Received: 1 Oct. 2026 | Revised: 4 Nov. 2026 | Accepted: 12 Nov. 2026 | Published online: 25 Nov. 2026
RESEARCH ARTICLE

Contemporary Education and Teaching Research
2026,Vol. 7(11)381-394
DOI: 10.61360/BoniCETR262021001103

Interpreting Body Gestures and Hand Signals: An Exploratory Study of Selected Meanings and Perceived Understanding

Manuel Alejandro Romero Hung[1] & Yanfei Li[1,*]

[1]*Tarim University, P. R. China*

Abstract: Visual gestures can convey meaning while allowing several plausible interpretations. This exploratory cross-sectional study examined selected meanings and perceived understanding for ten static images among 40 volunteers in Alar, China. Participants completed English-language tasks comprising five body-gesture images and five hand-signal images. Each image was accompanied by verbal response alternatives and a yes/no judgment of understanding. Analyses examined image-specific response distributions and an exploratory employment-group comparison of the number of images participants reported understanding. The most frequently selected response accounted for 27.5%–50.0% of selections across the ten images. For the open-palm image, "Stop" received 47.5% of selections and "Wait" received 30.0%; for the circled-finger image, "Okay" received 50.0% and "Good" received 30.0%. Mean perceived-understanding counts were 7.53 for students, 9.00 for workers, and 9.43 for freelancers. An exploratory Welch comparison indicated a difference across employment groups, $F(2, 23.8) = 5.21$, $p = .013$, alongside differences in their age composition. The findings describe variation in selected verbal labels alongside participants' judgments of understanding. They inform the development of visual communication materials and gesture-assessment tasks through clearer specification of context, response alternatives, and criteria for successful interpretation.

Keywords: gesture interpretation, hand signals, perceived understanding, visual communication, multimodal communication, exploratory study

Introduction

A visible gesture could imply a specific action or communicative intent but allow for multiple descriptions of what is observed. Open hand could be seen as a sign of stopping, waiting, or greeting. Hand formation may be seen as a form of evaluation, as a letter, or a number. It is important to understand how viewers discriminate among these options in regard to the psychology of communication since the communication process relies on creating sufficient shared meaning.

Gesture facilitates communication by means of coordinating with speech, action, and situation around it. Experimental studies have shown that people recognize words by means of the informativeness of gestures and movement of the mouth in cases of difficult listening environment (Krason et al., 2022). In addition, interpretation of the gestures relies on the discourse context. Fritz et al. (2021) have shown that discourse context constrains the interpretations given by the recipients of gestures and makes interpretation easier. Thus, gesture becomes a part of meaning construction process.

In this experiment, a more limited context of gesture recognition is examined. Verbal descriptions are used to describe images of gestures and hands in the given condition. This approach allows receiving data on the frequency of interpretations given by viewers for images in the specified response conditions. Moreover, viewers get the chance to report if they think that they understand the image. These results are of importance for two aspects of interpretation.

Gesture forms, conventions, and linguistic experience

The connection between form and meaning differs depending on the communicative phenomena. Emblematic signs feature conventionalized forms and meanings in certain communities, with examples being common signs that convey approval, rejection, and requests to be silent. In other words, interpretation of such signs relies on conventional knowledge that exists in a community, which can stretch beyond or inside the limits of a nation or a language. Depictive gestures relate to features of

Corresponding Author: Yanfei Li
Tarim University, P. R. China
©The Author(s) 2026. Published by BONI FUTURE DIGITAL PUBLISHING CO.,LIMITED This is an open access article under the CC BY License(https://creativecommons.org/licenses/by/4.0/).

actions, objects, and spatial relations. Their interpretation is usually based on speech and context.

Sign languages as natural languages differ in terms of structure and history of use, which makes it necessary to analyze them separately from each other. Study of sign-language acquisition recognizes the difference between signed modality experience and sign-language experience. Thus, Chen Pichler & Koulidobrova (2023) provide explanation of how learning another sign language differs from the experience of learning a first sign language after a spoken one. Hearing status, sign language proficiency, and multilingualism are separate factors of participants' communication background.

Psychological interpretation of forms needs to pay attention to iconicity and transparency. The former refers to the resemblance between a form and its aspects of meaning. The latter refers to how easily the meaning can be deduced from the form. According to Schiefner et al. (2026), iconicity and transparency need to be separated in the research of Dutch Sign Language. Their study demonstrates the differences in judgments of iconicity and its relationship to transparency depending on the background of raters.

