Çaglar Solak[59], Sinem Söylemez[59], Sami Çoksan[60], John Protzko[61], Ilker Dalgar[62], Vinka Mlakic[63], Elisabeth Oberzaucher[64], Stefan Stieger[63], Selina Volsa[63], Janis Zickfeld[65], & Christopher R. Chartier[27]

[1] Department of Human Development and Psychology, Tzu-Chi University, Hualien, Taiwan

[2] Harrisburg University of Science and Technology, Harrisburg, PA, USA

[3] Institute of Psychology, ELTE, Eotvos Lorand University, Budapest, Hungary

[4] Department of Psychology, Lund University, Lund, Sweden

[5] Department of Psychology, Willamette University,Salem OR, USA

[6] Institute of Cognition and Culture, Queen's University Belfast, UK

[7] Department of Psychology, Lancaster University, Lancaster, United Kingdom

[8] LIP/PC2S, Université Grenoble Alpes, Grenoble, France

[9] Busara Center for Behavioral Economics, Nairobi, Kenya

[10] Department of Clinical Psychology, United Arab Emirates University, Al Ain, UAE

[11] Royal Commission for Jubail and Yanbu, Jubail, Saudi Arabia

[12] Psychology Department, College of Education, Sultan Qaboos University, Muscat, Oman

[13] Psychology Department, Faculty of Arts, Menoufia University, Shebin El-Kom, Egypt

[14] Department of Experimental Psychology, Institute of Psychology, University of Sao Paulo, Sao Paulo, Brazil

[15] School of Psychology, University of Queensland, Brisbane, Australia

[16] Department of Psychology, Macquarie University, Sydney, Australia

[17] Department of Psychology, Nottingham Trent University, Nottingham, UK

[18] Department of Psychology, Carleton University, Ottawa, Canada

[19] Department of Psychology, University of Essex, Colchester, UK

[20] School of Social, Psychological and Behavioural Sciences, Coventry University, Coventry,

UK

[21] Division of Psychology, School of Social and Health Sciences, Abertay University, Dundee, UK

[22] School of Psychology, Faculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK.

[23] Department of Psychiatry and Psychological Health, Universiti Malaysia Sabah, Sabah, Malaysia

[24] School of Psychology, University of Nottingham Malaysia, Selangor, Malaysia

[25] Department of Psychology, Alex Ekwueme Federal University, Ndufu-Alike, Nigeria

[26] School of Psychology, University of Auckland, Auckland, NZ

[27] Department of Psychology, Ashland University, Ashland, OH, USA

[28] School of Psychological and Behavioral Sciences, Southern Illinois University, Carbondale, IL, USA

[29] Department of Psychology, Michigan State University, East Lansing, MI, USA

[30] Department of Psychology, Ithaca College, Ithaca, NY, USA

[31] Department of Psychology, University of Louisiana at Lafayette, Lafayette, LA, USA

[32] Department of Cognitive Science, Occidental College, Los Angeles, USA

[33] Department of Psychology, University of Denver, Denver, CO, USA

[34] Department of Psychology, Franklin and Marshall College, Lancaster, PA, USA

[35] Faculty of Psychology, University of Vienna, Wien, Austria

[36] Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Wien, Austria

[37] Department of Psychology and Psychotherapy, Witten/Herdecke University, Germany

[38] School of Education, National and Kapodistrian University of Athens, Athens, Greece

[39] Department of Psychology, Ben Gurion University, Beersheba, Israel

[40] School of Studies in Life Science, Pt. Ravishankar Shukla University, Raipur, India

[41] Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway

[42] Department of Psychology, University of Oslo, OSLO, Norway

[43] Department of Psychology, UiT - The Arctic University of Norway, Tromsø, Norway

[44] Institute of Psychology, Jagiellonian University, Krakow, Poland

[45] Iscte-University Institute of Lisbon, CIS-IUL, Lisbon, Portugal

[46] Laboratory for Neurocognition and Applied Cognition, Faculty of Philosophy, University of Belgrade, Belgrade, Serbia

[47] Department of Psychology, Hubei University, Wuhan, China

[48] School of Psychology, Nanjing Normal University, Nanjing, China

[49] Faculty of Arts and Humanities, University of Macau, Macau, China

[50] Department of Psychology, Faculty of Arts, Pavol Jozef Šafarik University in Košice, Košice, Slovakia

[51] Institute of Psychology, University of Presov, Prešov, Slovakia

[52] Institute for Research and Development of Education, Faculty of Education, Charles university, Prague, Czechia

[53] Faculty of Education, University of Presov, Prešov, Slovakia

[54] Faculty of Psychology, Fundación Universitaria Konrad Lorenz, Bogotá, Colombia

[55] Ecosystem Engineer, Universidad Regional Amazónica Ikiam, Tena, Ecuador

[56] Faculty of Psychology, Chulalongkorn University, Bangkok, Thailand

[57] Department of Psychology, National Cheng Kung University, Tainan, Taiwan

[58] Department of Psychology, Üsküdar University, Istanbul, Turkey

[59] Department of Psychology, Manisa Celal Bayar University, Manisa,Turkey

[60] Department of Psychology, Middle East Technical University, Ankara, Turkey

[61] Department of Psychological Science, Central Connecticut State University, New Britain, CT, USA

[62] Department of Psychology, Ankara Medipol University, Ankara, Turkey.

[63] Department of Psychology and Psychodynamics, Karl Landsteiner University of Health Sciences, Krems an der Donau, Austria

[64] Department of Evolutionary Anthropology, University of Vienna, Wien, Austria

[65] Department of Management, Aarhus University, Aarhus, Denmark

Author Note

Author contributions: Sau-Chin Chen contributed to the study concept, the design analysis protocol and wrote the initial report draft. Patrick Forscher, Pablo Bernabeu, Balazs Aczel and Attila Szuts improved the analysis protocol. Zoltan Kekecs, Jeremy K. Miller and Anna Szabelska managed the project administration which was established by Christopher R. Chartier. All the rest of authors contributed to the material prepation and data collection. All authors commented on previous versions of the manuscript, read and approved the final manuscript.

Funding statement. Below authors had the individual funds supporiting their participations. Glenn P. Williams was supported by the Leverhulme Trust Research Project Grant (RPG-2016-093). Krystian Barzykowski was supported by the National Science Centre, Poland (2019/35/B/HS6/00528). Zoltan Kekecs was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Science. Erin Buchanan was supported by the National Institute on Mental Health (1R03MH110812-01). Patrícia Arriaga was supported by the Portuguese National Foundation for Science and Technology (UID/PSI/03125/2019). Gabriel Baník was supported by Charles University Grant Agency (PRIMUS/20/HUM/009).

Ethical approval statement. Authors who collected data on site and online had the ethical approval/agreement from the local institute. The latest status of ethical approval for all the participating authors is available at the public OSF folder (https://osf.io/e428p/ "IRB approvals" in Files).

Acknowledgement. We appreciated the major contributions from the contributors as below. Chris Chartier and Jeremy Miller managed and monitored progress. Erin Buchanan provided guidelines to improve the inter-lab progress website management and managed the JATOS server for online data collection. Arti Parganiha, Asil Özdoğru, Attila Szuts, Babita Pande, Danilo Zambrano Ricaurte, Gabriel Baník, Harry Manley, Jonas Kunst, Krystian Barzykowski, Marco Antonio Correa Varella, Marietta Papadatou Pastou, Niv Reggev, Patrícia Arriaga, Stefan Stieger, Vanja Ković and Zahir Vally managed the material translation from English to the other languages. Roles of each collaborator are available in the public table (https://osf.io/mz97h/). We thank the suggestions from the editor and two reviewers on our first and second proposals.

Correspondence concerning this article should be addressed to Sau-Chin Chen, No. 67, Jei-Ren St., Hualien City, Taiwan. E-mail: csc2009@mail.tcu.edu.tw

Abstract

Mental simulation theories of language comprehension propose that people automatically create mental representations of the objects mentioned in sentences. One of the relevant paradigms is the sentence-picture verification task, in which participants first read a sentence and, on the following screen, see a picture of an object. Participants must verify whether the latter object had been mentioned in the sentence. Crucially, two covert conditions exist: the sentence and the picture can either match or mismatch in terms of a certain perceptual property. Usually, visual properties have been used, including object orientation, shape, color and size. The key finding obtained in some studies is the match advantage, whereby responses were faster in the match condition. The property of object orientation is noteworthy due to inconsistent findings across languages. After considering lexical and experimental explanations for those inconsistencies, this registered report describes our investigation of the match advantage of object orientation across 18 languages, which was undertaken by 33 laboratories and organized by the Psychological Science Accelerator. The preregistered analysis revealed that the match advantage was not significant either overall or in any specific language. We discuss the need for sample sizes that are far larger than usual, which are unequally accessible in different languages.

Keywords: mental simulation, object orientation, mental rotation, language comprehension

Word count: 5,138 words in total; Introduction: 1,242 words

Investigating Object Orientation Effects Across 18 Languages

Introduction

The simulation of object properties is a major topic in conceptual processing research (Ostarek & Huettig, 2019; Scorolli, 2014). One well-known method for assessing mental simulation during sentence reading is the sentence-picture verification task. The task requires participants to read a probe sentence displayed on the screen. On the following screen, the participants see a picture of an object and must verify whether the object was mentioned in the probe sentence. For example, the probe sentence, 'Tom hammered the nail into the wall' might be followed by a picture of a nail. Of central interest in the sentence-picture verification task is the match advantage, which occurs when people are faster to verify pictured objects whose properties match those of objects mentioned in the probe sentences. For example, in the sentence involving nails hammered into walls, an object orientation match advantage occurs when people are faster to verify horizontal nails than vertical ones. First discovered by Stanfield and Zwaan (2001), object orientation match advantages suggest that people mentally simulate objects during semantic processing (e.g., Barsalou, 1999, 2009). Researchers have found match advantages for shape (Zwaan et al., 2002) and color (Zwaan & Pecher, 2012; but see Connell, 2007). Consistent effects have appeared in English (Zwaan & Madden, 2005; Zwaan & Pecher, 2012), Chinese (Li & Shang, 2017), Dutch (De Koning et al., 2017; Engelen et al., 2011; Pecher et al., 2009; Rommers et al., 2013), German (Koster et al., 2018), Croatian (Šetić & Domijan, 2017), and Japanese (Sato et al., 2013). Object orientation, on the other hand, has produced mixed results across languages (Chen et al., 2020; De Koning et al., 2017; Koster et al., 2018; Zwaan & Madden, 2005; Zwaan & Pecher, 2012). To scrutinize the discrepancies across languages, we tested the match advantage of object orientation across 18 languages by means of a multi-lab collaboration. Among object properties, shape and color are described as intrinsic, which mean that these properties are relatively independent from the observer's state. That is, regardless of the observers' location or state, the object keeps its intrinsic properties constant. In contrast, an extrinsic property such as object orientation is dependent on the observer's state. Studies on visual simulation have suggested that match advantages for intrinsic properties are more consistent than those for extrinsic properties (De Koning et al., 2017; Koster et al., 2018). On the other hand, studies on motor simulation (Beilock et al., 2008; Glenberg & Kaschak, 2002) showed the sentences that imply extrinsic properties (e.g., Joe sends the card to you) facilitated the responses to a congruent target (e.g., a "card" presented in a larger size). A later study suggests that orientation match advantages are stronger for large objects (Chen et al., 2020). Taken together, these results suggest that both intrinsic and extrinsic properties produce match advantages, though the advantage may be smaller for extrinsic properties.

Cross-linguistic and Experimental Factors

Several factors might contribute to cross-linguistic differences in the match advantage of orientation, including cultural, lexical and experimental ones. First, Ghandhari et al. (2020) found that cultural differences between Persian and Italian participants influenced the action sentence compatibility effect (for further differences studied in psycholinguistics, see Norcliffe et al., 2015). Second, languages differ in how they encode motion and placement events in sentences. Last, the potential role of mental rotation as a confound has been considered. We expand on the lexical and experimental factors below.

Lexical Factors. The probe sentences used in object orientation studies usually contain several motion events (e.g., "The ant walked towards the pot of honey and tried to climb in."). The languages we probed in this study encode motion events in different ways, and grammatical differences between them could explain the different match advantage results. According to Verkerk (2014), Germanic languages (e.g., Dutch, English, German) generally encode the manner of motion in the verb (e.g., 'The ant dashed'), while conveying the path information through satellite adjuncts (e.g., 'towards the pot of honey'). In contrast, other languages, such as the Romance family (e.g., Portuguese, Spanish) more often encode path in the verb (e.g., ‘crossing,’ ‘exiting’). Crucially, the past research on the match advantage of object orientation is exclusively based on Germanic languages, and yet, there were differences across those languages, with English being the only one that consistently yielded the match advantage. As a minor difference across Germanic languages in this regard, Verkerk notes that path-only constructions (e.g., ‘The ant went to the feast’) are more common in English than in other Germanic languages.

Another topic to be considered is the lexical encoding of placement in each language, as the stimuli contains several placement events (e.g., ‘Sara situated the expensive plate on its holder on the shelf.’). Chen et al. (2020) and Koster et al. (2018) noted that some Germanic languages, such as German and Dutch, often make the orientation of objects more explicit than English. Whereas in English one could use the verb “put” in both “She put the book on the table” and “She put the bottle on the table,” in both Dutch and German, one could instead say “She laid the book on the table,” and “She stood the bottle on the table.” In these literal translations from German and Dutch, the verb “lay” encodes a horizontal orientation, whereas the verb “stand” encodes a vertical orientation. This distinction extends to verbs indicating existence. As Newman (2002) exemplified, an English speaker would be likely to say “There’s a lamp in the corner,” whereas a Dutch speaker would be more likely to say “There ‘stands’ a lamp in the corner.” Nonetheless, we cannot conclude that these cross-linguistic differences are affecting the match advantage across languages because there is no previous theory allowing predictions, and because placement events are not sufficiently prevalent in the stimuli.

