The Role of Cognitive Load and Spatial Abilities on Gesture Production

by

Bayimkhanim Bahmani

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Science

Department of Psychology

University of Alberta

Abstract

Cognitive load theory posits that due to our limited cognitive resources, gestures can be used as external storage to maintain some of the information as the load increases. In this study, we used a narrative task to examine the significance of cognitive load and spatial abilities on gesture production. We designed three conditions, each with an increasing level of difficulty and predicted that the highest level of difficulty would result in the highest rate of gestures. We also investigated the role of spatial abilities, specifically spatial visualization and spatial short-term memory, measured by the Mental Rotation and Corsi Block tests. Previous studies have found that cognitive load affects representational gesture production and we expected to replicate those results. Additionally, we coded for non-representational gestures and made no predictions about their relationship to cognitive load. The results showed that higher levels of cognitive load were associated with increased rates of both representational and nonrepresentational gestures. We found no evidence for the effect of spatial abilities on either gesture type. These findings demonstrate that increased cognitive demands, such as remembering and recalling complex narratives, are associated with higher rates of gestures.

Preface

This thesis is an original work by Bayimkhanim Bahmani. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "The complexity of event properties and their effect on gesture use", No. 00111943, November 11, 2021.

In loving memory of my father, Bahadur.

Thank you for being so strong for such a long time.

Acknowledgements

I would like to express my endless gratitude to Dr Elena Nicoladis, for her continued patience, guidance and support throughout my graduate education. I am grateful for the opportunity to learn from her wealth of experience and knowledge in the field. Thank you for creating a positive learning environment in which I was encouraged to make mistakes and learn from them. I am fortunate for the opportunities Dr Nicoladis has provided for me. I would also like to thank my committee members Dr Cor Baerveldt and Dr Kimberly Noels for their time, attention and contributions to this project.

Table of Contents

Abstract ii

Preface iii

Dedication iv

Acknowledgements v

Table of Contents vi

List of Tables viii

List of Figures ix

Introduction 1

Types of Gestures 1

Cognitive Load Theory 2

Spatial Abilities 4

This study 6

Method 8

Participants 8

Materials and Procedures 8

Coding 10

Analysis 12

Results 13

Descriptive Statistics 13

Correlational Analyses 15

Total gestures 16

Multiple Regression 16

One-way ANOVA 17

Representational and Nonrepresentational gestures……………………………………..18

MANOVA…………………… ……………………………………………....18

Word Count Analysis ……………………………………………………..…….….….20

Discussion…………………………………………………………………………..………22

References………………………………………………………….…………..…………30

Appendix A: Correlational matrices divided by condition……………………………….….….36

Appendix B: Analysis of nonrepresentational gestures by each category ……………………....39

Appendix C: Correlational matrices between gesture categories, word counts and time……..….41

List of Tables

Table 1……………………………………………………………………………………14

Table 2……………………………………………………………………………………16

Table 3……………………………………………………………………………………17

Table 4……………………………………………………………………………………19

Table 5……………………………………………………………………………………22

List of Figures

Figure 1 ……………………………………………………………………………………….18

Figure 2 ……………………………………………………………………………………….20

Figure 3 ……………………………………………………………………………………….22

Over the past few years, gestures have received increased attention from researchers interested in understanding how complex thoughts and language can be manifested in hand movements. *Co-speech gestures* (hereby referred to as gestures) can be defined as hand movements that accompany our speech and are critical to expressing our thoughts and attitudes as well as a wide range of ideas, be it a simple description of a route or an abstract concept of unity by moving our hands or body parts (e.g., head) (Goldin-Meadow, 2005; McNeill, 1992). Gesture production is likely influenced by many factors, including semantic and structural differences across languages (Colletta et al., 2014; Pika et al., 2006), knowledgeability of an audience (Holler & Stevens, 2007; Parrill, 2012), the content of the depicted event (Masson-Carro et al., 2015, 2017), personality (Hostetter & Potthoff, 2012; O'Carroll et al., 2015) and many other variables. The current paper focuses on two possible predictors of gesture production: the level of task difficulty and spatial abilities. These predictors might affect different types of gestures differently. Therefore, before discussing in more detail how these variables might affect gesture production, first, the types of gestures will be discussed.

Types of Gestures

Based on the functions of gestures and the contexts they occur, many researchers have proposed a classification system for the gestures (Kendon, 2004; McNeill, 1992). One of the most notable and utilized classifications was developed by McNeill (1992). He distinguished gestures into two main categories: *representational* and *nonrepresentational*. Representational gestures were defined as gestures that relay the shape or manner (or other semantic properties) of objects or actions present in the speech and as gestures that represent abstract concepts physically. This category encompasses both *iconic* and *metaphoric* gestures and is studied the most due to its ability to concretely represent ideas and spatial processing. The second category of gestures is known as nonrepresentational gestures, and these gestures serve many diverse functions. Depending on the classification scheme, there may be lots of different kinds of non-representational gestures. For example, McNeill (1992) placed only the *beat* gestures, rhythmic and rapid hand movements that do not contain any meaning, within the nonrepresentational gesture category. On the other hand, Kendon (2004) included both beat and *deictic* or pointing gestures, movements that would point out concrete or abstract entities, ideas, etc. as part of the nonrepresentational gestures. He also introduced a new gesture category and presented it as *pragmatic* gestures (Kendon, 2004, 2017). According to Kendon (2004, p. 5), pragmatic gestures emphasize the nonrepresentational part of speech and act as "semantic operators or as punctuators or parsers of the spoken discourse". For instance, these gestures can support social discourse by providing context cues, highlighting ideas, conveying the speaker's attitude, maintaining the interaction with addressees, etc. Thus, pragmatic gestures are also frequently categorized as nonrepresentational gestures by other researchers (Chu et al., 2014; Mol et al., 2009).

Cognitive Load Theory

One possible function of gestures is to decrease cognitive load (Goldin-Meadow et al., 2001; Maricchiolo et al., 2014; Mol et al., 2009; Pouw et al., 2014). The basic premise of the cognitive load theory is that working memory (WM) has limited capacity, and upon performing novel and complex tasks individuals employ some mechanisms to reduce the strain of information processing (Paas & van Merriënboer, 2020). Gesturing is one of the mechanisms that allows speakers to manage the WM resources via two possible processes. The first possibility is that gestures act as external storage by physically maintaining the novel information and preventing its loss, while the WM allocates the freed-up resources to other processes (Goldin-Meadow et al., 2001; Ping & Goldin-Meadow, 2010; Sepp et al., 2019). The cognitive load can be intrinsic (e.g., finding the solution to the problem) or extrinsic (e.g., construction noises), and with the increasing demand of the task, offloading cognition onto the physical world (by using gestures more) can improve the cognitive performance despite limited resources (Risko & Gilbert, 2016). Alternatively, gestures can serve as an auxiliary modality to formulate the current problem from a different perspective, thus allowing the individuals to grasp the problem faster (Sepp et al., 2019). For example, Cook et al. (2008) showed that children who were taught to use the correct strategy in their gestures while solving a mathematical problem were able to use the same strategy a month later. They concluded that the use of correct strategy in gestures solidified the knowledge.

Many previous studies have supported the notion that gesture production is related to cognitive load. In a dual-task paradigm, Ping and Goldin-Meadow (2010) demonstrated the benefits of gesturing while solving a task. They instructed children to explain the solution to a Piagetian conservation task while they were asked to keep two random words in their minds. The children's gesturing was manipulated such that some participants were not allowed to gesture or move their hands while others were allowed to gesture freely while explaining the solution to the task. The presence of the objects from the conservation task was also manipulated to create two conditions with differing levels of difficulty (present - easy vs absent - difficult). The results showed that those who gestured and were assigned to the difficult condition showed better recall of words than those who were not allowed to do so. They concluded that gesturing reduced cognitive load, thus improving performance.