The same result was achieved by Tsang & Perniss (2026) when they studied Hong Kong Sign Language. Although groups had the same patterns of iconicity ratings for signs, there were differences in the average iconicity ratings of signs between groups. These results support the idea that the visual resemblance is evaluated through experience. In addition, they explain the necessity to use different methods of assessing visible resemblance when selecting a plausible meaning, evaluating resemblance after receiving a meaning, and understanding conventional sign.

Response selection and perceived understanding

A gesture-image task involves the combination of the information contained within the image and the options presented by the task. Written labels can cue attention to certain actions or communicative functions. Labels can also provide descriptions in varying levels of detail. "Walking" specifies a particular action, while "Moving" leaves room for interpretation. The choice between such labels can indicate the level of description used by viewers in their account of an image without implying that any viewer has understood the image better than another.

Gesture naming research provides an example of this difference. Cairney et al. (2024) collected the meaningfulness ratings and verbal descriptions of gesture video clips. The researchers found that the differences in the vocabulary diversity could coexist with the similarity of the semantic content of the labels. Variation in wording, in other words, needs careful interpretation. Response diversity could be due to lexical preferences, levels of description, and actual differences in meanings attributed to images.

The observer's attitude toward the task also affects access to meanings. Akers et al. (2025) found differences in the neural response to iconic signs between hearing nonsigners who were expecting to learn the signs and those who were not. The results are in line with the idea of the relation between visual form processing and attention and task expectations. This research underscores the need to specify carefully the tasks that participants perform when viewing gestures.

Understanding judgments constitute an additional complication of measurements. An individual can feel that a certain image is understandable to him or her but assign it a different meaning compared to another viewer's meaning. The research into instructional visuals indicates that metacognitive judgments of understandings depend on the criteria of evaluation (Jaeger & Fiorella, 2024). Understanding judgments thus require interpretation themselves and, if calibration of these judgments is needed, comparison with a criterion of performance.

Relevance to language and communication

Research on gesture has implications for practical use in relation to language learning and communication in different linguistic contexts. The results from instructed learning indicate that effectiveness depends on the connection between the forms, meanings, and learning actions (Karadöller et al., 2024). Belío-Apaolaza and Hernández Muñoz (2024) demonstrated the variability in the learning of emblematic gestures depending on the gesture type and testing task.

The current results call for the careful consideration of the task-related aspects in the application of research on gestures to education or intercultural communication. Selection of a label for a static image can serve as an example of evidence of a particular kind. Interpretation is dependent on the choice of available labels, context information, and the assessed outcome. Formulated accounts of meaning of gestures pay attention to the visibility of their contribution into speech and discourse (Ebert, 2024; Lascarides & Stone, 2009).

The present pilot study investigated responses to five body-gesture images and five hand-signal

images in a sample with diverse communication background. The questions under investigation were the following:

1. What was the distribution of labels among the ten gesture images, and what were the differences between images concerning the concentration and variety of selections?

2. How many images were understood by participants, and was there any variation across the reported employment groups?

The current study is devoted to selected meanings and perception of understanding. Contribution of the study includes interpretation of visual data under certain response conditions.

Methods

Design and setting

This research followed the cross-sectional exploratory design approach. The sample size consisted of forty participants who were subjected to two image interpretation tests conducted in Alar, Xinjiang, China. Participants were recruited using the convenience sampling technique.

The test included the interpretation of five images with gestures as well as five images with signals. The participants had to give labels for all the images and indicate their understanding. Figure 1 illustrates the process.