Experimental factors. Stanfield and Zwaan (2001) found that participants who sufficiently understood the probe sentences showed a match advantage of object orientation. Later studies on this topic have examined the association between the match advantage and cognitive abilities. Spatial cognition is one of the relevant areas, which may be measured with mental rotation tasks. Studies have suggested that mental rotation tasks offer valid reflections of previous spatial experience (Frick & Möhring, 2013) and of current spatial cognition (Chu & Kita, 2008; Pouw et al., 2014). De Koning et al. (2017) suggested that mental rotation, as an alternative process to mental simulation, could quickly erase the mismatched orientation, replacing it with the orientation that matches the one described in the sentence (Cohen & Kubovy, 1993; Yaxley & Zwaan, 2007). Chen et al. (2020) investigated the relationship between the match advantage and mental rotation across three languages: English, Dutch and Chinese. They introduced the picture-picture verification task to examine how individuals process the target pictures regardless of their native language. This picture-picture verification task was designed using the mental rotation paradigm (Cohen & Kubovy, 1993). In each trial of this task, two pictures appear on opposite sides of the screen. Participants have to verify whether the pictures represent identical or different objects. The verification times for pictures of identical objects presented in the same orientation (i.e., two identical pictures presented in horizontal or vertical orientation) were shorter than those presented in different orientations (one horizontal; one vertical). Chen et al.'s findings suggested that crosslinguistic differences in the match advantage of object orientation are not confounded by mental rotation strategies.

Purposes of this study

Several explanations have been proposed for the inconsistent findings on the match advantage of object orientation, including the procedural details of some studies, such as participants not being required to verify the probe sentences they had read (see Zwaan, 2014). Without such a verification, participants might have paid less attention to the meaning of the probe sentences, in which case they would have been less likely to form a mental representation of the objects (e.g., Zwaan & van Oostendorp, 1993). In this regard, it is relevant to acknowledge that variability originating from individual differences and other characteristics of experiments can substantially influence the results (Barsalou, 2019).Thus, this study followed the original methods from Stanfield and Zwaan (2001) and addressed two primary questions: (1) How much of the match advantage of object orientation can be obtained within different languages and (2) How do differences in mental rotation affect the match advantage across languages?

Methods

Hypotheses and Design

Both the sentence-picture verification task and the picture-picture verification task involve a between-participant and a within-participant independent variable. The between-participant variable is the 18 languages registered in this study. The within-participant variable is the match or mismatch in object orientation. This binary factor, in the sentence-picture verification task, reflects the matching between the sentence and the picture, whereas in the picture-picture verification, it reflects the orientation settings between two pictures. The only dependent variable for both tasks is the response time. In the sentence-picture verification task, we expect response time to be shorter for matching compared to mismatching orientations. We expect to see the match advantage within each language. We did not select languages systematically, but instead based on who our collaborators could recruit. We did not have any specific hypotheses about the relative size of the object orientation match advantage in different languages. In the picture-picture verification task, we expect shorter response time for identical orientation compared to different orientations. We computed an imagery score by subtracting the verification time for identical orientation from the verification time for different orientations. Based on the assumption that the mental rotation is a general cognitive aspect, we expect imagery scores to be the same on average across languages, and can be used to predict a possible match advantage (see Chen et al., 2020).

Participant

Through the collaboration of The Psychological Science Accelerator (Moshontz et al., 2018), we collected data in 18 languages. Our priori power analysis recommended a language would have at least one thousand participants based on the current design[1]. Only English data approached this number because 17 laboratories recruited native English speakers. Based on the preregistered plan, the available participants' accuracy had to reach 70%. Before the pandemic outbreak, 2,340 participants (1,104 women; M = 21.46 years old) from 33 laboratories joined and finished the study. After the study migrated online, there were 1403 participants (926 women; M = 23.75 years old) from 1,403 laboratories completed the study. Web-based participants at the beginning heard the auditory instruction and had to correctly answer at least 2 of 3 comprehension check questions about the instructions. All participating laboratories had ethical approval before data collection. Appendix 1 summarizes the average characteristics by language and laboratory.

General Procedure and Materials

Participating laboratories conducted the tasks as follows. In the beginning of the sentence-picture verification task, participants had to correctly answer all the practice trials as the instruction. Each trial started with a left-justified and horizontally centered fixation point displayed for 1000 ms, immediately followed by the probe sentence. The sentence was presented until the participant pressed the space key, acknowledging that they understood the sentence. Then, the object picture was presented in the center of the screen until the participant responded otherwise it disappeared after 2 seconds. Participants were instructed to verify the object picture mentioned in the probe sentence as quickly and accurately as they could. Following the original study (Stanfield & Zwaan, 2001), a memory check test was carried out after every three to eight trials to ensure that the participants had read each

[1] See details of power analysis in the preregistered plan, p. 13 ~ 15. https://psyarxiv.com/t2pjv/

sentence carefully.

The picture-picture verification task used the same object pictures. In each trial, two objects appeared on either side of the central fixation point until either the participant indicated that the pictures displayed the same object or two different objects or until 2 seconds elapsed. Two pictures showing the same critical object appeared in each "yes" trial; two pictures showing two different objects from the filler items appeared in each "no" trial.

All the procedures are compiled in OpenSesame scripts (Mathôt et al., 2012). Before the Covid-19 pandemic broke out, 29 participating laboratories had completed data collection. The remaining laboratories had to stop data collection because of local lockdowns. The project team decided to move data collection online. To minimize the differences between on-site and web-based studies , we converted the original Python code to Javascript and collected the data through a JATOS server (Lange et al., 2015). After the changes in the procedure were approved by the journal editor and reviewers, we proceeded with the online study from February to June 2021. For the remote version, a recorded set of verbal instructions was played at first. Participants had to confirm they were native speakers of the targeted language. All verbal briefings were packaged in the language-specific scripts. Appendix 2 describes the deployment of the scripts and the results of participants' fluency tests. Following the literature, we did not anticipate any theoretically important differences between the two data sources (see Anwyl-Irvine et al., 2020; Bridges et al., 2020; de Leeuw & Motz, 2016). The instructions and experimental scripts are available at the public OSF folder (https://osf.io/e428p/ "Materials" in Files).

Analysis plan

Confirmatory Analysis According to our preregistered analysis plan[2], this study used meta-analysis and mixed-effect models to estimate the match advantage across

[2] See the analysis plan in the preregistered plan, p. 19 ~ 20. https://psyarxiv.com/t2pjv/

languages. The meta-analysis summarized the median reaction times by match condition to determine the global effect size. This approach was compatible with ANOVA used by the original study (Stanfield & Zwaan, 2001). The mixed-effect models involved the actual response time and analysed the fixed effects using mixed-effects models (Baayen et al., 2008). This approach was used by recent studies (Chen et al., 2020; Koster et al., 2018). Without the systematic comparison of pros and cons between the two approaches, the current analysis employed two approaches to estimate the match advantage. The statistical analyses were conducted by R packages including *metafor* for meta analysis (Viechtbauer, 2010), *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) for mixed-effects models, as well as multiple regression through R base package (Version 4.1.1; R Core Team, 2021).

Imagery scores are the dependent measure of the picture-picture verification responses. Tidied response times were summarized by the difference between the identical and different orientation. According to our preregistered analysis plan,[3] we first evaluated the equality of imagery scores across languages in use of the mixed-effects models. Our other linear regression analysis evaluated the imagery scores as the predictor of match advantage. In a best fit model having the imagery score as the predictor, the slope would indicate its accountability.

Exploratory Analysis In one of the cases below we conducted the mixed-effect models for some language dataset. At first the total sample size reached recommended sample size as our prior power analysis. Otherwise the meta-analysis indicated a language dataset showed a significant match advantage. Although this analysis was not in the preregistered analysis plan, the authors contributed to the methodology agreed this analysis could improve the reliability of the linguistic-specific result.

Decision criterion P values were interpreted using the preregistered alpha level of .05. Because in our preregistered plan each language was assumed a standalone group, P values

[3] See the analysis plan in the preregistered plan, p. 21. https://psyarxiv.com/t2pjv/

of the analysis by each language were not corrected (Armstrong, 2014). All the final mixed-effects models were selected by pursuing a maximal random-effects structure whilst allowing the model to converge (Bates et al., 2015). P values for each effect were calculated using the Satterthwaite approximation for degrees of freedom (Luke, 2017).

Results

Within the data collected on-site, 1,979 participants finished the sentence-picture verification task and met the preregistered inclusion criterion (accuracy percentile > 70%); 2,007 participants finished the picture-picture verification task. Raw data files containing data for twenty-eight participants were lost due to human error. Within the data sets collected online, 1,337 participants finished the sentence-picture verification task and met the preregistered inclusion criterion; 1,402 participants finished the picture-picture verification task. All data and analyses are available on the source files (https://osf.io/p7avr/).

Confirmatory analysis: Intra-lab analysis during data collection

Before data collection, each lab decided whether they wanted to apply a sequential analysis (Schönbrodt et al., 2017) or whether they wanted to settle for a fixed sample size. The preregistered protocol for labs applying sequential analysis established that they could stop data collection upon reaching the preregistered criterion ($BF_{10} = 10 \; or \; -10$), or the maximal sample size. Most laboratories either chose a fixed sample size without applying sequential analysis, or applied sequential analysis and reached their maximal sample size.

Two laboratories (HUN 001, TWN 001) stopped data collection at the preregistered criterion. Some laboratories did not conduct the sequential analysis on all their data because of one of the following reasons: (1) their data collection was interrupted by the pandemic outbreak; (2) participants performed worse in the online study; (3) too many of their participants were non-native speakers. Lab-specific results were reported on a public website as each laboratory completed data collection (details available in Appendix 2).

Confirmatory analysis: Inter-lab analysis of final data

Identification of outliers. For each laboratory, outliers were identified by the third quantile of the grand intercept in the simplest mixed-effects model. This mixed-effects model contained the response times as the dependent measure, matching condition as the only fixed effect, and the participant as the only random intercept. Among the data sets showing outliers, the averaged proportion of outliers was 0.25. Table S4 in Appendix 1 illustrates the distribution of outliers by laboratory. Table 1 and Table 2 respectively summarise the match advantages by language. All the below data analysis depended on the datasets excluding the outliers.

(Insert Table 1 about here)

(Insert Table 2 about here)

Meta-analysis of match advantages across laboratories. Because the preregistered analysis plan did not consider the data collected online, we conducted the overall meta-analyses for all the datasets combined data sources. In this analysis, we computed the effect size by data set and estimated the global effect size. Since data from small samples may contribute to a biased estimate, nine datasets with sample sizes smaller than 25 were excluded from the analyses. The overall meta-analysis found no match advantage (Figure 1).Among the languages that had at least two datasets, we conducted the meta-analysis for English, German, Norway, Traditional Chinese, Slovak, and Turkey. Only Traditional Chinese showed a significant meta-analytic effect across laboratories(see Figure 2). Results of the other languages are available in Appendix 3.

(Insert Figure 1 about here)

(Insert Figure 2 about here)

Evaluating match advantages using linear mixed-effects models. Considering the bias of small sample size, we excluded the languages with below 25 participants in each data source before conducting the mixed-effects models. Thus we excluded Portuguese in the on-site data and Norwegian in the web-based data. Because the sources of data collection included the labs and the web, we had to evaluate whether one mixed-effects model sufficiently fitted all the data. Otherwise, separate models would be needed for each data set. This analysis showed a significant difference between data sources: b = -744.624, SE = 119.447, t(9.188) = -6.234, $p < .001$. Thus, the on-site and the web-based data had to be analyzed separately.

The final models examined the interaction between language and match advantage in each data source, as reported below. All other models are reported in Appendix 4. It must be acknowledged that the languages with larger sample sizes (see Tables 1 and 2) have more reliable results. Furthermore, most of the languages were underpowered, being far from the 1,200 participants suggested by an a priori power analysis.

In each data source, we compared the fit of the models with and without the random slope of matching condition. Both indicated that the models without the random slope had the best fit. The model from the on-site data revealed no significant effect of match advantage: b = 5.729, SE = 3.595, t(32924.494) = 1.594, p = 0.111. The model from the web-based data also failed to reveal a significant effect: b = -3.921, SE = 255.165, t(25484) = -0.015, p = 0.988. The latter model had a negative coefficient, unlike the on-site data. Although neither effect was significant, the difference in direction resounds with the match advantages and disadvantages found in experiments using the property of color (cf. Connell, 2007; Zwaan & Pecher, 2012).

Figure 3 illustrates the response times from the on-site data. Ten languages presented significant intercepts (see “Models including languages” section in Appendix 4).

(Insert Figure 3 about here)

Figure 4 illustrates the response times in the web-based data. Four languages presented significant effects (see "Models included languages" section in Appendix 4).

(Insert Figure 4 about here)

Anecdotal evidence on the match advantage. In the on-site data, only Greek presented a match advantage, b = 21.718, SE = 9.997, t(32951.681) = 2.173, p = 0.030. It should be noted, however, that these results are not robust due to the underpowered sample sizes (see Discussion).

The mean response times in Greek (M = 739.70, SD = 263.87) and Serbian (M = 1,990.70, SD = 10,335.91) was longer than the average across languages (M = 955.37, SD = 3,204.99). This might not be coincidental, as according to Yap et al. (2014), longer response times have been associated with larger effects in psycholinguistics (Schilling et al., 1998; Seidenberg, 1985; Tainturier, 1992).

Analysis of imagery scores. Prior to data collection, we assumed the imagery scores of every language group would be nearly equal. The best-fitting model included random intercepts for participants, targets and laboratories but no slopes for orientation. The fixed effect of orientation match was significant, b = 33.348, SE = 5.442, t(59435.963) = 6.127, p $<$ .001. The response times illustrated in Figure 5 indicated that the imagery scores measured for each language were consistently positive supporting our hypothesis. The coefficients of all evaluated mixed-effects models are reported in Appendix 5.

(Insert Figure 5 about here)

The above analyses suggested that data sources did not influence the imagery scores but did influence the match advantage. Therefore, we evaluated the fit of the model with languages and imagery scores and the model with languages only. Both models included match advantage as the dependent variable. If imagery scores predict match advantage, the model with languages and imagery scores should fit the data better than the model with languages only. Because the random slopes for items in the analyses of the match advantage were zero (see Appendix 5), the data for building the regression models were the aggregated data by participants.