Results to date suggest that cognitive load might affect representational gestures in particular. Hoetjes and Masson Carro (2017) investigated how different types of loads (i.e., baseline, verbal and motoric) affected the production of representational and nonrepresentational gestures. They used the dual-task situation, and in addition to remembering the main task (verbal or motoric load), they concurrently performed a secondary task by describing complex geometric shapes. Compared to the baseline condition in which only one task was completed, both load conditions resulted in higher representational gesture rates and yielded a similar increase in gestures. However, nonrepresentational gesture rates did not differ across the three conditions. They concluded that increasing the cognitive load of either verbal or motoric tasks would result in more representational gestures regardless of the task's nature.

Although many researchers support the idea that difficult tasks increase gesture rates, some contradictory results were also demonstrated. One study found results opposite to those predicted by cognitive load. Mol et al. (2009) investigated the offloading properties of gestures under different levels of cognitive load. The load was controlled using either one long cartoon (high load) or the short version of the same cartoon cut into shorter scenes (low load). Similar to previous findings, representational gestures were used for the study; however, one crucial difference was the inclusion of nonrepresentational gestures. The results suggested that the representational gesture rates were the highest when the task demand was easier while the task load did not affect the rate of nonrepresentational gestures.

In sum, most (but not all) previous studies have shown that representational gesture production increases with increasing cognitive load.

Spatial Abilities

Do individuals' spatial abilities play a role in gesture production? McNeill (1992) pointed out that gestures contain imagistic elements of ideas, and that mental images or representations underlie the emergence of these elements. Several studies have found a link between gestures and spatial processing (Hostetter & Alibali, 2007; Kita & Özyürek, 2003) and have shown that gestures can facilitate the maintenance of spatial imagery in memory. For example, in a picture description task, Wesp et al. (2001) found that participants who described a picture from memory produced significantly more gestures than participants who described a picture while looking at it. Morsella and Krauss (2004) also conducted an analogous picture description experiment. Similar to previous findings, participants gestured more when images were not present. Moreover, the complex images (compared to simple ones) also elicited more gestures indicating some association between spatial abilities and gesture production.

Several studies indicate that variations in individual spatial abilities may influence the number of gestures and types of gestures produced. Given that representational gestures allow us to map elements from the physical world (e.g., routes, shapes) onto our hands, it made them more accessible to study spatial abilities (Alibali et al., 2000; Chu et al., 2014). In a study conducted by Hostetter and Alibali (2007), the authors demonstrated that participants with high spatial visualization abilities and low verbal skills were more likely to produce representational gestures (but not beat) while narrating a cartoon. They suggested that an increase in representational gestures could be the result of compensation for poorer verbal abilities. Such that, individuals who would struggle with organizing their train of thoughts in a verbal form could instead translate their mental representations into representational gestures and support the communication via additional mechanisms.

However, two studies found that poorer spatial abilities were associated with more gesture production. Göksun et al. (2013) asked participants to solve a Mental Rotation (MR) task and later describe the solution to the experimenter. They showed that individuals with low spatial visualization capabilities (as measured by paper-and-pencil MR test) were more likely to gesture while describing the solution to the task than those with high capabilities. Similarly, Chu et al. (2014) also investigated how individual differences in cognitive abilities affected the rates of gestures while describing abstract concepts. Specifically, they measured spatial and visual WM and spatial transformation abilities. Their results indicated that spatial WM was negatively associated with representational gestures, but it only reached a marginal significance. On the other hand, they found that poorer spatial transformation abilities resulted in significantly more representational and nonrepresentational gesture production.

Still another study found no relationship between individual spatial abilities and gesture production. Smithson and Nicoladis (2013) investigated the role of visuospatial and verbal abilities in predicting iconic gesture production. Participants were first given a battery of tests measuring both verbal and visuospatial abilities (short-term memory and WM). After completing the measures, participants then were asked to watch cartoon clips and narrate them. The results showed that no measures of visuospatial abilities were associated with representational gesture productions. Moreover, only individuals with poorer verbal abilities were more likely to produce iconic gestures.

In sum, studies have shown variable results as to whether individual spatial abilities are related to gesture production. One of the possible causes of this variation could be due to the difference in measurements used for quantifying spatial abilities. Some studies have focused on spatial working memory (i.e., the mental ability to maintain dynamic visual forms in working memory) (Chu et al., 2014; Smithson & Nicoladis, 2013), whereas others have focused on spatial visualization skills (i.e., the mental ability to generate and manipulate spatial objects) (Carroll, 1993). In the present study, we utilized two measures of spatial abilities, short-term memory and visualization, measured by the Corsi Block test and MR task, respectively, to investigate the role of spatial abilities and gesture production.

This study

The first objective of this study was to test whether cognitive load affects the rates of gesture production. Cognitive load was operationalized as a function of task difficulty. The difficulty was manipulated in two ways: by changing the number of videos watched at a time and whether participants narrated while simultaneously watching the videos or after finishing. Most previous studies have shown that gesture production increases with increased cognitive load (Goldin-Meadow et al., 2001; Hoetjes & Masson Carro, 2017; Ping & Goldin-Meadow, 2010). We expected to replicate those results here. We hypothesized that narrating the task as it unfolds (i.e., condition A) would put the least strain on participants' cognition and thus elicit the least number of total gestures and representational gestures. We also expected that watching all videos and describing them from memory would be highly taxing on cognition and result in the highest number of total and representational gestures to compensate for the load. Finally, we included a third condition, in which participants narrated each video excerpt right after watching them and predicted that its load would be neither too taxing nor too easy and elicit a moderate number of total and representational gestures would serve as a baseline. Because of the limited data in literature, no predictions were made for the effect of cognitive load on the nonrepresentational gestures. Although we did not include any specific measures to assess the cognitive load experienced by participants, Zhou et al. (2018) suggested that variables such as word count can be used to determine the level of participants' cognitive load. That is, higher levels of load would translate into a lower number of words used by participants. Thus, we used word count as a way to measure cognitive load.

A second objective of this study was to test whether an individual's spatial abilities (spatial visualization and short-term memory) are related to gesture production. Short-term memory measures were reported to be more robust predictors for gesture production (Chu et al., 2014; Pyers et al., 2021; Smithson & Nicoladis, 2013). Therefore, in addition to measuring spatial visualization skills, we also assessed participants' short-term memory. Due to the opposing results found in the literature, we made no predictions regarding spatial abilities and the rates of gesture production.

The data to date suggests that nonrepresentational gestures are not necessarily affected by cognitive load (Hoetjes & Masson Carro, 2017). Most previous studies have linked both cognitive load and individual spatial abilities specifically with representational gestures. However, Chu et al. (2014) argued that nonrepresentational gestures can reduce memory load. Our last objective was to include both representational and nonrepresentational gestures (as well as total gestures) and explore their association with task difficulty and spatial abilities.

Method

Participants

Only native English speakers were recruited for this study. In total, data from 104 participants (72 female, 32 male) were used in all analyses, and the mean age was 21 (*SD* = 4.75), ranging from 18 to 46. Participants were randomly assigned to one of three conditions varying on the degree of cognitive load.

The data collection consisted of two phases. In the first phase, 17 university students participated in the study in exchange for $10. During the study's second phase, data included 121 first-year psychology students who participated in exchange for course credits. Data from four non-native speakers and from 30 additional participants who experienced technical problems (e.g., weak internet connection, poor audio and video) or whose arm movements and torso were not visible were excluded (not included in *N*). Due to the nature of the conditions and their effect on gesture rates both gesturers and non-gestures were included during the analyses.

Materials and Procedures

We used 12 short animations constructed by Kita and his colleagues (2007) to elicit gestures. These videos are comprised of two figures (i.e., red circle and green triangle) that interact with one another while performing different motions (e.g., jumping, tumbling, pushing, etc.). The length of each video ranged from 6 seconds to 14 seconds.