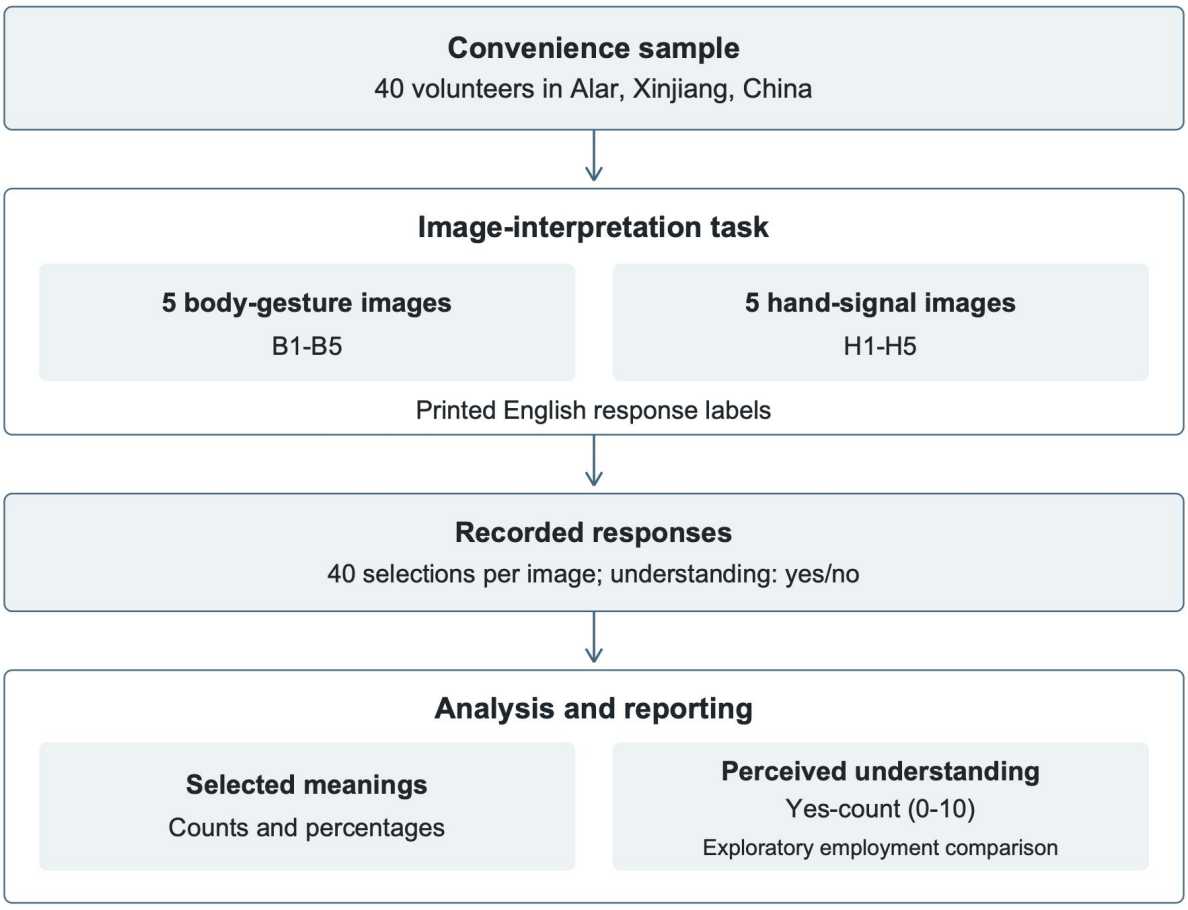

Fig. 1 Study and analysis flow.

Forty volunteers recruited by convenience sampling in Alar completed two tasks comprising five body-gesture images and five hand-signal images with printed English response labels. Each image had 40 recorded selections. Selected response categories are summarized using image-specific counts and percentages. Perceived understanding is represented by the number of "yes" responses across the ten images, with an exploratory comparison across employment groups

Participants

The sample was composed of 10 deaf and hard-of-hearing individuals, as well as 30 hearing people. In the latter case, 19 people were monolingual, whereas 11 others belonged to the multilingual category. The sample consisted of 17 students, 16 workers, and 7 freelancers, with an average age of 27.1 years (SD = 11.1).

These are the types that define the composition of the sample. There was no standard measure of proficiency in sign language and second language in the image tasks. The results of image selection were aggregated throughout the whole sample.

Table 1. Sample characteristics.

Characteristic	Category or summary	n	Percentage
Total sample	Participants	40	100.0
Hearing status	Deaf or hard of hearing	10	25.0
	Hearing	30	75.0
Language background within the hearing group	Monolingual	19	63.3
	Multilingual	11	36.7
Employment category	Student	17	42.5
	Worker	16	40.0
	Freelancer	7	17.5
Age, years	M = 27.1, SD = 11.1	40	—

Note. Percentages for language background use the 30 hearing participants as the denominator. Other percentages use the full sample of 40.

Materials

The materials consisted of ten static pictures made by the researcher and divided into two five-item groups. One group of pictures showed body postures or gestures, and the other – hand postures or gestures. All pictures were accompanied by printed response choices in English, an “Other” choice for any other meanings, and a yes/no “Understood? ” question.

The original image-interpretation task comprised two parts. Task 1 (T1) contained five body-gesture pictures, originally labeled T1-img1 to T1-img5 and referred to here as B1-B5. Task 2 (T2) contained five hand-signal pictures, originallylabeled T2-img1 to T2-img5 and referred to here as H1-H5. Figure 2 shows the reproductions of the original pictures in the published form, depicting configurations, orientation and the order of items. Tables 2 and 3 show the response categories and their frequencies.