In the linear regression analysis, we selected the best fit model from the model with only one predictor, language, and the model with two predictors, language and imagery scores. Because the analysis of match advantage revealed a difference between on-site versus web-based data, we conducted separate regression analyses for the two data sources. In the analysis of the on-site data, the model with language and imagery scores had yet fit better than the model with language only, $F = 1.369$, $p = 0.161$. In contrast, in the analysis of the web-based data, the model with language and imagery scores had a better fit than the model with language only, F (8,1132) = 2.604, p =0.161. In the latter case, the effect of imagery scores was nonsignificant, $b = -0.28$, 95% CI $[-1.68, 1.12]$, $t(1132) = -0.40$, $p = .691$. Appendix 5 summarized the coefficients of the models included in these analyses.

Exploratory analysis: language-specific match advantages

Based on the policy to conduct the linguistic-specific mixed-effect models, we selected the English datasets (N = 1,216) and the Traditional Chinese datasets (N = 150). For both languages, we are interested in whether the data sources could inhibit the match advantage. Another topic of interest is if the match advantage changed with English dialects, namely American English and British English.

Using the data from 1,216 English speaking participants, we ran a mixed-effects model for the English data containing orientation match condition, English dialects (American vs. British) and data sources (on-site vs. web-based) as fixed effects. Following Brauer and Curtin (2018), English dialects and data sources were numerically recoded. The best fitted model indicated that only data source (on-site vs. web-based) was significant, b = -869.698, SE = 131.861, t(4314.646) = -6.596, p < .001. Although the match advantage of orientation was nonsignificant, this exploratory analysis indicated the interaction of orientation match condition and English dialects: b = -446.255, SE = 192.479, t(26059.644) = -2.318, p = 0.020 (see the detailed report in Appendix 4).

We conducted another exploratory mixed-effect model on Traditional Chinese data because his was the only language to show a significant result in the preregistered meta-analysis. The best fit model had orientation match condition and data sources as the fixed effects. This model indicated that data source was significant, b = -480.819, SE = 43.81, t(326.144) = -10.975, $p < .001$. The match advantage of orientation was nearly significant: b = -78.485, SE = 40.065, t(3186.213) = -1.959, p = 0.053, but the interaction of match advantage and data sources was nonsignificant: b = 82.833, SE = 50.91, t(3186.075) = 1.627, p = 0.104 (see the detailed report in Appendix 4). This result suggested that Traditional Chinese study could have a robust estimation in the circumstance multiple teams conducted the study in terms of one the same protocol. Combined with the previous results of Traditional Chinese (Chen et al., 2020), future research on this language could explore any potential linguistic aspects that might result in the match advantage of object orientation and other properties. Although this study is unable to provide further advice, the advantage for the future Traditional Chinese studies would be a precise sample size justification on the participants and stimulus items.

Discussion

After estimating the match advantage of object orientation across 18 languages, no evidence for a global effect was found, but the meta-analysis and mixed-effect models indicated the marginal match advantage was present in the investigated languages that had at least two datasets. Traditional Chinese, especially, showed marginal results which are consistent with the findings of Chen et al. (2020). This suggests that the match advantage of object orientation for many languages are small. Thus large sample sizes are needed to determine if the match advantage for a language is significantly different from zero. One exception might be a language has some unique features that amplify the match advantage.

This requirement is especially onerous in cross-linguistic studies when it is difficult to reach the desirable sample size. In sum, the present results put into question the robustness of cross-linguistic studies.

The second question addressed whether the mental simulation of object orientation could be predicted by mental rotation, which was operationalized as imagery scores. The mixed-effects models indicated that imagery scores were hardly affected by the different languages and data collection procedures. Regarding the planned regression analysis, the imagery scores underpredicted the effect of orientation match. In conclusion, the current findings barely confirm the predictions based on the mental simulation theory.

Measurement issues across platforms

The precision of web-based experiments has been previously investigated (Anwyl-Irvine et al., 2020; Bridges et al., 2020). In the present study, the responses to the sentence-picture verification task collected on the web (M = 1,573, SD = 6,945) were roughly twice as long as those collected in labs (M = 695, SD = 258). Previous studies have also found online responses to be longer than on-site ones, but the current difference is larger. For instance, de Leeuw and Motz (2016) collected response times of just under 100 ms, and found that online responses were 10–40 ms longer than on-site ones. Our primary concern was whether OpenSesame could have caused a higher measurement error in the on-site data than in the web-based data. Of the frequently used desktop applications, OpenSesame Windows version has the highest precision and relatively low variation (see Table 2 in Bridges et al., 2020). Although Bridges et al. did not evaluate the performance of OSWeb, PsychoPy (Peirce et al., 2019), which is the basis for OpenSesame, had higher precision than OpenSesame desktop version. Specifically, it had a 25 to 50 ms lag (see Table 3 of Bridges et al., 2020) in many combinations of operating systems and web browsers. This lag is shorter than the response time difference between the data sources, suggesting that measurement precision was not the source of the timing discrepancies between on-site and web-based data.

According to our meta-analysis, the data from 19 teams revealed marginal match advantages: 12 of these teams collected data on site (e.g., NOR 003) and 5 collected data online (e.g., NZL 005). The former example of the online data is from a lab testing in Norwegian, a language in which the match advantage of orientation had not been studied before, to our knowledge. The latter example is from a lab testing in English, a language that has yielded marginal match advantages of object orientation before (Chen et al., 2020; Stanfield & Zwaan, 2001; Zwaan & Pecher, 2012).

Generalizability and Limitations

To acknowledge deviations from the preregistration, it must be noted that the final data included four more languages than were initially planned, and that some data were collected on the web due to the Covid-19 pandemic, instead of in labs as planned. We do not know of any research suggesting how these deviations could have affected our results. For instance, we reviewed research suggesting that on-site and web-based data need not substantially differ (e.g., de Leeuw & Motz, 2016).

This study reflected the difficulty of investigating cognition across languages, especially when dealing with effects that require large sample sizes (see Loken & Gelman, 2017; Vadillo et al., 2016). Indeed, a fundamental challenge for our project was substantial variation in the number of participants available for the languages we investigated. It must also be noted that the mixed-effects models could have been more conservative by prioritizing the maximal random-effects structure over the achievement of model convergence (Brauer & Curtin, 2018).

Some languages in this study had one participating team only or did not have sufficient data for the exploratory analysis: Arabic, Brazilian Portuguese, European Portuguese, Greek, Hebrew, Hindi, Hungarian, Polish, Serbian, Simplified Chinese and Thai. As a consequence, we do not know if these languages would show a robust match advantage if they were studied by more than two laboratories plus a sample size as large as that which we had for English.

Researchers could use stimuli employed for the present study to launch new studies that focus on these specific languages. With rigorous power analyses and large multi-site collaborative projects, future research on specific languages has the potential to provide robust estimates of match advantages associated with a variety of object properties. Such research could serve as a firm foundation for developing a strong theory of mental simulation.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barsalou, L. W. (2019). Establishing generalizable mechanisms. *Psychological Inquiry*, *30*(4), 220–230. https://doi.org/10.1080/1047840X.2019.1693857

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660. https://doi.org/10.1017/S0140525X99002149

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*, 1281–1289. https://doi.org/10.1098/rstb.2008.0319

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C., & Small, S. L. (2008). Sports experience changes the neural processing of action language. *Proceedings of the National Academy of Sciences*, *105*(36), 13269–13273. https://doi.org/10.1073/pnas.0803424105

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, *23*(3), 389–411. https://doi.org/10.1037/met0000159

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

Chen, S.-C., de Koning, B. B., & Zwaan, R. A. (2020). Does object size matter with regard to the mental simulation of object orientation? *Experimental Psychology*, *67*(1), 56–72. https://doi.org/10.1027/1618-3169/a000468

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, *137*(4), 706–723. https://doi.org/10.1037/a0013157

Cohen, D., & Kubovy, M. (1993). Mental rotation, mental representation, and flat slopes. *Cognitive Psychology*, *25*, 351–382. https://doi.org/10.1006/cogp.1993.1009

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*, 476–485. https://doi.org/10.1016/j.cognition.2006.02.009

De Koning, B. B., Wassenburg, S. I., Bos, L. T., & Van der Schoot, M. (2017). Mental simulation of four visual object properties: Similarities and differences as assessed by the sentence-picture verification task. *Journal of Cognitive Psychology*, *29*(4), 420–432. https://doi.org/10.1080/20445911.2017.1281283

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*(1), 1–12. https://doi.org/10.3758/s13428-015-0567-2

Engelen, J. A. A., Bouwmeester, S., de Bruin, A. B. H., & Zwaan, R. A. (2011).

Perceptual simulation in developing language comprehension. *Journal of Experimental Child Psychology*, *110*(4), 659–675. https://doi.org/10.1016/j.jecp.2011.06.009

Frick, A., & Möhring, W. (2013). Mental object rotation and motor development in 8- and 10-month-old infants. *Journal of Experimental Child Psychology*, *115*(4), 708–720. https://doi.org/10.1016/j.jecp.2013.04.001

Ghandhari, M., Fini, C., Da Rold, F., & Borghi, A. M. (2020). Different kinds of embodied language: A comparison between Italian and Persian languages. *Brain and Cognition*, *142*, 105581. https://doi.org/10.1016/j.bandc.2020.105581

Glenberg, A., & Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565. https://doi.org/10.3758/bf03196313

Koster, D., Cadierno, T., & Chiarandini, M. (2018). Mental simulation of object orientation and size: A conceptual replication with second language learners. *Journal of the European Second Language Association*, *2*(1). https://doi.org/10.22599/jesla.39

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lange, K., Kühn, S., & Filevich, E. (2015). "Just Another Tool for Online Studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLOS ONE*, *10*(6), e0130834. https://doi.org/10.1371/journal.pone.0130834

Li, Y., & Shang, L. (2017). An ERP study on the mental simulation of implied object color information during Chinese sentence comprehension. *Journal of Psychological Science*, *40*(1), 29–36. https://doi.org/10.16719/j.cnki.1671-6981.20170105

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, *355*(6325), 584–585. https://doi.org/10.1126/science.aal3618

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Moshontz, H., Campbell, L., Ebersole, C. R., IJzerman, H., Urry, H. L., Forscher, P. S., Grahe, J. E., McCarthy, R. J., Musser, E. D., Antfolk, J., Castille, C. M., Evans, T. R., Fiedler, S., Flake, J. K., Forero, D. A., Janssen, S. M. J., Keene, J. R., Protzko, J., Aczel, B., … Chartier, C. R. (2018). The Psychological Science Accelerator: Advancing psychology through a distributed collaborative network. *Advances in Methods and Practices in Psychological Science*, *1*(4), 501–515. https://doi.org/10.1177/2515245918797607

Newman, J. (2002). 1. A cross-linguistic overview of the posture verbs 'Sit,' 'Stand,' and 'Lie.' In J. Newman (Ed.), *Typological Studies in Language* (Vol. 51, pp. 1–24). John Benjamins Publishing Company. https://doi.org/10.1075/tsl.51.02new

Norcliffe, E., Harris, A. C., & Jaeger, T. F. (2015). Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition and Neuroscience*, *30*(9), 1009–1032. https://doi.org/10.1080/23273798.2015.1080373

Ostarek, M., & Huettig, F. (2019). Six Challenges for Embodiment Research. *Current Directions in Psychological Science*, *28*(6), 593–599. https://doi.org/10.1177/0963721419866441

Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *The Quarterly Journal of Experimental Psychology*, *62*(6), 1108–1114. https://doi.org/10.1080/17470210802633255

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00359

Rommers, J., Meyer, A. S., & Huettig, F. (2013). Object shape and orientation do not routinely influence performance during language processing. *Psychological Science*, *24*(11), 2218–2225. https://doi.org/10.1177/0956797613490746

Sato, M., Schafer, A. J., & Bergen, B. K. (2013). One word at a time: Mental representations of object shape change incrementally during sentence processing. *Language and Cognition*, *5*(04), 345–373. https://doi.org/10.1515/langcog-2013-0022

Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–1281. https://doi.org/10.3758/BF03201199

Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322–339. https://doi.org/10.1037/met0000061

Scorolli, C. (2014). Embodiment and language. In L. Shapiro (Ed.), *The Routledge* *handbook of embodied cognition* (pp. 145–156). Routledge.

Seidenberg, M. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*(1), 1–30. https://doi.org/10.1016/0010-0277(85)90029-0

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Šetić, M., & Domijan, D. (2017). Numerical Congruency Effect in the Sentence-Picture Verification Task. *Experimental Psychology*, *64*(3), 159–169. https://doi.org/10.1027/1618-3169/a000358

Tainturier, M. (1992). Educational level and the word frequency effect: A lexical decision investigation*1. *Brain and Language*, *43*(3), 460–474. https://doi.org/10.1016/0093-934X(92)90112-R

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, *23*(1), 87–102. https://doi.org/10.3758/s13423-015-0892-6

Verkerk, A. (2014). *The evolutionary dynamics of motion event encoding.* [Radboud Universiteit Nijmegen]. https://repository.ubn.ru.nl/handle/2066/127455

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(1), 1–48. https://doi.org/10.18637/jss.v036.i03

Yap, M. J., Sibley, D. E., Balota, D. A., Ratcliff, R., & Rueckl, J. (2014). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, No Pagination Specified. https://doi.org/10.1037/xlm0000064

Yaxley, R. H., & Zwaan, R. A. (2007). Simulating visibility during language comprehension. *Cognition*, *105*(1), 229–236.

https://doi.org/10.1016/j.cognition.2006.09.003

Zwaan, R. A. (2014). Replications Should Be Performed With Power and Precision: A Response to Rommers, Meyer, and Huettig (2013). *Psychological Science*, *25*(1), 305–307. https://doi.org/10.1177/0956797613509634

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 224–245). Cambridge University Press.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PLoS ONE*, *7*, e51382. https://doi.org/10.1371/journal.pone.0051382

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*, 168–171. https://doi.org/10.1111/1467-9280.00430

Zwaan, R. A., & van Oostendorp, H. (1993). Do readers construct spatial representations in naturalistic story comprehension? *Discourse Processes*, *16*(1-2), 125–143. https://doi.org/10.1080/01638539309544832

Table 1

Median reaction times and accuracy percentages (in parentheses) per match condition (Mismatching, Matching); Match advantage (difference in response times) by language in the on-site data.