Three conditions were designed, each with a different difficulty level (or cognitive load). The first condition (i.e., A) was designed to have the least cognitive load. Participants watched each video one at a time and were instructed to narrate the videos while watching them. The second condition (i.e., B) had moderate levels of load, and participants were instructed to watch one video at a time and then narrate from memory by the end of each video. Lastly, the third condition (i.e., C) had the highest levels of load, and it was achieved by asking participants to watch all 12 videos at once, which was one minute and 55 seconds long, and then narrate everything they could remember. The videos in all three conditions had the same order. Condition A had 32 participants in total, whereas conditions B and C had 35 and 37 participants, respectively.

Another aim of this study was to examine how spatial abilities affect gesture rates. The MR task and the Corsi Block test were used to achieve this. The online versions of these tests were obtained from the free-to-use website PsyToolkit, and the links to these tests were provided to participants (Stoet, 2010, 2017). The MR task consisted of two sessions. The first session was a training session with five trials, whereas the second part was the main session and consisted of 10 trials. Participants were shown a figure for this task and asked to select the matching figure from two options. For the Corsi Block test, the software requires participants to mimic the order by which the blocks change colors by tapping the correct order at the end of the demonstration. The number of blocks ranged from two (the easiest sequence to follow) to nine blocks, resulting in a final score ranging between two and nine.

Due to the limitations brought by COVID-19, all data were collected online by trained research assistants for safety. All participants were sent a Zoom link and were given informed consent regarding the nature of the study. They were informed that the study required audio and video recordings of their narrations. At the beginning of each session, participants were instructed to position their cameras such that their hands and torso were in the frame at all times. After instructions were given, the PowerPoint slide containing the videos was sent to them, and the recording was started. We did not put any time restrictions on narrating the videos, thus allowing a natural flow of narrations. Once all the videos were finished, the recording was stopped, and the participants were sent the link for the Corsi Block test and the MR task. In the end, all participants were debriefed and thanked for their participation.

Coding

The speech recorded in the videos was transcribed and gestures were coded using EUDICO Linguistic Annotator (ELAN) 6.3 program (ELAN, 2022). According to Kendon (2004), gestures are comprised of several phases such as preparation, stroke and retraction, of which the strokes are the main part that is analyzed for its content. The strokes were defined as the phases of gesture that represented the gist of gestures (e.g., shape, trace, rhythmic movement) and required the most effort to carry out the movement (McNeill, 1992). For the current study, the total number of gestures, as well as two categories (i.e., representational and nonrepresentational), were analyzed. Representational gestures were operationalized according to the definitions provided by McNeill (1992) and included only iconic gestures since no metaphoric gestures were produced by participants. Nonrepresentational gestures included beat, deictic and pragmatic gestures and were originally coded more finely and then included in the analysis under the broad category in order to address the research questions. Except for the pragmatic gestures, which were coded per Kendon's (2004, 2017) classification, the other categories were coded according to McNeill's (1992) specifications. The total number of gestures was obtained by adding up two categories of gestures. For the analysis, only gesture strokes were examined, and each complete gesture stroke was assigned to either of the categories. Only in the instances in which participants produced beat gestures in addition to other categories was the stroke assigned to two categories. For example, if a participant used repeating rhythmic movements while pointing in the direction that a figure was moving, it was coded as both deictic and beat. Self-adaptors (e.g., head-scratching, playing with items, clothes, touching face) were omitted.

To account for the differences in the length of narrations, gesture rate per 100 words was used. That is, the gesture rate was obtained by multiplying the total number of gestures by 100 and then dividing it by the total number of words used to describe the video tasks. This calculation was done for both categories of gestures. For the word count calculations, we added the number of all the words used by each participant while describing the content of the videos. The fillers such as *um, hmm, ehm* were not included in the count. For partial correlational analyses, word counts, and raw gesture counts were used while controlling for the time that took to narrate events. The time was measured in seconds and included only the time participants described the video contents.

For the Corsi Block test, the highest span achieved by participants was used for the analyses. The range of the span changes between two and nine, and each number denotes the highest sequence remembered by a participant. The MR task yielded two scores: accuracy rate and response time for each trial. Our analysis, measured by Kendall rank correlation, indicated no trade-off between the accuracy rates and average response times (τ = -.04, p = .61). Moreover, Hirschfeld et al. (2013) found that the accuracy did not improve by increasing the response time for the task; thus, for this study, we only used the response time as an indicator of spatial abilities. Given that each participant had 10 trials, we averaged their response times to obtain one score for each. For both tests, outliers were identified, and the results of participants (i.e., four) that deviated from the mean within two standard deviations were excluded.

Analysis

The first part of the analysis explored the correlations between all variables using the Pearson correlation method. During the first correlation analysis, we did not separate the variables based on conditions. For the subsequent analyses, correlations among variables were measured for each condition separately to make sure that the condition effect was not interfering with the results (see Appendix A).

We wanted to understand how spatial short-term memory and WM contribute to the frequency of total gestures. Thus, a multiple regression analysis using three variables, load and short-term spatial abilities and visualization as predictor variables and total gesture rate as outcome variable was conducted. Due to the categorical nature of cognitive load, we were unable to obtain standardized coefficients (β) and reported only unstandardized coefficients (*B*). The initial model (i.e., Model 1) excluded 12 observations out of 104, due to the missing values in short-term spatial abilities and/or visualization, and explained little variation observed in the results ($R^2_{multiple}$ =.20, $R^2_{adjusted}$ =.16). Therefore, an additional analysis looked at only the difficulty levels of conditions (cognitive load) and how they affected the total rate of gestures. This new model (i.e., Model 2) did not have any missing values (i.e., 104 observations), so it produced better adjusted R^2 ($R^2_{multiple}$ =.21, $R^2_{adjusted}$ =.20). We also ran another model (i.e., Model 3) in which cognitive load was the only predictor, and we also excluded the values that were previously missing from Model 1 (i.e., the very same 12 observations). We ran Model 3 to make sure that participants with missing values in spatial abilities did not bias the results. The results of Model 3 showed that cognitive load as a predictor was still significant, $F(2,89)$ = 10.62, $p < .001$, and the adjusted R^2 (but not multiple R^2) improved ($R^2_{multiple}$ =.19, $R^2_{adjusted}$ =.17). The difference between Model 1 and Model 3 was not significant (p = .73), indicating that spatial abilities did not improve model. For the one-way Analysis of Variance (ANOVA) we used only cognitive load as the predictor and included the observations of all 104 participants.

Looking into only total gesture rates might not allow seeing how other categories of gestures are affected by the strain on cognition. Thereby, two types of gestures - representational and nonrepresentational - were chosen as dependent variables, while only the cognitive load levels were used as independent variables. Multivariate Analysis of Variance (MANOVA) was implemented for this purpose. Due to the robust nature of Pillai's Trace, it was used as the method of multivariate statistics. After obtaining a significant effect of cognitive load levels, the next step was understanding which of the two categories of gestures differed across conditions. Due to the unbalanced design across conditions a robust alternative for the ANOVA, the Kruskal – Wallis test, was chosen as the test for the univariate analysis. The multiple pairwise comparisons were conducted after identifying the gestures that yielded significant results. Due to the violation of homogeneity of variance, the Games – Howell post hoc test was used.

Finally, we compared the number of words used by participants across the conditions of cognitive load. Due to the violations in linearity and homogeneity of variance, Kruskal – Wallis rank sum test was used to compare three levels of task difficulty. For the post hoc analyses, we administered pairwise t-test comparisons with the Benjamini – Hochberg (BH) adjustment for uneven group sizes (Benjamini & Hochberg, 1995). We also measured the partial correlations between gesture categories and word count while controlling for the time.

Results

Descriptive statistics

In total, participants produced 3,228 gestures. Upon looking at the production of total gesture rates we found that most of the participants produced some gesture at some point (88%). When we compared all three cognitive load conditions, noticeable differences emerged in the production of overall gesture rates. Almost 22% of participants in condition A did not produce any gestures at all, compared to 14% and 3% non-gestures in conditions B and C, respectively. The majority of gestures were representational gestures (56%), followed by nonrepresentational gestures (44%). We analyzed the subcategories of representational gestures (i.e., hand-as-hand, hand-as-object, trajectory) and found that nearly 63% of all iconic gestures were trajectory movements. The mean value for the Corsi Block test span was 5.97 (*SD*= 1.12) and the results ranged from 3 to 8. The average response time for the completion of the MR task was 5,982 milliseconds (ms) (*SD*=2,285) ranging from 1,525 ms to 13,400 ms. Table 1 summarizes the descriptive statistics for all the variables for each condition separately.

Table 1

Descriptive statistics for the gesture rates and spatial abilities as a function of cognitive load.

Variable	*M*	*SD*	Minimum	Maximum	*N*
Condition A					
Representational	3.25	4.15	0.00	13.89	32
Nonrepresentational	1.72	2.38	0.00	10.19	32
Total gestures	4.97	6.03	0.00	24.07	32
Corsi Block test	5.97	1.03	4.00	8.00	30
Mental Rotation task (ms)	5695.50	2313.15	1633.00	8357.00	30
Word count	401.94	133.93	153	676	32
Condition B					
Representational	4.58	3.84	0.00	12.78	35
Nonrepresentational	3.52	3.22	0.00	10.81	35
Total gestures	8.10	6.57	0.00	21.09	35
Corsi Block test	6.00	1.24	4.00	8.00	32
Mental Rotation task (ms)	5953.74	1776.64	2794.00	9449.00	34
Word count	497.71	266.19	187	1558	35

Condition C					
Representational	6.44	4.46	0.00	16.41	37
Nonrepresentational	6.61	4.50	0.00	18.80	37
Total gestures	13.05	6.89	0.00	25.15	37
Corsi Block test	5.94	1.10	3.00	8.00	34
Mental Rotation task (ms)	6261.82	2706.55	1525.00	13400.00	34
Word count	192.54	105.83	46	448	37

Note. Gesture rates were calculated as the number of gestures per 100 words. Response time for the Mental Rotation task is measured in milliseconds (ms). Corsi Block task results were represented by the highest span achieved (ranging from 2-9).

Correlation analyses

Table 2 shows the correlations between all dependent and predictor variables not controlled for the conditions. Only the correlation between representational and nonrepresentational gestures was significant. The positive correlation indicated that those who used more representational gestures also produced more nonrepresentational gestures. Next, we looked at correlations between predictor and dependent variables for each condition separately. Similar to the overall results, both types of gestures showed positive significant correlations in conditions A and B, $r(102) = .68$, $p < .01$, and $r(102) = .73$, $p < .01$, respectively. In condition B, representational gestures also positively correlated with Corsi block test results indicating that those who showed high visuospatial short-term memory skills also tended to produce more representational gestures, $r(102) = .42$, $p < .05$. Another significant correlation was found between the total number of words used during the experiment and nonrepresentational gestures in condition B, $r(102) = .39$, $p < .05$. No other significant correlations were found across three conditions among the variables (for more details please see Appendix A).

Table 2

Pearson correlation coefficients for Predictor and Dependent variables with confidence intervals

Variable	1	2	3
1. Corsi Block test			
2. Mental Rotation task	-.11 [-.31, .10]		
3. Representational	.10 [-.11, .29]	.07 [-.13, .27]	
4. Nonrepresentational	.02 [-.18, .22]	.08 [-.12, .28]	**.51**** [.35, .64]

Note. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). ** $p < .01$.

Total gestures

Multiple regression

While there was little evidence for correlations between gesture production and spatial abilities (see Table 2 and Appendix A), different levels of cognitive load, the Corsi Block test results (the highest span) and the MR task results (average response time) were the main predictors of the total gesture rate. Due to some missing values, the model with spatial abilities and cognitive load included only 92 participants (12 observations were deleted by the model). As predicted, the different levels of cognitive load indeed had a significant effect on the total gesture rates. Corsi Block test did not significantly predict the gesture rates, $B = .42$, $p = .52$. Similarly, the average response time required to solve the MR task also had no main effect on the gesture rates, $B = .00$, $p = .59$. Cognitive load, specifically condition C compared to condition A, significantly predicted the total gesture rates, $B = 7.64$, $p < .001$. The results of regression showed that the model explained 16% of the variance ($R^2_{multiple} = .20$, $R^2_{adjusted} = .16$, $F(4, 87) = 5.39$, $p < .000$) The more detailed results of multiple regression are shown in Table 3.

Table 3

Multiple regression results using total rate of gestures as the criterion

Predictor	*B*	*B* 95% CI		*SE B*	*t*	*p*
		LL	UL			
(Intercept)	1.95	-7.19	11.10	4.60	0.42	.67
Condition B	2.37	-1.11	5.85	1.75	1.35	.18
Condition C	**7.64**	**4.19**	**11.10**	**1.74**	**4.40**	**.00 *****
MR task	0.00	-0.00	0.00	0.00	0.53	.59
Corsi Block test	0.42	-0.87	1.71	0.65	0.65	.52

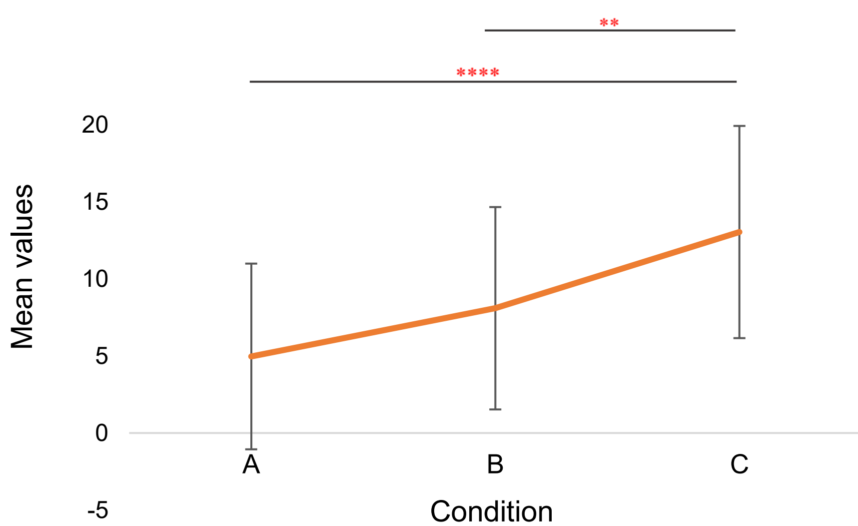
Note. *B* represents unstandardized regression weights. *SE B* represents the standard error of *B*. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. Since the cognitive load was a categorical variable, the condition A were used as the reference for comparison and represented as intercept. *** p < .000.

One-way ANOVA

To test for possible effects of cognitive load on gesture production, we conducted a one-way Analysis of Variance (ANOVA). The results showed a significant main effect of condition, $F(2, 101) = 13.54$, $p < .000$, $\eta^2_g = .21$ and a higher variation explained by the model ($R^2_{multiple} = .21$, $R^2_{adjusted} = .20$). Post hoc analyses using the Tukey's honest significance difference (HSD) adjustment for multiple pairwise comparisons indicated that two comparisons yielded statistically significant results. No gesture rate difference was found between the medium and low load conditions, $p = .127$. On average, participants in condition C with the highest cognitive load ($M = 13.05$, $SD = 6.89$) gestured significantly more than the conditions with medium ($M = 6$, $SD = 1.24$) and low loads ($M = 4.97$, $SD = 6.03$), $p = .005$ and $p < .000$, respectively (refer to Figure 1).