Language	N	Mismatching	Matching	Match Advantage
English	473	567(93.73)	561(94.63)	6.00
German	74	553(95.50)	555(95.61)	-2.00
Greek	73	701(90.53)	692(91.10)	9.00
Hebrew	109	556(97.09)	568(95.57)	-12.00
Hindi	59	594(87.57)	627(91.24)	-32.50
Hungarian	97	607(95.45)	614(95.62)	-7.50
Norwegian	92	577(96.29)	581(96.47)	-3.75
Polish	37	561(94.82)	585(95.72)	-24.00
Portuguese	5	588(96.67)	580(98.33)	8.00
Simplified Chinese	60	632(91.67)	619(92.78)	13.75
Slovak	103	589(96.60)	606(94.82)	-17.00
Spanish	95	646(92.63)	644(92.54)	2.00
Thai	37	637(90.77)	596(87.61)	41.50
Traditional Chinese	78	611(94.23)	616(93.27)	-5.25
Turkish	137	631(95.26)	618(94.89)	12.50

Table 2

Median reaction times and accuracy percentages (in parentheses) per match condition (Mismatching, Matching); Match advantage (difference in response times) by language in the web-based data.

Language	N	Mismatching	Matching	Match Advantage
Arabic	79	522(77.95)	544(76.69)	-21.50
Brazilian Portuguese	50	778(94.00)	779(94.00)	-0.75
English	630	644(93.16)	645(93.51)	-1.00
German	127	657(95.60)	668(95.21)	-11.50
Norwegian	16	844(96.88)	636(93.75)	208.25
Portuguese	40	672(96.46)	664(94.79)	9.00
Serbian	129	712(93.48)	722(94.38)	-9.50
Traditional Chinese	34	616(91.18)	664(95.34)	-47.75
Turkish	59	779(92.80)	780(90.68)	-1.00

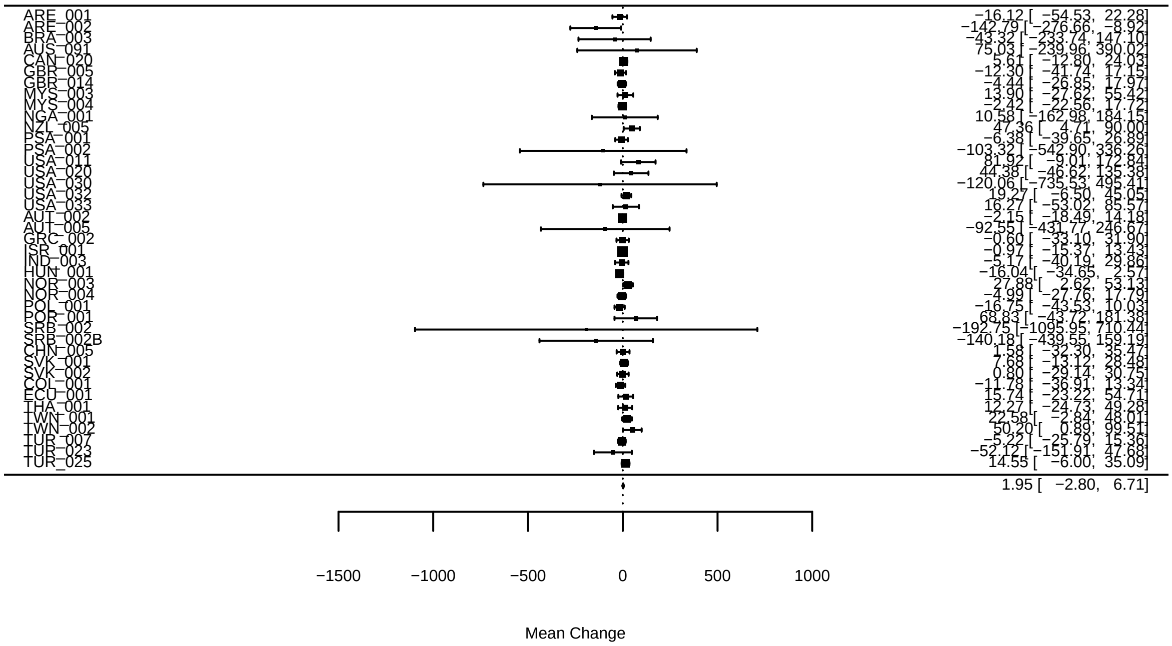

Figure 1. Meta-analysis on match advantage of object orienation for all datasets

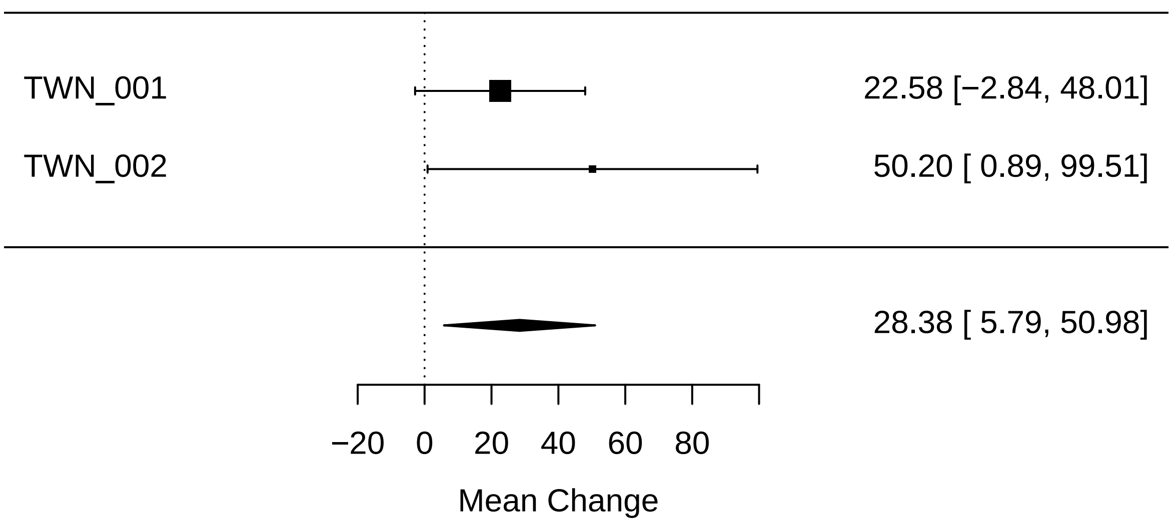

Figure 2. Meta-analysis on match advantage of object orienation for Traditional Chinese datasets.

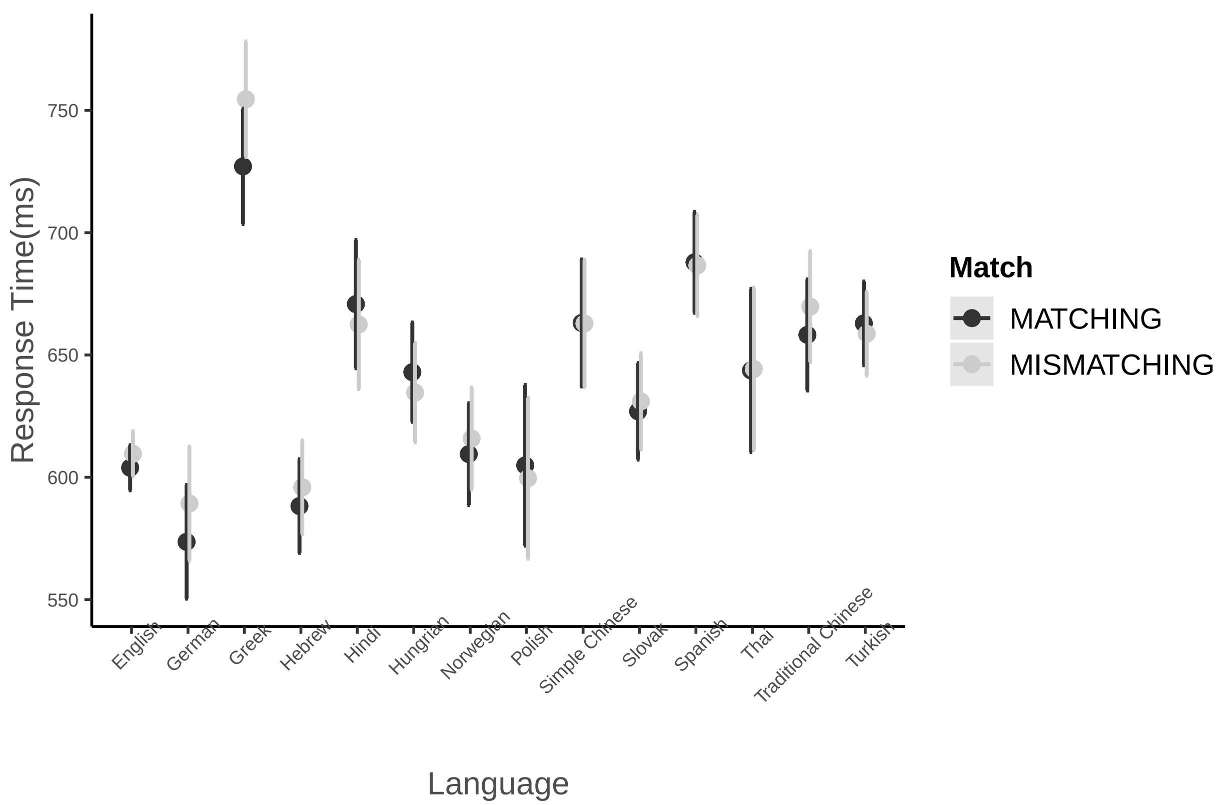

Figure 3. Response times and standard error in the sentence-picture verification task by match condition in each language (on-site data only).

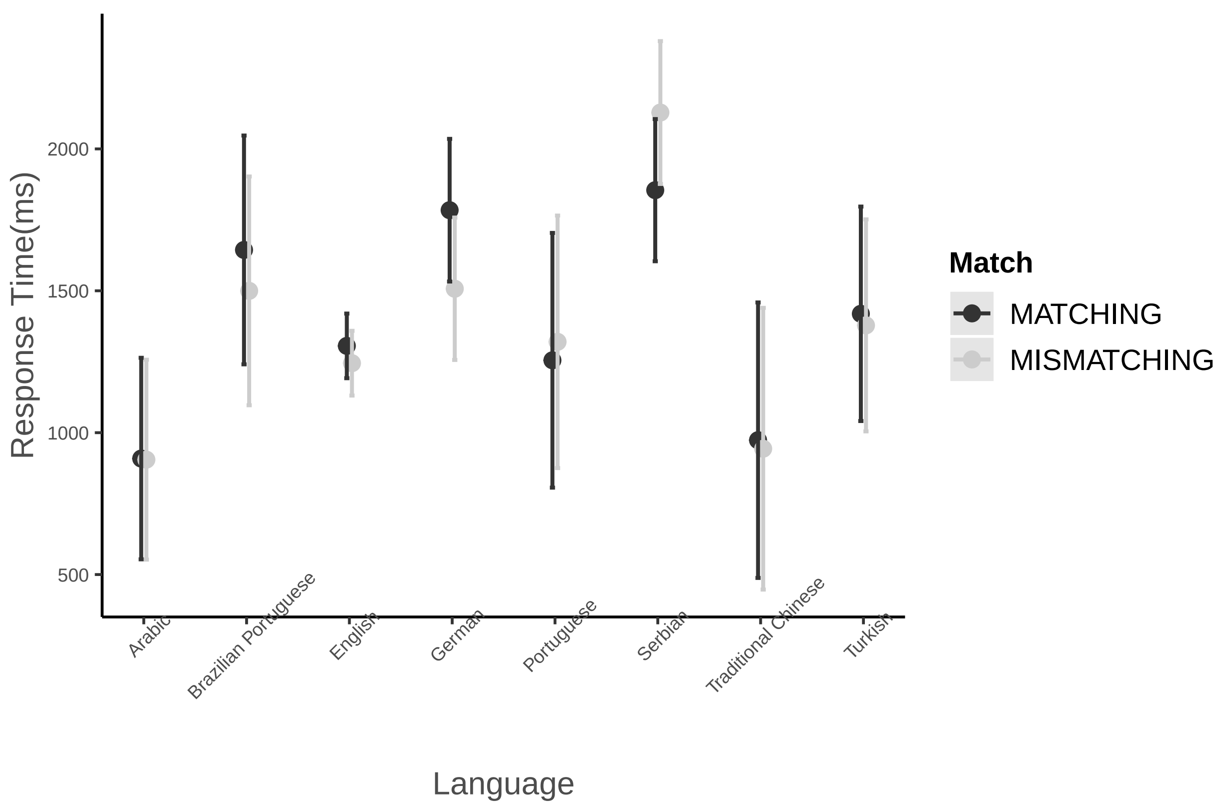

Figure 4. Response times and standard error in the sentence-picture verification task by match condition in each language (web-based data only).