Figure 1

Mean values for the total gesture rates across three conditions.



Note. The error bars represent the standard deviation. The cognitive load of conditions increases from A to C alphabetically. That is, condition A represents the least load, whereas condition C represents the highest load level. $^{**}p < .01$, $^{***}p < .001$, $^{****}p < .0001$

Representational and Nonrepresentational gestures

MANOVA

In the second part of the analysis, we wanted to look at whether different categories of gestures are affected when the cognitive load of the task is manipulated. Given that both the Corsi Block test and the MR task did not yield significant results, we assumed their contribution would be minimal for this analysis phase and included only the cognitive load conditions. In the

MANOVA, the effect of conditions yielded statistically significant results, $F(4, 202) = 7.41$, $p < .001$, therefore, we performed a univariate analysis to identify which categories differed across the conditions.

The results of the Kruskal-Wallis test showed that both representational and nonrepresentational gestures demonstrated statistically significant differences across conditions (refer to Table 4). Since the homogeneity of variance assumption was violated, multiple pairwise comparisons were performed by utilizing the Games - Howell test. The results indicated that when compared to condition A, representational and nonrepresentational rates were produced significantly more during condition C, $p = .008$, and $p < .000$, respectively. Nonrepresentational gestures were also elicited significantly more during condition C, $p = .004$, compared to condition B. The final significant difference in rates was observed between conditions A and B for the nonrepresentational gestures, and more nonrepresentational gestures were produced in the latter one, $p = .03$ (Figure 2).

Table 4

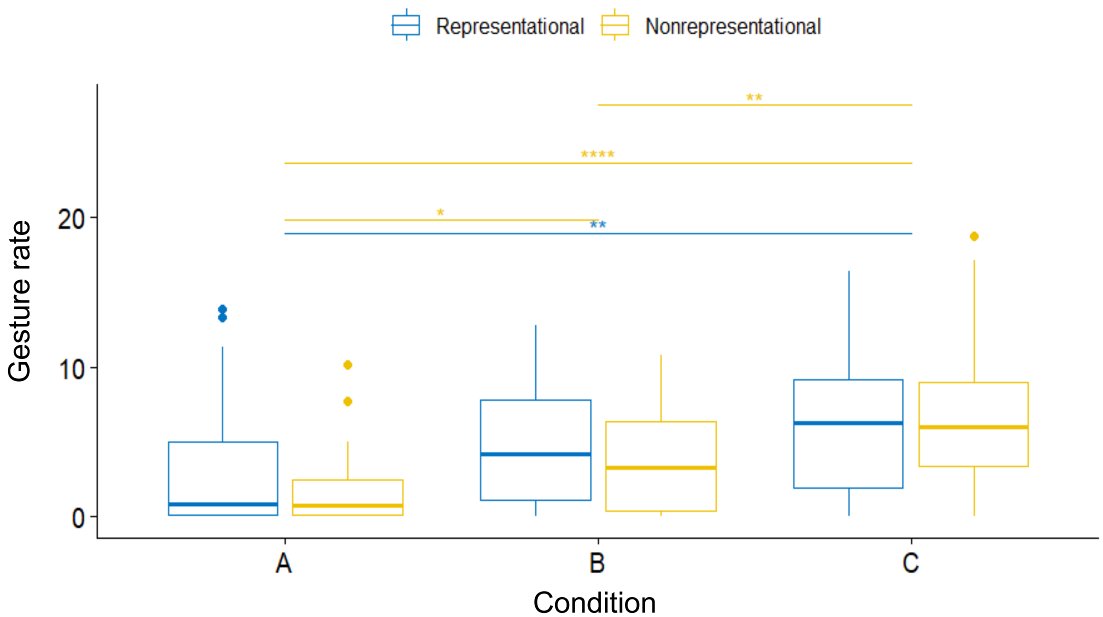
Kruskal-Wallis test results for gesture categories.

Categories	df_{Num}	n	H	p
Representational	2	104	11.2	**.004****
Nonrepresentational	2	104	26.68	**.000*****

Note. $^{**}p < .01$, $^{***}p < .001$. df_{Num} indicates the degrees of freedom numerator.

Figure 2

Boxplots demonstrating significant gesture rate differences.



Note. Gesture rates per 100 words for each category across conditions. The lower and upper limits of the box represent 25th and 75th percentile of data whereas the midline represents median. The whiskers show 1.5 times the interquartile from top or bottom to the furthest data point. The cognitive load of conditions increases from A to C alphabetically. That is, condition A represents the least load, whereas condition C represents the highest load level. $*p < .05$, $**p < .01$, $****p < .0001$

Word count analysis

Initial results of descriptive analyses showed a distinct difference in the number of words that the participants used across conditions. That is, the mean number of words used by participants in condition C (M = 192.54, SD = 105.83) was much lower than the participants in conditions A (M = 401. 94, SD = 133.93) and B (M =497.71, SD = 266.19). The Kruskal – Wallis test result indicated that there was a significant difference in the total number of words used by participants as a function of task difficulty levels, $H(2) = 46.28$, $p < .001$. Pairwise t-test comparisons with the BH adjustment were used to compare all the groups. The difference between conditions A and B reached only a marginal significance, $p = .065$, however, the other two comparisons, B vs C and A vs C yielded statistically significant results, $p < .000$ for both results (Figure 3). Table 5 shows the correlations and partial correlations between gestures and the number of words. The correlations between representational and nonrepresentational gestures were large and significant regardless of the conditions. The partial correlations among gesture categories and word counts were measured by excluding the time (seconds) required to narrate the contents of the videos (see Appendix C correlations between time and other variables). Representational gestures and word counts reached statistical significance and large correlation for condition B, and medium correlations for conditions A and C. Nonrepresentational gestures were correlated strongly with word counts in conditions B and C and reached statical significance in both. However, the correlation with word counts for condition A was weak.

Table 5

Partial correlations across conditions

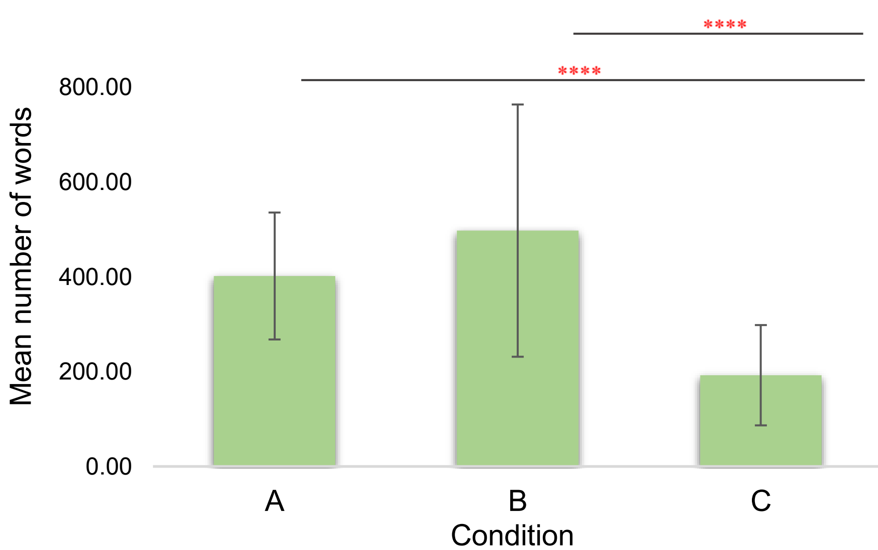
Variable	1	2
Condition A		
1. Representational		
2. Nonrepresentational	**.56****	
3. Number of words	.33	.06
Condition B		
1. Representational		
2. Nonrepresentational	**.83****	
3. Number of words	**.58****	**.43***

Condition C			
	1. Representational		
	2. Nonrepresentational	**.75****	
	3. Number of words	.31	**.62****

Note. Raw gesture numbers (instead of gesture rates) were used for both categories. All correlations except between gesture categories are partial correlations in which time was partialled out. * $p < .05$. ** $p < .01$.

Figure 3

The bar chart for the mean values of word count across conditions.



Note. The error bars represent the standard deviation. The cognitive load of conditions increases from A to C alphabetically. That is, condition A represents the least load, whereas condition C represents the highest load level. ****$p < .000$

Discussion

Cognitive load theory suggested that upon performing a complex task, limited WM resources are depleted, and gesturing can alleviate the strain on cognitive processes by acting either as an external source or by providing a different modality for problem-solving (Sepp et al., 2019). Previous research has demonstrated that the level of cognitive load can affect the rate of gestures produced by individuals (Goldin-Meadow et al., 2001; Hoetjes & Masson Carro, 2017; Ping & Goldin-Meadow, 2010).

Our results were consistent with our hypothesis and previous literature and demonstrated that the number of total and representational gestures vary as a function of the load experienced by the participants. Specifically, we manipulated the load by increasing the amount of information to be remembered and whether participants needed to remember the videos from memory or just narrated the video as they saw it. The results indicated that as the difficulty of a task increased, overall, individuals would gesture more. Due to the nature of our study, certain inferences regarding cognitive load and gesture use are hard to make. Since we did not specifically measure the cognitive load experienced by participants, it is hard to conclude whether our manipulation truly yielded three separate task difficulty levels. However, one possible cognitive load measurement based on speech cues was suggested by several researchers (Berthold & Jameson, 1998; Zhou et al., 2018). They pointed out that under heavy loads, the word count would decrease to offset the cognitive demands of planning longer speeches. We found that the average number of words used in conditions A and B were significantly more compared to condition C, however, this was not true when we compared conditions A and B. The lower count of words and the increased gesture use in condition C would be more likely a result of heavy cognitive load. However, these results still do not answer the question of whether gesturing more, offloads the cognitive load. To test this, we looked at the partial correlations between the number of gestures and the words while controlling for the time. The results showed that for conditions B and C, increased gesture use positively correlated with word counts. This potentially indicates that using gestures indeed frees up the resources allocated to constructing a narration indicated by the increasing number of words used for the story. One interesting finding was the similar word counts between conditions A and B. If we were to follow the same route of reasoning, it would imply that the task difficulty between conditions A and B was not sensitive enough to yield any differences in the word count, however, nonrepresentational gestures were elicited more in condition B than in condition A. An alternative explanation for the similarities in the word count could be ascribed to the fact that participants in the narrating-while-watching condition (i.e., A) did not need to construct a coherent narrative. On the other hand, narrating after watching each animation (i.e., B), provided participants with an opportunity to construct a narrative, thereby leading to nonrepresentational gesture use.

Some studies have shown that the variability in spatial abilities across individuals could be another potential candidate for explaining the gesture rate differences observed across individuals (Chu et al., 2014; Göksun et al., 2013; Hostetter & Alibali, 2007). Our study attempted to understand the role of two spatial ability measures in the production of gesture rates. Unlike some other studies, we found no evidence that supports the effect of spatial abilities (measured by the Corsi Block task and MR task) in gesture production when participants were narrating videos. We also found no association (measured by correlational analyses) between spatial visualization and spatial short-term memory. Two spatial constructs were negatively (but not significantly) related. This could indicate that each task tapped into a fundamentally different ability. Alongside mixed results found in literature in relation to spatial abilities and gesture use, our results support the idea that spatial abilities are not associated with the number of gestures (Pyers et al., 2021; Smithson & Nicoladis, 2013). One explanation for the contradictory results could be attributed to the different nature of the MR task used for this study which was two-dimensional (instead of the three-dimensional that was used in some other studies). Jolicœur et al. (1985) looked at the response time differences between 2D and 3D stimuli and found that participants were faster in responding to 2D stimuli but only when the angular differences were above 60°. Our stimuli were presented in many different angular degrees (including 60° and above), which raises the possibility that our measure was not sensitive enough to capture the individual differences. Nevertheless, we also measured spatial short-term memory with a reliable task (Corsi Block task) and did not find any connection with gesture production. Pyers et al. (2021) proposed that measures targeting short-term memory are more robust measures of spatial abilities than those targeting WM only. Moreover, they looked at the resolution of the tip-of-their-tongue (TOT) phenomenon and demonstrated that spatial short-term memory did not predict the use of representational gestures. Similarly, Smithson and Nicoladis (2013) also did not observe any significant relationship between spatial abilities (short-term memory and WM) and gesture production. Despite finding clear evidence for the effect of cognitive load, the role of spatial abilities in gesturing remains somewhat ambiguous. These findings could be attributed to the ease of studying and controlling circumstantial events (e.g., task difficulty) compared to individual characteristics (e.g., spatial abilities). Nevertheless, our results suggest that the role of spatial abilities in gesture production is not as definitive and robust as it is sometimes assumed.

The current study extended the results of previous studies by including several nonrepresentational gesture subcategories in addition to representational gestures. Our results found that as the cognitive load of tasks increased, participants used more representational and nonrepresentational gestures. Some researchers have proposed that using representational gestures will confer more cognitive benefits than other types of gestures (Cook et al., 2012). For example, Cook et al. (2012) suggested that while solving a task with a heavy load, individuals could use representational gestures to convert verbal propositional information into a visuospatial format and organize it spatially. By making use of two different modalities (verbal and spatial) representational gestures could lighten the load by allocating sources efficiently. Our results did indicate that the representational gestures seem to diminish the burden imposed by the task at hand. Based on these suggestions, one could assume that since nonrepresentational gestures do not convey visuospatial information, they would not help offload the cognitive load. However, our data showed some support for the offloading properties of nonrepresentational gestures since their rates were higher in conditions with the highest cognitive load. While some studies have demonstrated that only representational gestures increase when individuals experience high levels of cognitive load (Hoetjes & Masson Carro, 2017; Mol et al., 2009), other studies have shown that nonrepresentational types of gestures also increase during high-load tasks (Chu et al., 2014). Chu et al. (2014) found a positive association between poor visual WM and some nonrepresentational gestures and suggested that such gestures can also help focus on relevant aspects of tasks while ignoring the irrelevant ones.

Within the context of this study, the increase in nonrepresentational gesture rates with increasing cognitive load could be due to a few possible mechanisms. One mechanism that led to the increase in nonrepresentational gestures could be due to the similarity of animations that we asked participants to retain and then recall. Most of the animation interactions took place in very similar settings and environments (e.g., around a tree, nearby a cliff and a pool) could lead to more errors and hesitations on the part of participants, which in turn would lead to instances that would require non-semantic functions of gestures (e.g., pragmatic). For example, participants could correct themselves (e.g., negating gestures), give an example (e.g., palm up open hand) or shrug their shoulders while palms up to indicate that they cannot remember anything else or that they are uncertain about the correctness of the recalled information.