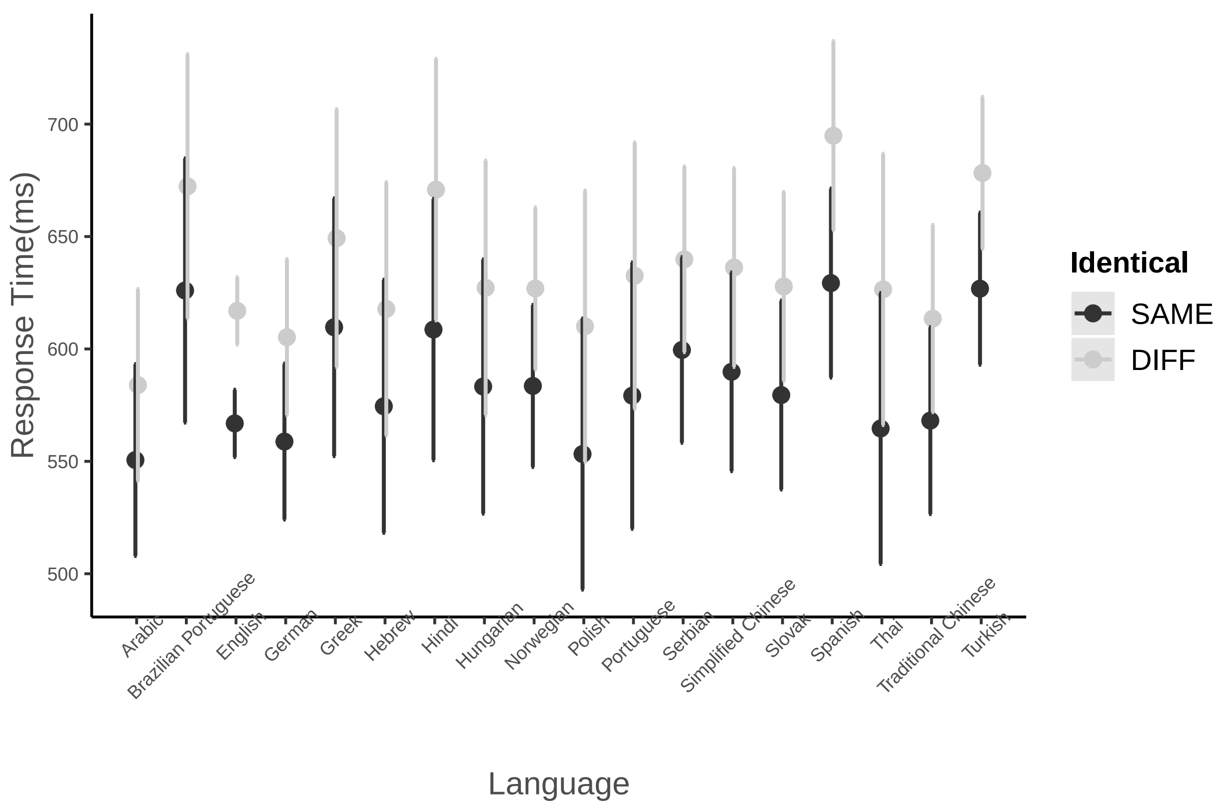

Figure 5. Response times and standard error in the picture-picture verification task by match condition in each language (both on-site and web-based data).

Appendix 1: Summary of participant characteristics

Table S 1. Characteristics of on-site participants.

Lab ID	N(Number of Females)	Average age (years)
AUT_002	102(77)	21
CAN_020	93(52)	20
CHN_005	51(29)	19
CHN_019	31(13)	25
COL_001	68(39)	21
ECU_001	63(31)	22
GBR_005	52(39)	20
GBR_006	50(37)	20
GBR_014	51(42)	19
GBR_043	29(15)	26
GRC_002	100(8)	33
HUN_001	110(2)	21
IND_003	78(51)	21
ISR_001	148(81)	23
MYS_003	47(37)	22
MYS_004	88(59)	20
NOR_003	49(1)	24
NOR_004	77(0)	22
POL_001	51(1)	24
PSA_001	55(45)	19
SVK_001	97(1)	21
SVK_002	60(0)	22
THA_001	50(39)	22
TUR_007	88(0)	21
TUR_007E	9(6)	20
TUR_025	99(0)	22
TWN_001	61(48)	21

Lab ID	N(Number of Females)	Average age (years)
TWN_002	30(21)	21
USA_020	305(235)	19
USA_032	51(30)	19
USA_033	45(34)	20
USA_065	49(31)	19
USA_173	3(0)	20

Table S 2. Characteristics of online participants.

Lab ID	N(Number of Females)	Average age (years)	Proficiency accuracy(%)
ARE_001	52(0)		59
ARE_002	54(42)	27	79
AUS_091	160(127)	26	86
AUT_005	108(80)	22	88
BRA_003	50(36)	31	92
DEU_020	26(18)	24	87
NGA_001	52(24)	24	90
NOR_002	21(12)	30	87
NZL_005	320(244)	23	91
POR_001	55(26)	31	87
PSA_002	64(49)	20	89
SRB_002	82(67)	21	94
SRB_002B	48(41)	22	86
TUR_023	79(36)	23	83
TWN_002	57(19)	21	88
TWN_002E	12(6)	21	89
USA_011	63(30)	22	88
USA_020	26(16)	20	85
USA_030	27(20)	21	87
USA_033	47(33)	20	89

Note. No participants from ARE_001 reported their birth year.

Table S 3. Number of participants completed the experimental tasks.

		on sites		online	
Language	Lab ID	SP	PP	SP	PP
Arabic	ARE_001			16	52
Arabic	ARE_002			52	54
Brazilian Portuguese	BRA_003			49	50
English	AUS_091			159	160
English	NGA_001			48	52
English	NZL_005			311	320
English	PSA_002			64	64
English	TWN_002E			12	12
English	USA_011			60	63
English	USA_020	23	23	25	26
English	USA_030			26	27
English	USA_033	31	31	45	46
German	AUT_005			108	108
German	DEU_020			26	26
Norwegian	NOR_002			21	21
Portuguese	POR_001	5	5	54	55
Serbian	SRB_002			81	82
Serbian	SRB_002B			47	48
Traditional Chinese	TWN_002	33	33	57	57
Turkish	TUR_023			76	79
English	CAN_020	93	97		
English	GBR_005	50	52		
English	GBR_006	25	25		
English	GBR_014	50	50		
English	GBR_043	25	25		
English	MYS_003	44	47		
English	MYS_004	99	100		
English	PSA_001	40	41		
English	TUR_007E	5	8		
English	USA_032	49	50		
English	USA_065	32	32		

SP: Sentence-picture verifiaction task; PP: Picture-picture verifiaction task

		on sites		online	
Language	Lab ID	SP	PP	SP	PP
English	USA_173	29	31		
German	AUT_002	99	99		
Greek	GRC_002	97	98		
Hebrew	ISR_001	145	145		
Hindi	IND_003	78	78		
Hungarian	HUN_001	129	129		
Norwegian	NOR_003	51	51		
Norwegian	NOR_004	72	70		
Polish	POL_001	50	50		
Simplified Chinese	CHN_005	50	50		
Simplified Chinese	CHN_019	30	30		
Slovak	SVK_001	90	91		
Slovak	SVK_002	47	47		
Spanish	COL_001	67	69		
Spanish	ECU_001	54	58		
Thai	THA_001	48	50		
Traditional Chinese	TWN_001	59	60		
Turkish	TUR_007	83	84		
Turkish	TUR_025	97	98		

SP: Sentence-picture verifiaction task; PP: Picture-picture verifiaction task

Table S 4. Number and proportion of outliers by laboratory.

Lab ID	Numbers of excluded participants	Proportion of excluded participants
ARE_001	13	0.25
ARE_002	14	0.26
AUS_091	0	0.00
AUT_002	25	0.25
AUT_005	0	0.00
BRA_003	0	0.00
CAN_020	24	0.25
CHN_005	13	0.26
CHN_019	8	0.26
COL_001	17	0.25
DEU_020	7	0.27
ECU_001	15	0.26
GBR_005	13	0.25
GBR_006	6	0.24
GBR_014	13	0.26
GBR_043	6	0.24
GRC_002	25	0.26
HUN_001	32	0.25
IND_003	20	0.25
ISR_001	37	0.25
MYS_003	12	0.26
MYS_004	25	0.25
NGA_001	13	0.25
NOR_002	5	0.24
NOR_003	13	0.25
NOR_004	18	0.25
NZL_005	80	0.25
POL_001	13	0.26
POR_001	15	0.25
PSA_001	10	0.24
PSA_002	0	0.00
SRB_002	0	0.00
SRB_002B	0	0.00

Lab ID	Numbers of excluded participants	Proportion of excluded participants
SVK_001	23	0.25
SVK_002	12	0.26
THA_001	13	0.26
TUR_007	21	0.25
TUR_007E	2	0.25
TUR_023	20	0.25
TUR_025	25	0.25
TWN_001	15	0.25
TWN_002	23	0.26
TWN_002E	0	0.00
USA_011	16	0.25
USA_020	12	0.24
USA_030	0	0.00
USA_032	13	0.26
USA_033	19	0.25
USA_065	8	0.25
USA_173	8	0.25

Appendix 2: Public records of intra-lab analysis

Raw data and Logs

The link to access the public site: https://scgeeker.github.io/PSA002_log_site/index.html

If you want to check the sequential analysis result of a team, at first you have to identify the ID and language of this team from "Overview" page. Next you will navigate to the language page under the banner "Tracking Logs". For example, you want to see the result of "GBR_005". Navigate "Tracking Logs -> English". Search the figure by ID "GBR_005".

The source files of the public site are available in the github repository: https://github.com/SCgeeker/PSA002_log_site

All the raw data and log files are compressed in the project OSF repository. Direct access link: https://osf.io/rg8a3/

The R code to conduct the Bayesian sequential analysis is available at "data_seq_analysis.R". Direct access link: https://github.com/SCgeeker/PSA002_log_site/blob/master/data_seq_analysis.R

Note 1 AUS_002 was unavailable because of the incorrect practices. Their raw data are accessible at the OSF (https://osf.io/j3qba). USA_067, BRA_004 and POL_004 were unavailable because the teams withdrew.

Note 2 Some mistakes happened between the collaborators' communications and required advanced data wrangling. For example, some AUS_091 participants were assigned to NZL_005. The Rmd file in NZL_005 folder were used to identify the AUS_091 participants' data then move them to AUS_091 folder.

Tidy data

All the planned analysis were conducted in terms of the filtered raw data. We split the raw data into three parts: Sentence-picture verification responses (https://osf.io/msd97/); Responses of memory trials (https://osf.io/4duce/); Picture-picture verification responses (https://osf.io/qfrhu/).

Appendix 3: Meta-analysis of match advantage by language

We conducted the respective meta-analysis for the language datasets of more than two laboratories collected at least 25 available participants' data. In addition to the Traditional Chinese, we illustrated the other 5 languages as below: English, German, Norway, Slovak and Turkey.

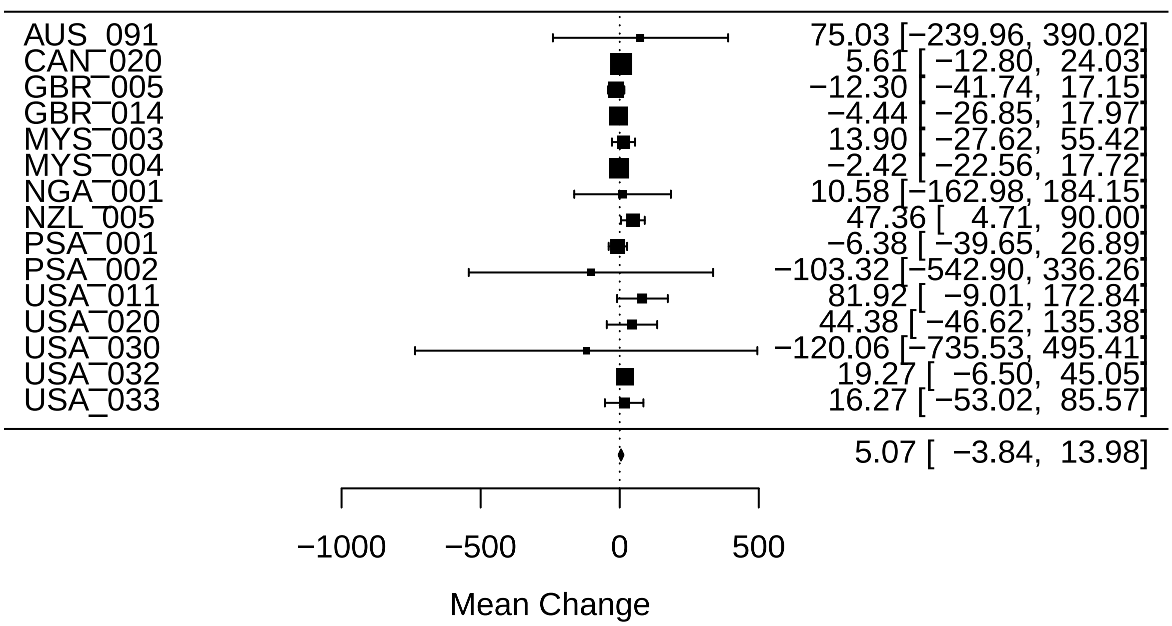

Figure S 1. Meta-analysis on match advantage of object orienation for English datasets.

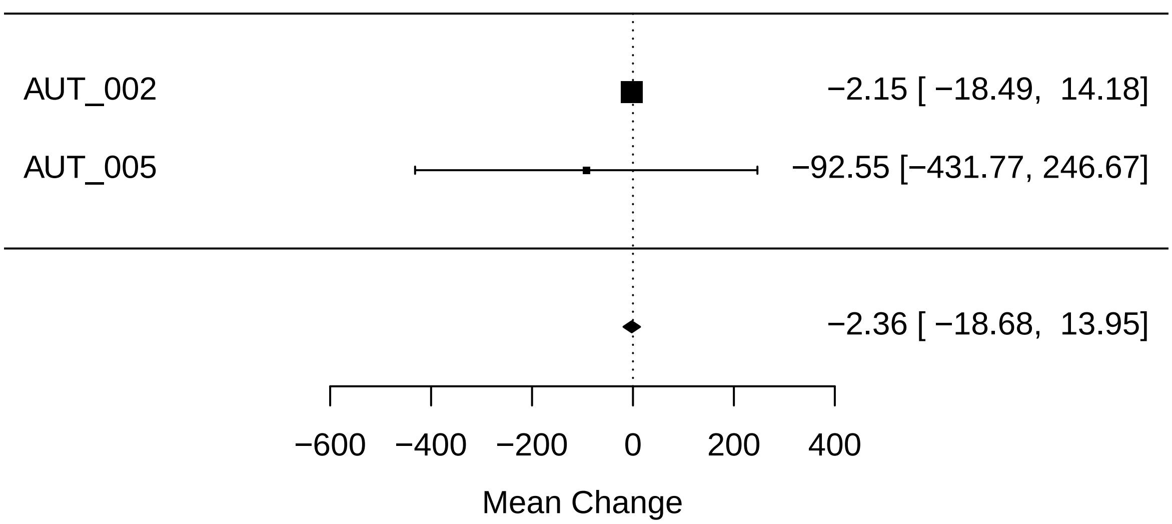

Figure S 2. Meta-analysis on match advantage of object orienation for German datasets.