Alternatively, nonrepresentational gestures could be used for the benefit of a listener and serve only a communicative function. Participants could produce more nonrepresentational gestures to improve the understanding of the listener by highlighting the key events. It was found that observing nonrepresentational gestures (i.e., beat) was found to help with the recall of spatial information and words, and improved the narrative structure of stories told by children (Austin &

Sweller, 2014; So et al., 2012; Vilà-Giménez et al., 2019). Jacobs and Garnham (2007) emphasized that narratives are mainly constructed for the listeners. However, attributing the increases in nonrepresentational gestures to only communicative functions would not explain why as the load of conditions increased (from A to C), so was the rate of nonrepresentational gestures. Moreover, correlational analyses further supported the connection between the task load and nonrepresentational gestures. The strength of associations between nonrepresentational gestures and the number of words, got stronger as the load increased. One might ask how nonrepresentational gestures could be involved in improving the performance of the narrators. We suggest that nonrepresentational gestures could support the story construction and produce better-structured narratives. All three conditions demanded different levels of story construction. The first condition did not have the element of recalling events from memory, and the descriptions did not need to be constructed or organized in a way a narrative would be. The second condition only required the construction of very brief and episodic narratives, and the episodes had a beginning and an ending. The last condition was the most challenging in terms of connecting the storylines into one coherent narrative. Separately, each video had a simple story. Connecting all videos, where geometric figures perform a series of motions and activities with no definitive storyline or story structure could be perceived as very challenging. The activities did not follow any logical pattern (e.g., tumbling up and down the hill) and the manner of the motions was more emphasized throughout the shorts. The differences in the narrative construction across three conditions could increase the demands of planning and thereby cognitive load. The simple animations used in our study could encourage participants to actively explore and construct story structures more, thus leading to increases in nonrepresentational gestures. Differing narrative demands of conditions could also explain why some previous studies could not find the increases in nonrepresentational gestures as a function of cognitive load (Mol et al., 2009). Most of the cartoon videos (e.g., Sylvester & Tweety) used by previous studies had a coherent storyline that did not require participants to construct a logical sequence for the narrative as much. However, these speculations cannot be confirmed based on our study design.

The current study had several limitations, one being the process of data collection. Due to safety precautions, all of our data was collected via Zoom. During recordings, in order to record the movements of participants clearly, we did not ask them to share their screens (which would significantly shrink their video screen); hence, the research assistants relied on the participants to follow instructions throughout the experiment. Furthermore, some participants joined the Zoom meetings from the comfort of their homes, whereas others were in public places. Although we removed data that was not recorded in a quiet environment, setting differences could still contribute to some variations between participants. One other limitation of our study could be the nature of videoconferences and the constraints that come while using them. Specifically, Bailenson (2021) argued that problems related to sending and receiving nonverbal cues and prolonged exposure to the listener's gaze could increase cognitive load. The author also mentioned how viewing oneself while talking could increase the attention to self and make one more self-conscious. In our study, the participants assigned to the first and the second conditions had little opportunity to be affected by the abovementioned problems since they needed to keep the slides with video tasks open at all times. However, for the last condition, participants could open the Zoom meetings (while narrating to the research assistants) on full-screen mode once all videos were finished and potentially see themselves. The data collection via Zoom meetings may or may not have contributed to the differences observed among the conditions. Thus, these results must be interpreted with caution.

In sum, we investigated the roles of cognitive load and individual spatial abilities in the production of gestures. We found that remembering and narrating a large amount of information increased participants' gesture production consistent with the findings of previous literature. We also found some intriguing evidence, similar to representational gestures, nonrepresentational gestures can also decrease individuals' cognitive load. Future studies should investigate how cognitive load affects gesture rates and further research the conditions under which this relationship is pronounced more.

References

Austin, E. E., & Sweller, N. (2014). Presentation and production: The role of gesture in spatial communication. *Journal of Experimental Child Psychology*, *122*, 92–103. https://doi.org/10.1016/j.jecp.2013.12.008

Bailenson, J. N. (2021). Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior*, *2*(1). https://doi.org/10.1037/tmb0000030

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, *15*(4), 469–489. https://doi.org/10.1080/01638539209544823

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Berthold, A., & Jameson, A. (1998). Interpreting symptoms of cognitive load in speech input. In J. Kay (eds), *UM99 User Modeling. CISM International Centre for Mechanical Sciences,* (Vol. 407). Springer, Vienna. https://doi.org/10.1007/978-3-7091-2490-1_23

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies* (Illustrated ed.). Cambridge University Press.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, *143*(2), 694–709. https://doi.org/10.1037/a0036311

Colletta, J. M., Guidetti, M., Capirci, O., Cristilli, C., Demir, O. E., Kunene-Nicolas, R. N., & Levine, S. (2014). Effects of age and language on co-speech gesture production: an investigation of French, American, and Italian children's narratives. *Journal of Child Language*, *42*(1), 122–145. https://doi.org/10.1017/s0305000913000585

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. https://doi.org/10.1016/j.cognition.2007.04.010

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes*, *27*(4), 594–610. https://doi.org/10.1080/01690965.2011.567074

ELAN (Version 6.3) [Computer software]. (2022). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Göksun, T., Goldin-Meadow, S., Newcombe, N., & Shipley, T. (2013). Individual differences in mental rotation: what does gesture tell us? *Cognitive Processing*, *14*(2), 153–162. https://doi.org/10.1007/s10339-013-0549-1

Goldin-Meadow, S. (2005). *Hearing Gesture: How Our Hands Help Us Think* (New Ed). Belknap Press: An Imprint of Harvard University Press.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, *12*(6), 516–522. https://doi.org/10.1111/1467-9280.00395

Hirschfeld, G., Thielsch, M. T., & Zernikow, B. (2013). Reliabilities of Mental Rotation Tasks: Limits to the Assessment of Individual Differences. *BioMed Research International*, *2013*, 1–7. https://doi.org/10.1155/2013/340568

Hoetjes, M., & Masson Carro, I. (2017). Under load: The effect of verbal and motoric cognitive load on gesture production. *Journal of Multimodal Communication Studies*, *4*, 29–36.

Holler, J., & Stevens, R. (2007). The Effect of Common Ground on How Speakers Use Gesture and Speech to Represent Size Information. *Journal of Language and Social Psychology*, *26*(1), 4–27. https://doi.org/10.1177/0261927x06296428

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial. *Gesture*, *7*(1), 73–95. https://doi.org/10.1075/gest.7.1.05hos

Hostetter, A. B., & Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture*, *12*(1), 62–83. https://doi.org/10.1075/gest.12.1.04hos

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, *56*(2), 291–303. https://doi.org/10.1016/j.jml.2006.07.011

Jolicœur, P., Regehr, S., Smith, L. B. J. P., & Smith, G. N. (1985). Mental rotation of representations of two-dimensional and three-dimensional objects. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *39*(1), 100–129. https://doi.org/10.1037/h0080118

Kendon, A. (1997). Gesture. *Annual Review of Anthropology*, *26*(1), 109–128. https://doi.org/10.1146/annurev.anthro.26.1.109

Kendon, A. (2004). *Gesture: Visible Action as Utterance* (Illustrated ed.). Cambridge University Press.

Kendon, A. (2017). Pragmatic functions of gestures. *Gesture*, *16*(2), 157–175. https://doi.org/10.1075/gest.16.2.01ken

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16–32. https://doi.org/10.1016/s0749-596x(02)00505-3

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, *22*(8), 1212–1236. https://doi.org/10.1080/01690960701461426

Maricchiolo, F., de Dominicis, S., Cancellieri, U. G., di Conza, A., Gnisci, A., & Bonaiuto, M. (2014). Co-speech gestures: Structures and Functions. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body - Language - Communication. Volume 2 (Handbooks of Linguistics and Communication Science Book 38.2.)* (pp. 1461–1473). De Gruyter Mouton.

Masson-Carro, I., Goudbeek, M., & Krahmer, E. (2015). Can you handle this? The impact of object affordances on how co-speech gestures are produced. *Language, Cognition and Neuroscience*, *31*(3), 430–440. https://doi.org/10.1080/23273798.2015.1108448

Masson-Carro, I., Goudbeek, M., & Krahmer, E. (2017). How What We See and What We Know Influence Iconic Gesture Production. *Journal of Nonverbal Behavior*, *41*(4), 367–394. https://doi.org/10.1007/s10919-017-0261-4

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.

Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009). Communicative Gestures and Memory Load. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *31*, 1569–1574.

Morsella, E., & Krauss, R. M. (2004). The Role of Gestures in Spatial Working Memory and Speech. *The American Journal of Psychology*, *117*(3), 411. https://doi.org/10.2307/4149008

O'Carroll, S., Nicoladis, E., & Smithson, L. (2015). The effect of extroversion on communication: Evidence from an interlocutor visibility manipulation. *Speech Communication*, *69*, 1–8. https://doi.org/10.1016/j.specom.2015.01.005

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-Load Theory: Methods to Manage Working Memory Load in the Learning of Complex Tasks. *Current Directions in Psychological Science*, *29*(4), 394–398. https://doi.org/10.1177/0963721420922183

Parrill, F. (2012). Interactions between discourse status and viewpoint in co-speech gesture. In B. Dancygier & E. Sweetser (Eds.), *Viewpoint in Language: A Multimodal Perspective* (pp. 97–112). Cambridge University Press.

Pika, S., Nicoladis, E., & Marentette, P. F. (2006). A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer? *Bilingualism: Language and Cognition*, *9*(3), 319–327. https://doi.org/10.1017/s1366728906002665

Ping, R., & Goldin-Meadow, S. (2010). Gesturing Saves Cognitive Resources When Talking About Nonpresent Objects. *Cognitive Science*, *34*(4), 602–619. https://doi.org/10.1111/j.1551-6709.2010.01102.x

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

Pyers, J. E., Magid, R., Gollan, T. H., & Emmorey, K. (2021). Gesture Helps, Only If You Need It: Inhibiting Gesture Reduces Tip-of-the-Tongue Resolution for Those With Weak Short-Term Memory. *Cognitive Science*, *45*(1). https://doi.org/10.1111/cogs.12914

Risko, E. F., & Gilbert, S. J. (2016). Cognitive Offloading. *Trends in Cognitive Sciences*, *20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Sepp, S., Howard, S. J., Tindall-Ford, S., Agostinho, S., & Paas, F. (2019, February 5). Cognitive Load Theory and Human Movement: Towards an Integrated Model of Working Memory. *Educational Psychology Review*, *31*(2), 293–317. https://doi.org/10.1007/s10648-019-09461-9

Smithson, L., & Nicoladis, E. (2013). Verbal memory resources predict iconic gesture use among monolinguals and bilinguals. *Bilingualism: Language and Cognition*, *16*(4), 934–944. https://doi.org/10.1017/s1366728913000175

So, W. C., Sim Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, *27*(5), 665–681. https://doi.org/10.1080/01690965.2011.573220

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/brm.42.4.1096

Stoet, G. (2017). PsyToolkit. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Vilà-Giménez, I., Igualada, A., & Prieto, P. (2019). Observing storytellers who use rhythmic beat gestures improves children's narrative discourse performance. *Developmental Psychology*, *55*(2), 250–262. https://doi.org/10.1037/dev0000604

Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures Maintain Spatial Imagery. *The American Journal of Psychology*, *114*(4), 591. https://doi.org/10.2307/1423612

Appendix A

Correlation matrices divided by condition

Table A-1

Correlations with confidence intervals for Condition A

Variable	1	2	3
1. Corsi Block test			
2. Mental Rotation task	-.25 [-.56, .13]		
3. Representational	.03 [-.34, .38]	.14 [-.23, .48]	
4. Nonrepresentational	.07 [-.30, .42]	-.01 [-.37, .35]	.68** [.44, .83]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * $p < .05$. ** $p < .01$.

Table A-2

Correlations with confidence intervals for Condition B

Variable	1	2	3
1. Corsi Block test			
2. Mental Rotation task	.02 [-.33, .37]		
3. Representational	.42* [.09, .67]	.00 [-.33, .34]	
4. Nonrepresentational	.21 [-.15, .52]	.05 [-.30, .38]	.73** [.52, .85]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * $p < .05$. ** $p < .01$.

Table A-3

Correlations with confidence intervals for Condition C

Variable	1	2	3
1. Corsi Block test			
2. Mental Rotation task	-.11 [-.44, .25]		
3. Representational	-.13 [-.45, .22]	.00 [-.34, .34]	
4. Nonrepresentational	-.12 [-.44, .22]	.06 [-.29, .39]	.18 [-.15, .48]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * $p < .05$. ** $p < .01$.

Appendix B

Analysis of nonrepresentational gestures by each category

Table B-1

Kruskal-Wallis test results for gesture categories.

Categories	df_{Num}	n	H	p
Iconic	2	104	11.2	$.004^{**}$
Pragmatic	2	104	34	$.000^{***}$
Deictic	2	104	3.70	.157
Beat	2	104	14.1	$.000^{***}$

Note. *p < .05, **p < .01, ***p < .001.

Table B-2

Descriptive statistics for the dependent variables

Variable	M	SD	Minimum	Maximum	N
Condition A					
Iconic gestures	3.25	6.03	0.00	24.07	32
Pragmatic gestures	0.67	4.15	0.00	13.89	32
Deictic gestures	0.79	0.89	0.00	2.61	32
Beat gestures	0.26	1.56	0.00	6.68	32
Total gestures	4.97	0.55	0.00	2.31	32
Condition B					
Iconic gestures	4.58	3.84	0.00	12.78	35
Pragmatic gestures	1.29	1.59	0.00	5.57	35
Deictic gestures	1.52	1.66	0.00	7.35	35
Beat gestures	0.71	0.91	0.00	3.98	35
Total gestures	8.10	6.57	0.00	21.09	35

Condition C					
Iconic gestures	6.44	4.46	0.00	16.41	37
Pragmatic gestures	4.16	3.67	0.00	15.04	37
Deictic gestures	1.20	1.12	0.00	4.27	37
Beat gestures	1.24	1.41	0.00	5.13	37
Total gestures	13.05	6.89	0.00	25.15	37

Note. All variables were calculated as the number of gestures per 100 words.

Appendix C

Correlational matrices between gesture categories, word counts and time.

Table C-1

Means, standard deviations, and correlations with confidence intervals for Condition A

Variable	*M*	*SD*	1	2	3
1. Representational	13.66	18.50			
2. Nonrepresentational	6.38	7.43	**.56**** **[.26, .76]**		
3. Word count	401.94	133.93	.30 [-.06, .58]	.04 [-.31, .38]	
4. Time (in seconds)	188.38	47.25	-.02 [-.37, .33]	-.06 [-.40, .30]	**.37*** **[.02, .63]**

Note. M and *SD* are used to represent mean and standard deviation, respectively. Raw gesture numbers were used. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * $p < .05$. ** $p < .01$.

Table C-2

Means, standard deviations, and correlations with confidence intervals for Condition B

Variable	*M*	*SD*	1	2	3
1. Representational	25.23	24.73			
2. Nonrepresentational	20.80	24.31	**.83**** **[.69, .91]**		
3. Word count	497.71	266.19	**.73**** **[.52, .85]**	**.80**** **[.63, .89]**	
4. Time (in seconds)	210.06	140.76	**.61**** **[.35, .79]**	**.75**** **[.55, .86]**	**.95**** **[.90, .97]**

Note. M and *SD* are used to represent mean and standard deviation, respectively. Raw gesture numbers were used. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * $p < .05$. ** $p < .01$.

Table C-3

Means, standard deviations, and correlations with confidence intervals for Condition C

Variable	*M*	*SD*	1	2	3
1. Representational	13.59	12.74			
2. Nonrepresentational	12.78	11.92	**.75**** **[.56, .86]**		
3. Word count	192.54	105.83	**.76**** **[.58, .87]**	**.74**** **[.54, .86]**	
4. Time (in seconds)	85.38	40.40	**.75**** **[.56, .86]**	**.58**** **[.31, .76]**	**.91**** **[.83, .95]**

Note. M and *SD* are used to represent mean and standard deviation, respectively. Raw gesture numbers were used. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * $p < .05$. ** $p < .01$.