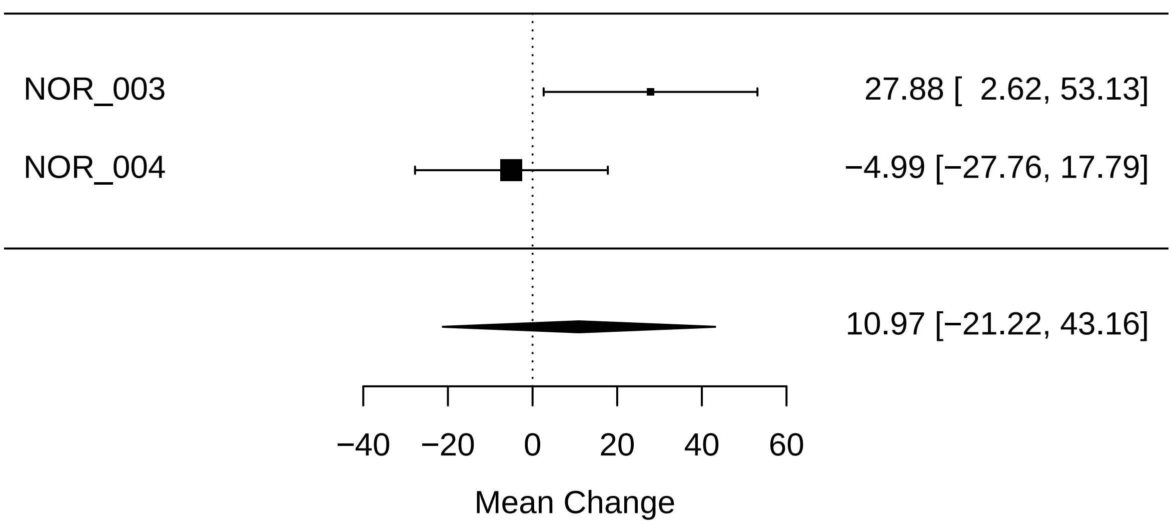

Figure S 3. Meta-analysis on match advantage of object orienation for Norwegian datasets.

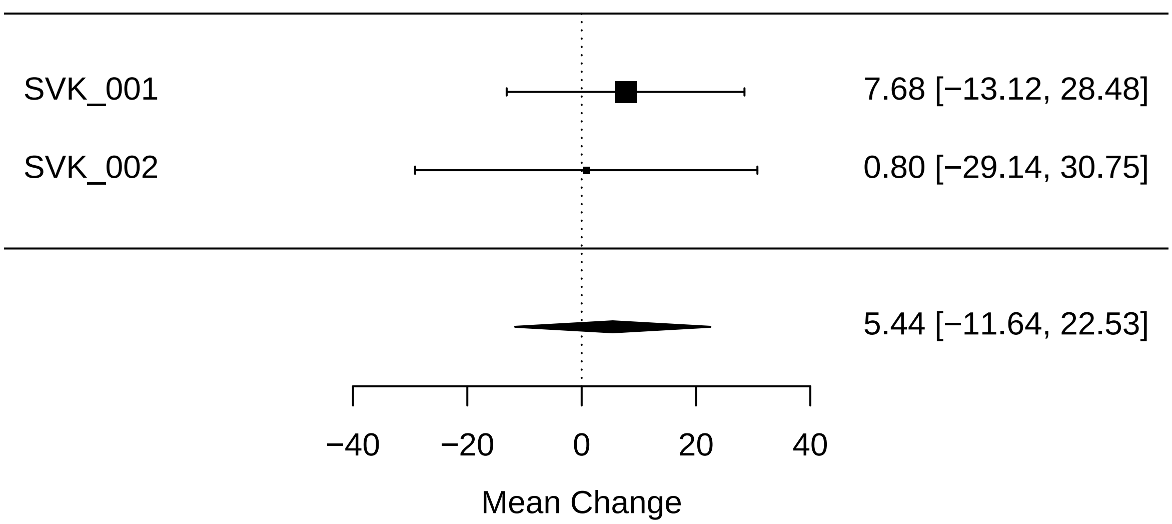

Figure S 4. Meta-analysis on match advantage of object orienation for Slovak datasets.

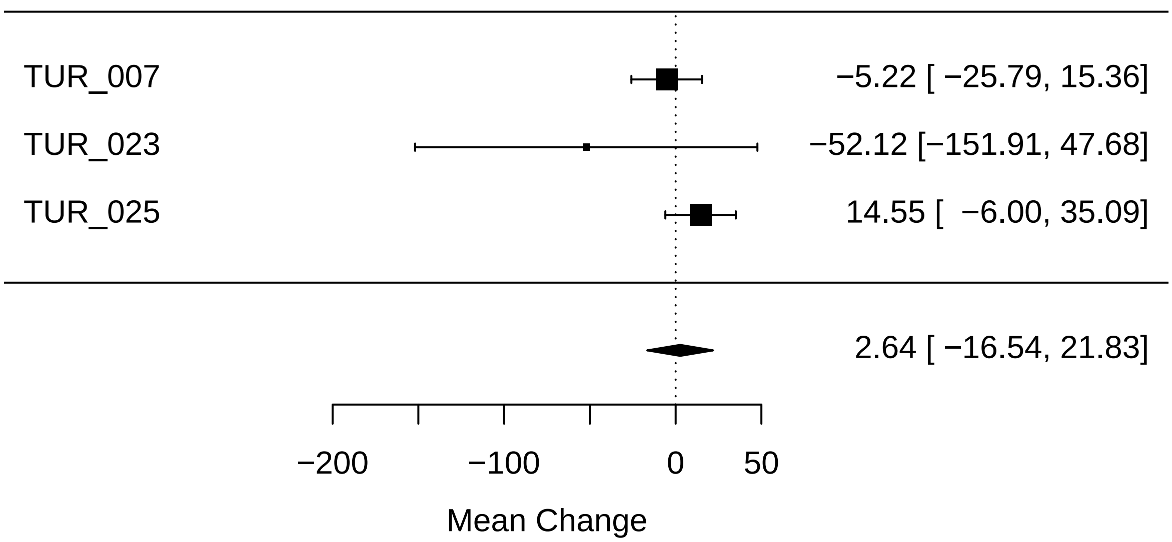

Figure S 5. Meta-analysis on match advantage of object orienation for Turkish datasets.

Appendix 4: Mixed-effect models of match advantages

Planned analysis: Matching was the only one independent variable

1. The model with four random effects: participants, targets, laboratories, and languages.

```
SP_all_random.lmer = lmer(response_time ~ Match + (1|Subject) + (1|Target) + (1|PSA_ID) + (1|Language), data = SP_V_lme_data) ## build mixed-effec
```

```
tab_model(SP_all_random.lmer, title = "Coefficients")
```

Coefficients

	response_time		
Predictors	Estimates	CI	p
(Intercept)	907.20	742.26 – 1072.13	**<0.001**
Match [MISMATCHING]	-14.42	-65.11 – 36.27	0.577
Random Effects			
σ^2	10100135.94		
τ_{00} Subject	0.00		
τ_{00} PSA_ID	122977.83		
τ_{00} Target	0.00		
τ_{00} Language	50155.84		
N Subject	2692		
N Target	48		
N PSA_ID	50		
N Language	18		
Observations	60404		
Marginal R^2 / Conditional R^2	0.000 / NA		

2. The model excluded the targets and languages from the random effect structure. The final report decided `SP_reduced_random.lmer` the best fitted model.

```
SP_reduced_random.lmer = lmerTest::lmer(response_time ~ Match + (1|Subject) + (1|PSA_ID), data = SP_V_lme_data)
SP_slope_nocor_reduced_random.lmer = lmer(response_time ~ Match + (1|Subject) + (Match||PSA_ID), data = SP_V_lme_data)
SP_slope_cor_reduced_random.lmer = lmer(response_time ~ Match + (1|Subject) + (Match|PSA_ID), data = SP_V_lme_data)
```

```
tab_model(SP_reduced_random.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	905.23	788.92 – 1021.53	**<0.001**
Match [MISMATCHING]	-14.43	-65.12 – 36.26	0.577
Random Effects			
σ^2	10100120.51		
τ_{00} Subject	0.00		
τ_{00} PSA_ID	155340.45		
N Subject	2692		
N PSA_ID	50		
Observations	60404		
Marginal R^2 / Conditional R^2	0.000 / NA		

```
tab_model(SP_slope_nocor_reduced_random.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	904.35	786.48 – 1022.21	**<0.001**
Match [MISMATCHING]	-12.75	-63.54 – 38.05	0.623
N Subject	2692		
N PSA_ID	50		
Observations	60404		

```
tab_model(SP_slope_cor_reduced_random.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	904.35	786.52 – 1022.17	**<0.001**
Match [MISMATCHING]	-12.75	-63.55 – 38.04	0.623
Random Effects			
σ^2	10100085.72		
τ_{00} Subject	0.00		
τ_{00} PSA_ID	159848.14		
τ_{11} PSA_ID.MatchMISMATCHING	133.25		
ρ_{01} PSA_ID	-1.00		
N Subject	2692		
N PSA_ID	50		
Observations	60404		
Marginal R^2 / Conditional R^2	0.000 / NA		

Models included data source

We evaluated the interaction of match advantage and data collection sources in three models. Based on the recommended practices(Barr et al., 2013; Brauer & Curtin, 2018), the models used the optimizer `bobyqa`. The final report decided `source_cor.lmer` the best fitted model.

```
## standardized
SP_V_lme_data$r_Source = if_else(SP_V_lme_data$Source == "Site",1,0)

source_zero_slope_nocor.lmer =lmer(response_time ~ Match*r_Source + (1|Subject) + (1|PSA_ID) + (1|Language),
  control = lmerControl(optimizer = "bobyqa",
                        optCtrl = list(maxfun = 1e6)), data = SP_V_lme_data)


source_nocor.lmer =lmer(response_time ~ Match*r_Source + (1|Subject) + (r_Source||PSA_ID) + (r_Source||Language),
  control = lmerControl(optimizer = "bobyqa",
                        optCtrl = list(maxfun = 1e6)),
   data = SP_V_lme_data)

source_cor.lmer =lmer(response_time ~ Match*r_Source + (1|Subject) + (r_Source|PSA_ID) + (r_Source|Language),
  control = lmerControl(optimizer = "bobyqa",
                        optCtrl = list(maxfun = 1e6)),
                        data = SP_V_lme_data)

tab_model(source_zero_slope_nocor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	1360.41	1236.16 – 1484.67	**<0.001**
Match [MISMATCHING]	-39.57	-117.06 – 37.91	0.317
r_Source	-685.06	-793.42 – -576.71	**<0.001**
Match [MISMATCHING] * r_Source	43.94	-58.50 – 146.39	0.400
Random Effects			
σ^2	10099089.93		
τ_{00} Subject	0.00		
τ_{00} PSA_ID	13643.18		
τ_{00} Language	31903.35		
N Subject	2692		
N PSA_ID	50		
N Language	18		
Observations	60404		
Marginal R^2 / Conditional R^2	0.011 / NA		

```
tab_model(source_nocor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	1360.41	1236.16 – 1484.67	**<0.001**
Match [MISMATCHING]	-39.57	-117.06 – 37.91	0.317
r_Source	-685.06	-793.42 – -576.71	**<0.001**
Match [MISMATCHING] * r_Source	43.94	-58.50 – 146.39	0.400
Random Effects			
σ^2	10099089.92		
τ_{00} Subject	0.00		
τ_{00} PSA_ID	13643.25		
τ_{00} Language	31903.23		
τ_{11} PSA_ID.r_Source	0.00		
τ_{11} Language.r_Source	0.00		
ρ_{01}			
ρ_{01}			
N Subject	2692		
N PSA_ID	50		
N Language	18		
Observations	60404		
Marginal R^2 / Conditional R^2	0.011 / NA		

```
tab_model(source_cor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	1374.45	1145.16 – 1603.75	**<0.001**
Match [MISMATCHING]	-39.70	-117.17 – 37.77	0.315
r_Source	-744.57	-978.73 – -510.42	**<0.001**
Match [MISMATCHING] * r_Source	44.01	-58.41 – 146.43	0.400
Random Effects			
σ^2	10094277.36		
τ_{00} Subject	0.00		
τ_{00} PSA_ID	46683.47		
τ_{00} Language	79313.45		
τ_{11} PSA_ID.r_Source	47754.95		
τ_{11} Language.r_Source	79299.01		
ρ_{01} PSA_ID	-1.00		
ρ_{01} Language	-1.00		

N $_{Subject}$	2692
N $_{PSA\_ID}$	50
N $_{Language}$	18
Observations	60404
Marginal R^2 / Conditional R^2	0.013 / NA

Models included languages

We analyzed the interactions by the data sources separately.

on site data

We evaluated the interaction of match advantage and languagess in three models. Based on the recommended practices(Barr et al., 2013; Brauer & Curtin, 2018), the models used the optimizer `bobyqa`. The final report decided `lang_cor.lmer` the best fitted model.

```
## Check sample size of a language by site data
site_excluded_lang <- subset(SP_V_lme_data, Source=="Site") %>%
  group_by(Language, Subject) %>%
  summarise(N_trials = n()) %>%
  group_by(Language) %>%
  summarise(N = n()) %>%
  filter(N < 25) %>%
  pull(Language)

## Allocate the site data
SP_V_lang_lme_data = subset(SP_V_lme_data, Source=="Site" & !(Language %in% site_excluded_lang))

## Run the mixed effect model by site data
lang_cor.lmer =lmer(response_time ~ Language*Match + (1|Subject),
  control = lmerControl(optimizer = "bobyqa",
                        optCtrl = list(maxfun = 1e6)),
  data = SP_V_lang_lme_data)

lang_slope_nocor.lmer =lmer(response_time ~ Language*Match + (Match||Subject),
  control = lmerControl(optimizer = "bobyqa",
                        optCtrl = list(maxfun = 1e6)), data = SP_V_lang_lme_data)

lang_slope_cor.lmer =lmer(response_time ~ Language*Match + (Match|Subject),
  control = lmerControl(optimizer = "bobyqa",
                        optCtrl = list(maxfun = 1e6)), data = SP_V_lang_lme_data)
```

```
tab_model(lang_cor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	603.90	594.69 – 613.11	**<0.001**
Language [German]	-30.29	-55.28 – -5.30	**0.018**
Language [Greek]	123.21	97.90 – 148.51	**<0.001**
Language [Hebrew]	-15.66	-36.91 – 5.58	0.148
Language [Hindi]	66.91	39.13 – 94.69	**<0.001**
Language [Hungarian]	39.07	16.78 – 61.35	**0.001**
Language [Norwegian]	5.58	-17.18 – 28.33	0.631
Language [Polish]	0.98	-33.13 – 35.10	0.955
Language [Simplified Chinese]	59.20	31.69 – 86.70	**<0.001**
Language [Slovak]	23.06	1.29 – 44.82	**0.038**
Language [Spanish]	83.98	61.39 – 106.56	**<0.001**
Language [Thai]	39.78	5.19 – 74.38	**0.024**
Language [Traditional Chinese]	54.33	29.81 – 78.84	**<0.001**
Language [Turkish]	59.02	39.60 – 78.44	**<0.001**
Match [MISMATCHING]	5.73	-1.32 – 12.77	0.111
Language [German] * Match [MISMATCHING]	9.93	-9.10 – 28.96	0.306
Language [Greek] * Match [MISMATCHING]	21.72	2.12 – 41.31	**0.030**
Language [Hebrew] * Match [MISMATCHING]	1.95	-14.17 – 18.08	0.812
Language [Hindi] * Match [MISMATCHING]	-14.07	-35.75 – 7.61	0.203
Language [Hungarian] * Match [MISMATCHING]	-14.08	-31.05 – 2.90	0.104
Language [Norwegian] * Match [MISMATCHING]	0.66	-16.63 – 17.94	0.941
Language [Polish] * Match [MISMATCHING]	-10.98	-37.00 – 15.03	0.408
Language [Simplified Chinese] * Match [MISMATCHING]	-5.93	-27.13 – 15.28	0.584
Language [Slovak] * Match [MISMATCHING]	-1.65	-18.20 – 14.89	0.845
Language [Spanish] * Match [MISMATCHING]	-6.98	-24.34 – 10.38	0.430
Language [Thai] * Match [MISMATCHING]	-5.11	-31.96 – 21.73	0.709
Language [Traditional Chinese] * Match [MISMATCHING]	5.80	-12.96 – 24.56	0.544

Language [Turkish] * Match [MISMATCHING]	-9.95	-24.76 – 4.85	0.188
Random Effects			
σ^2	34460.57		
τ_{00} Subject	7389.48		
ICC	0.18		
N Subject	1524		
Observations	34441		
Marginal R^2 / Conditional R^2	0.034 / 0.204		

```
tab_model(lang_slope_nocor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	603.89	594.75 – 613.04	**<0.001**
Language [German]	-30.28	-55.09 – -5.47	**0.017**
Language [Greek]	123.20	98.09 – 148.32	**<0.001**
Language [Hebrew]	-15.65	-36.74 – 5.43	0.146
Language [Hindi]	66.89	39.31 – 94.46	**<0.001**
Language [Hungarian]	39.08	16.95 – 61.20	**0.001**
Language [Norwegian]	5.59	-17.00 – 28.18	0.628
Language [Polish]	0.98	-32.88 – 34.85	0.955
Language [Simplified Chinese]	59.21	31.90 – 86.51	**<0.001**
Language [Slovak]	23.06	1.46 – 44.66	**0.036**
Language [Spanish]	83.95	61.53 – 106.37	**<0.001**
Language [Thai]	39.76	5.42 – 74.10	**0.023**
Language [Traditional Chinese]	54.33	30.00 – 78.67	**<0.001**
Language [Turkish]	59.02	39.74 – 78.29	**<0.001**
Match [MISMATCHING]	5.73	-1.36 – 12.83	0.113
Language [German] * Match [MISMATCHING]	9.91	-9.24 – 29.07	0.310
Language [Greek] * Match [MISMATCHING]	21.75	2.03 – 41.47	**0.031**
Language [Hebrew] * Match [MISMATCHING]	1.94	-14.29 – 18.17	0.815
Language [Hindi] * Match [MISMATCHING]	-14.08	-35.90 – 7.74	0.206
Language [Hungarian] * Match [MISMATCHING]	-14.09	-31.17 – 3.00	0.106
Language [Norwegian] * Match [MISMATCHING]	0.65	-16.75 – 18.05	0.942
Language [Polish] * Match [MISMATCHING]	-10.98	-37.17 – 15.21	0.411
Language [Simplified Chinese] * Match [MISMATCHING]	-5.92	-27.26 – 15.42	0.587
Language [Slovak] * Match [MISMATCHING]	-1.64	-18.30 – 15.01	0.847
Language [Spanish] * Match [MISMATCHING]	-6.96	-24.43 – 10.51	0.435
Language [Thai] * Match [MISMATCHING]	-5.08	-32.10 – 21.93	0.712
Language [Traditional Chinese] * Match [MISMATCHING]	5.81	-13.08 – 24.69	0.547
Language [Turkish] * Match [MISMATCHING]	-9.94	-24.85 – 4.97	0.191
N Subject	1524		
Observations	34441		

```
tab_model(lang_slope_cor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	603.89	594.75 – 613.04	**<0.001**
Language [German]	-30.28	-55.09 – -5.47	**0.017**
Language [Greek]	123.20	98.09 – 148.32	**<0.001**
Language [Hebrew]	-15.65	-36.74 – 5.43	0.146
Language [Hindi]	66.89	39.31 – 94.46	**<0.001**
Language [Hungarian]	39.08	16.95 – 61.20	**0.001**
Language [Norwegian]	5.59	-17.00 – 28.18	0.628
Language [Polish]	0.98	-32.88 – 34.85	0.955
Language [Simplified Chinese]	59.21	31.90 – 86.51	**<0.001**
Language [Slovak]	23.06	1.46 – 44.66	**0.036**
Language [Spanish]	83.95	61.53 – 106.37	**<0.001**
Language [Thai]	39.76	5.42 – 74.10	**0.023**
Language [Traditional Chinese]	54.33	30.00 – 78.67	**<0.001**

Language [Turkish]	59.02	39.74 – 78.29	**<0.001**
Match [MISMATCHING]	5.73	-1.36 – 12.83	0.113
Language [German] * Match [MISMATCHING]	9.91	-9.24 – 29.07	0.310
Language [Greek] * Match [MISMATCHING]	21.75	2.03 – 41.47	**0.031**
Language [Hebrew] * Match [MISMATCHING]	1.94	-14.29 – 18.17	0.815
Language [Hindi] * Match [MISMATCHING]	-14.08	-35.90 – 7.74	0.206
Language [Hungarian] * Match [MISMATCHING]	-14.09	-31.17 – 3.00	0.106
Language [Norwegian] * Match [MISMATCHING]	0.65	-16.75 – 18.05	0.942
Language [Polish] * Match [MISMATCHING]	-10.98	-37.17 – 15.21	0.411
Language [Simplified Chinese] * Match [MISMATCHING]	-5.92	-27.26 – 15.42	0.587
Language [Slovak] * Match [MISMATCHING]	-1.64	-18.30 – 15.01	0.847
Language [Spanish] * Match [MISMATCHING]	-6.96	-24.43 – 10.51	0.435
Language [Thai] * Match [MISMATCHING]	-5.08	-32.10 – 21.93	0.712
Language [Traditional Chinese] * Match [MISMATCHING]	5.81	-13.08 – 24.69	0.547
Language [Turkish] * Match [MISMATCHING]	-9.94	-24.85 – 4.97	0.191
Random Effects			
σ^2	34438.84		
τ_{00} Subject	7238.67		
τ_{11} Subject.MatchMISMATCHING	83.90		
ρ_{01} Subject	0.17		
ICC	0.18		
N Subject	1524		
Observations	34441		
Marginal R^2 / Conditional R^2	0.034 / 0.205		

web-based data

We evaluated the interaction of match advantage and languagess in three models. Based on the recommended practices(Barr et al., 2013; Brauer & Curtin, 2018), the models used the optimizer `bobyqa`. The final report decided `osweb_cor.lmer` the best fitted model.

```
## Check sample size of a language by osweb data
osweb_excluded_lang <- subset(SP_V_lme_data, Source=="Internet") %>%
  group_by(Language, Subject) %>%
  summarise(N_trials = n()) %>%
  group_by(Language) %>%
  summarise(N = n()) %>%
  filter(N < 25) %>%
  pull(Language)  ## Exclude the languages less than 25 participants

## Allocate the osweb data
SP_V_osweb_lme_data = subset(SP_V_lme_data, Source=="Internet" & !(Language %in% osweb_excluded_lang))


## Run the mixed effect model by site data
osweb_cor.lmer = lmerTest::lmer(response_time ~ Language*Match + (1|Subject),
  control = lmerControl(optimizer = "bobyqa",
                      optCtrl = list(maxfun = 1e6)), # Increase maximum number of iterations to facilitate model convergence ,
        data = SP_V_osweb_lme_data)

osweb_slope_nocor.lmer =lmer(response_time ~ Language*Match + (Match||Subject),
  control = lmerControl(optimizer = "bobyqa",
        optCtrl = list(maxfun = 1e6)), data = SP_V_osweb_lme_data)

osweb_slope_cor.lmer =lmer(response_time ~ Language*Match + (Match|Subject),
  control = lmerControl(optimizer = "bobyqa",
        optCtrl = list(maxfun = 1e6)), data = SP_V_osweb_lme_data)


tab_model(osweb_cor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	908.78	553.58 – 1263.99	**<0.001**
Language [Brazilian Portuguese]	735.18	197.77 – 1272.59	**0.007**
Language [English]	397.20	24.14 – 770.25	**0.037**
Language [German]	875.24	440.05 – 1310.43	**<0.001**
Language [Portuguese]	346.34	-226.17 – 918.85	0.236
Language [Serbian]	945.80	511.11 – 1380.49	**<0.001**
Language [Traditional Chinese]	64.78	-536.87 – 666.42	0.833
Language [Turkish]	510.13	-8.57 – 1028.83	0.054
Match [MISMATCHING]	-3.92	-504.22 – 496.37	0.988

Language [Brazilian Portuguese] * Match [MISMATCHING]	-140.28	-898.94 – 618.38	0.717
Language [English] * Match [MISMATCHING]	-57.33	-583.00 – 468.34	0.831
Language [German] * Match [MISMATCHING]	-272.91	-886.48 – 340.66	0.383
Language [Portuguese] * Match [MISMATCHING]	68.91	-737.32 – 875.14	0.867
Language [Serbian] * Match [MISMATCHING]	277.47	-336.10 – 891.04	0.375
Language [Traditional Chinese] * Match [MISMATCHING]	-26.12	-882.09 – 829.84	0.952
Language [Turkish] * Match [MISMATCHING]	-36.82	-766.74 – 693.10	0.921
Random Effects			
σ^2	23878104.68		
$\tau_{00\ Subject}$	0.00		
$N_{Subject}$	1147		
Observations	25480		
Marginal R^2 / Conditional R^2	0.003 / NA		

```
tab_model(osweb_slope_nocor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	908.89	553.03 – 1264.76	**<0.001**
Language [Brazilian Portuguese]	735.08	196.58 – 1273.58	**0.007**
Language [English]	397.07	23.31 – 770.83	**0.037**
Language [German]	875.10	439.06 – 1311.15	**<0.001**
Language [Portuguese]	346.24	-227.45 – 919.93	0.237
Language [Serbian]	945.91	510.36 – 1381.45	**<0.001**
Language [Traditional Chinese]	64.65	-538.23 – 667.54	0.834
Language [Turkish]	510.01	-9.72 – 1029.74	0.054
Match [MISMATCHING]	-4.15	-506.95 – 498.65	0.987
Language [Brazilian Portuguese] * Match [MISMATCHING]	-140.02	-902.82 – 622.77	0.719
Language [English] * Match [MISMATCHING]	-57.10	-585.44 – 471.24	0.832
Language [German] * Match [MISMATCHING]	-272.66	-889.48 – 344.15	0.386
Language [Portuguese] * Match [MISMATCHING]	69.14	-741.55 – 879.84	0.867
Language [Serbian] * Match [MISMATCHING]	277.40	-339.40 – 894.20	0.378
Language [Traditional Chinese] * Match [MISMATCHING]	-25.88	-886.53 – 834.77	0.953
Language [Turkish] * Match [MISMATCHING]	-36.57	-770.40 – 697.26	0.922
$N_{Subject}$	1147		
Observations	25480		

```
tab_model(osweb_slope_cor.lmer)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	908.89	553.03 – 1264.76	**<0.001**
Language [Brazilian Portuguese]	735.08	196.58 – 1273.58	**0.007**
Language [English]	397.07	23.31 – 770.83	**0.037**
Language [German]	875.10	439.06 – 1311.15	**<0.001**
Language [Portuguese]	346.24	-227.45 – 919.93	0.237
Language [Serbian]	945.91	510.36 – 1381.45	**<0.001**
Language [Traditional Chinese]	64.65	-538.23 – 667.54	0.834
Language [Turkish]	510.01	-9.72 – 1029.74	0.054
Match [MISMATCHING]	-4.15	-506.95 – 498.65	0.987
Language [Brazilian Portuguese] * Match [MISMATCHING]	-140.02	-902.82 – 622.77	0.719
Language [English] * Match [MISMATCHING]	-57.10	-585.44 – 471.24	0.832
Language [German] * Match [MISMATCHING]	-272.66	-889.48 – 344.15	0.386
Language [Portuguese] * Match [MISMATCHING]	69.14	-741.56 – 879.84	0.867
Language [Serbian] * Match [MISMATCHING]	277.40	-339.40 – 894.20	0.378

Language [Traditional Chinese] * Match [MISMATCHING]	-25.88	-886.53 – 834.77	0.953
Language [Turkish] * Match [MISMATCHING]	-36.57	-770.40 – 697.26	0.922
Random Effects			
σ^2	23865260.44		
τ_{00} Subject	10220.10		
τ_{11} Subject.MatchMISMATCHING	51127.34		
ρ_{01} Subject	-1.00		
N Subject	1147		
Observations	25480		
Marginal R^2 / Conditional R^2	0.003 / NA		

Appendix 5: Planned analysis for imagery scores

```
# Load PP verification responses
PP <- dir(path = "..",
      pattern = "all_rawdata_PP",
      recursive = TRUE, full.names = TRUE) %>%
      read_csv() %>%
      subset(correct == 1 & Identical != "F")  %>%  ## Exclude the incorrect responses and filler trials
      inner_join(select(lab_info, PSA_ID, Language), by = "PSA_ID") %>%
    distinct() %>% ## Merge the language aspects
    mutate(Source = if_else(opensesame_codename == "osweb","osweb","site"),
           Subject = paste0(Source,"_",PSA_ID,"_",subject_nr)) ## Compose the unique participant id

## Tidy PP data for mixed linear model
PP_site_tidy <- PP %>%
    filter(Source!="osweb")

## Tidy PP data for mixed linear model
PP_osweb_tidy <-  PP %>%
      filter(Source=="osweb") %>%   # include jatos data
      subset(correct == 1 & Identical != "F")  %>%  ## Exclude the incorrect responses and filler trials
    distinct() %>% ## Merge the language aspects
    filter(!(PSA_ID == "USA_033" & subject_nr == 39)) ## exclude this participant who had not complete PP

PP_tidy = bind_rows(PP_site_tidy, PP_osweb_tidy)

## Dataset for mixed-effect model
PP_lme_data <- PP_tidy %>% left_join(outliers_table, by=c("PSA_ID" = "LAB", "Subject" = "Subject")) %>%  ## filter the outliers by SP_V data
  filter(Outlier == FALSE)

PP_lme_data$Language <- ifelse(PP_lme_data$Language == "Magyar","Hungarian",PP_lme_data$Language)
PP_lme_data$Language <- ifelse(PP_lme_data$Language ==  "Simple Chinese", "Simplified Chinese",PP_lme_data$Language)

PP_lme_data <- mutate(PP_lme_data,
                        Identical= factor(Identical,
                          levels = c("Y","N"),
                          labels = c("SAME","DIFF")))
```

Planned analysis: Interaction of imagery scores and langauges

We evaluated the interaction of imagery scores and languages in three models. Based on the recommended practices(Barr et al., 2013; Brauer & Curtin, 2018), the models used the optimizer `bobyqa`. The final report decided `PP.lang.zero_slope.cor.lme` the best fitted model.

```
## standardized the imagery score
PP_lme_data$Language <- if_else(PP_lme_data$Language=="English","0English",PP_lme_data$Language)

PP.lang.zero_slope.cor.lme <- lmerTest::lmer(response_time ~
                    Identical*Language +                # Fixed effect
                    (1 | Subject) +   # By-subject random intercept
                    (1 | Picture1) +   # By-item random intercept
                    (1 | PSA_ID),    # By-lab random intercept
#                    (z_Identical || PSA_ID),    # By-lab random slopes
                    data = PP_lme_data,
                    #method = 'KR',               # Calculate p values using Kenward-Roger method
                    control = lmerControl(optimizer = "bobyqa",optCtrl = list(maxfun = 1e6)) # Increase maximum number of iterations to facilitate
                    )


PP.lang.slopes.nocor.lme <- lmerTest::lmer(response_time ~
                    Identical*Language +                # Fixed effect
                    (1 | Subject) +   # By-subject random intercept
                    (1 | Picture1) +   # By-item random intercept
                    (Identical || PSA_ID),    # By-lab random intercept
                    data = PP_lme_data,
                    #method = 'KR',               # Calculate p values using Kenward-Roger method
                    control = lmerControl(optimizer = "bobyqa",optCtrl = list(maxfun = 1e6)) # Increase maximum number of iterations to facilitate
                    )

PP.lang.slopes.cor.lme <- lmerTest::lmer(response_time ~
                    Identical*Language +                # Fixed effect
                    (1 | Subject) +   # By-subject random intercept
                    (1 | Picture1) +   # By-item random intercept
                    (Identical | PSA_ID),    # By-lab random intercept
                    data = PP_lme_data,
                    #method = 'KR',               # Calculate p values using Kenward-Roger method
                    control = lmerControl(optimizer = "bobyqa",optCtrl = list(maxfun = 1e6)) # Increase maximum number of iterations to facilitate
                    )

tab_model(PP.lang.zero_slope.cor.lme)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	566.93	551.97 – 581.88	**<0.001**
Identical [DIFF]	50.10	47.24 – 52.97	**<0.001**
Language [Arabic]	-16.34	-60.42 – 27.73	0.467
Language [Brazilian Portuguese]	59.07	-0.69 – 118.83	0.053
Language [German]	-8.12	-44.52 – 28.28	0.662
Language [Greek]	42.72	-15.64 – 101.08	0.151
Language [Hebrew]	7.56	-49.79 – 64.92	0.796
Language [Hindi]	41.71	-17.44 – 100.86	0.167
Language [Hungarian]	16.36	-41.24 – 73.95	0.578
Language [Norwegian]	16.62	-21.09 – 54.33	0.388
Language [Polish]	-13.67	-74.94 – 47.59	0.662
Language [Portuguese]	12.30	-47.94 – 72.53	0.689
Language [Serbian]	32.61	-10.13 – 75.36	0.135
Language [Simplified Chinese]	22.91	-22.66 – 68.48	0.324
Language [Slovak]	12.57	-30.88 – 56.01	0.571

Language [Spanish]	62.37	18.91 – 105.84	**0.005**
Language [Thai]	-2.31	-63.57 – 58.95	0.941
Language [Traditional Chinese]	1.18	-41.85 – 44.22	0.957
Language [Turkish]	59.88	24.29 – 95.47	**0.001**
Identical [DIFF] * Language [Arabic]	-16.76	-27.80 – -5.71	**0.003**
Identical [DIFF] * Language [Brazilian Portuguese]	-3.76	-17.43 – 9.91	0.590
Identical [DIFF] * Language [German]	-3.70	-10.99 – 3.59	0.319
Identical [DIFF] * Language [Greek]	-10.48	-21.93 – 0.97	0.073
Identical [DIFF] * Language [Hebrew]	-6.80	-16.30 – 2.69	0.160
Identical [DIFF] * Language [Hindi]	12.12	-0.59 – 24.83	0.062
Identical [DIFF] * Language [Hungarian]	-6.19	-16.17 – 3.79	0.224
Identical [DIFF] * Language [Norwegian]	-6.76	-16.36 – 2.83	0.167
Identical [DIFF] * Language [Polish]	6.73	-9.00 – 22.46	0.402
Identical [DIFF] * Language [Portuguese]	3.26	-11.10 – 17.61	0.656
Identical [DIFF] * Language [Serbian]	-9.79	-18.60 – -0.97	**0.030**
Identical [DIFF] * Language [Simplified Chinese]	-3.70	-16.35 – 8.94	0.566
Identical [DIFF] * Language [Slovak]	-1.78	-11.51 – 7.95	0.720
Identical [DIFF] * Language [Spanish]	15.46	5.20 – 25.72	**0.003**
Identical [DIFF] * Language [Thai]	11.77	-4.04 – 27.57	0.144
Identical [DIFF] * Language [Traditional Chinese]	-4.72	-14.10 – 4.67	0.324
Identical [DIFF] * Language [Turkish]	1.36	-5.99 – 8.71	0.717
Random Effects			
σ^2	13509.16		
$\tau_{00\ Subject}$	5662.22		
$\tau_{00\ PSA\_ID}$	747.87		
$\tau_{00\ Picture1}$	613.78		
ICC	0.34		
$N_{Subject}$	2687		
$N_{Picture1}$	48		
N_{PSA_ID}	50		
Observations	62164		
Marginal R^2 / Conditional R^2	0.052 / 0.377		

```
tab_model(PP.lang.slopes.nocor.lme)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	566.96	552.64 – 581.27	**<0.001**
Identical [DIFF]	50.10	46.76 – 53.45	**<0.001**
Language [Arabic]	-16.38	-58.20 – 25.44	0.443
Language [Brazilian Portuguese]	59.03	2.44 – 115.62	**0.041**
Language [German]	-9.13	-43.47 – 25.21	0.602
Language [Greek]	42.69	-12.42 – 97.80	0.129
Language [Hebrew]	7.53	-46.52 – 61.58	0.785
Language [Hindi]	41.68	-14.27 – 97.63	0.144
Language [Hungarian]	16.33	-37.97 – 70.63	0.556
Language [Norwegian]	16.15	-19.62 – 51.92	0.376
Language [Polish]	-13.70	-71.88 – 44.48	0.644
Language [Portuguese]	12.26	-44.83 – 69.36	0.674
Language [Serbian]	31.52	-8.87 – 71.91	0.126
Language [Simplified Chinese]	22.79	-20.55 – 66.14	0.303
Language [Slovak]	13.06	-28.05 – 54.17	0.533
Language [Spanish]	62.21	21.04 – 103.39	**0.003**
Language [Thai]	-2.34	-60.52 – 55.84	0.937
Language [Traditional Chinese]	1.03	-39.68 – 41.73	0.961
Language [Turkish]	59.71	26.08 – 93.35	**0.001**

Identical [DIFF] * Language [Arabic]	-16.75	-29.14 – -4.36	**0.008**
Identical [DIFF] * Language [Brazilian Portuguese]	-3.73	-19.45 – 11.98	0.642
Identical [DIFF] * Language [German]	-3.87	-12.61 – 4.87	0.385
Identical [DIFF] * Language [Greek]	-10.48	-24.31 – 3.35	0.137
Identical [DIFF] * Language [Hebrew]	-6.81	-19.07 – 5.45	0.276
Identical [DIFF] * Language [Hindi]	12.13	-2.76 – 27.01	0.110
Identical [DIFF] * Language [Hungarian]	-6.20	-18.85 – 6.44	0.336
Identical [DIFF] * Language [Norwegian]	-7.22	-17.95 – 3.50	0.187
Identical [DIFF] * Language [Polish]	6.73	-10.81 – 24.27	0.452
Identical [DIFF] * Language [Portuguese]	3.27	-13.05 – 19.58	0.695
Identical [DIFF] * Language [Serbian]	-8.38	-18.85 – 2.09	0.117
Identical [DIFF] * Language [Simplified Chinese]	-3.52	-17.37 – 10.33	0.619
Identical [DIFF] * Language [Slovak]	-1.95	-13.20 – 9.31	0.735
Identical [DIFF] * Language [Spanish]	15.30	3.60 – 27.00	**0.010**
Identical [DIFF] * Language [Thai]	11.77	-5.84 – 29.37	0.190
Identical [DIFF] * Language [Traditional Chinese]	-4.56	-15.50 – 6.39	0.415
Identical [DIFF] * Language [Turkish]	1.56	-7.16 – 10.29	0.726
$N_{Subject}$	2687		
$N_{Picture1}$	48		
N_{PSA_ID}	50		
Observations	62164		

```
tab_model(PP.lang.slopes.cor.lme)
```

	response_time		
Predictors	Estimates	CI	p
(Intercept)	566.96	552.64 – 581.27	**<0.001**
Identical [DIFF]	50.10	46.76 – 53.45	**<0.001**
Language [Arabic]	-16.38	-58.20 – 25.44	0.443
Language [Brazilian Portuguese]	59.03	2.44 – 115.62	**0.041**
Language [German]	-9.13	-43.47 – 25.21	0.602
Language [Greek]	42.69	-12.42 – 97.80	0.129
Language [Hebrew]	7.53	-46.52 – 61.58	0.785
Language [Hindi]	41.68	-14.27 – 97.63	0.144
Language [Hungarian]	16.33	-37.97 – 70.63	0.556
Language [Norwegian]	16.15	-19.62 – 51.92	0.376
Language [Polish]	-13.70	-71.88 – 44.48	0.644
Language [Portuguese]	12.26	-44.83 – 69.36	0.674
Language [Serbian]	31.52	-8.87 – 71.91	0.126
Language [Simplified Chinese]	22.79	-20.55 – 66.14	0.303
Language [Slovak]	13.06	-28.05 – 54.17	0.533
Language [Spanish]	62.21	21.04 – 103.39	**0.003**
Language [Thai]	-2.34	-60.52 – 55.84	0.937
Language [Traditional Chinese]	1.03	-39.68 – 41.74	0.961
Language [Turkish]	59.71	26.08 – 93.35	**0.001**
Identical [DIFF] * Language [Arabic]	-16.75	-29.14 – -4.36	**0.008**
Identical [DIFF] * Language [Brazilian Portuguese]	-3.73	-19.45 – 11.98	0.642
Identical [DIFF] * Language [German]	-3.87	-12.61 – 4.87	0.385
Identical [DIFF] * Language [Greek]	-10.48	-24.31 – 3.35	0.137
Identical [DIFF] * Language [Hebrew]	-6.81	-19.07 – 5.45	0.276
Identical [DIFF] * Language [Hindi]	12.13	-2.76 – 27.01	0.110

Identical [DIFF] * Language [Hungarian]	-6.20	-18.85 – 6.44	0.336
Identical [DIFF] * Language [Norwegian]	-7.22	-17.95 – 3.50	0.187
Identical [DIFF] * Language [Polish]	6.73	-10.81 – 24.27	0.452
Identical [DIFF] * Language [Portuguese]	3.27	-13.05 – 19.58	0.695
Identical [DIFF] * Language [Serbian]	-8.38	-18.85 – 2.09	0.117
Identical [DIFF] * Language [Simplified Chinese]	-3.52	-17.37 – 10.33	0.619
Identical [DIFF] * Language [Slovak]	-1.95	-13.20 – 9.31	0.735
Identical [DIFF] * Language [Spanish]	15.30	3.60 – 27.00	**0.010**
Identical [DIFF] * Language [Thai]	11.77	-5.84 – 29.37	0.190
Identical [DIFF] * Language [Traditional Chinese]	-4.56	-15.50 – 6.39	0.415
Identical [DIFF] * Language [Turkish]	1.56	-7.16 – 10.29	0.726
Random Effects			
σ^2	13507.20		
$\tau_{00\ \mathrm{Subject}}$	5661.86		
$\tau_{00\ \mathrm{PSA\_ID}}$	656.95		
$\tau_{00\ \mathrm{Picture1}}$	614.01		
$\tau_{11\ \mathrm{PSA\_ID.IdenticalDIFF}}$	14.89		
$\rho_{01\ \mathrm{PSA\_ID}}$	1.00		
N_{Subject}	2687		
N_{Picture1}	48		
$N_{\mathrm{PSA\_ID}}$	50		
Observations	62164		
Marginal R^2 / Conditional R^2	0.078 / NA		