The materials evoked interpretations of visible configurations. This task did not have an independently developed scoring key that determined the one correct response for each picture. Therefore, selected responses were considered nominal response categories.

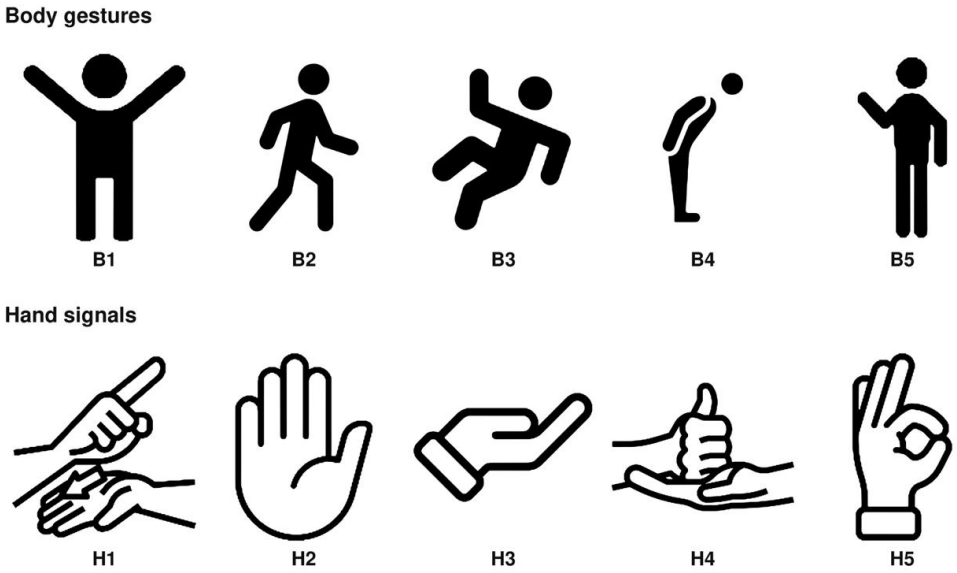

Fig. 2 Publication reproductions of the ten author-created stimulus illustrations.

The upper row presents the five body-gesture images (B1–B5), and the lower row presents the five hand-signal images (H1–H5). The original illustrations were reproduced as scalable outlines while preserving their depicted configurations, orientations, and item order. Tables 2 and 3 report the recorded response categories and frequencies

Procedure

Each participant undertook the task alone. The task included an informed consent process where the participant completed a demographic survey. The participants would then look at the body gesture and hand sign materials and provide their interpretations.

The tasks were done using the English language. This involved descriptive responses together with an opportunity for giving further explanation to the response. The participant also provided either yes or no to indicate whether he/she understands the image. The frequency distribution of the responses had 40 responses to each of the ten images.

Observations were made during the process where the evaluator would make notes of comments and/or questions from the participant. Such observations formed part of contextual information. The quantitatives included the frequencies and understanding responses.

Measures

The first result was the allocation of chosen response categories per each image, including "Nothing". Codes were assigned to each response category per image, item-specifically. The codes played a role of an identifier, and had no meaning regarding quantity or ordination. The perceived understanding was measured using a yes/no question.

The second result was the perceived understanding. The answer "yes" was rated as 1, and "no" as 0 for each image. Summation of the ten answers led to obtaining a score between 0 and 10. A higher score signified that the participant understood more images. It is a specific task understanding measure.

The age and employment categories were used as a description of the groups in the exploratory comparison. The hearing and language background categories were kept as sample characteristic features.

Statistical analysis

Statistical summarization was done using jamovi version 2.7.17.0. Frequencies and percentages per category were the main analyses conducted on all categories of responses. All image-specific percentages were based on 40 responses per image. Most frequent response category was used to define the distribution of responses, but not designated as the right response category.

Counts of perceived-understanding were summarized with means and standard deviation. An exploratory one-way ANOVA with Welch correction was done between students, workers, and freelancers, with allowance for unequal variances. The difference was done on unadjusted group means. Aged counts were provided along with perceived-understanding counts to help interpret the groups.

The overall result was done for three groups. Numerical results obtained were exploratory in nature considering the small sample size and range of perceptions.

Results

Selected meanings for body-gesture images

Several categories were used to describe every body-gesture stimulus. The most frequently chosen category made up 30.0%–42.5% of all choices for the B1–B5 stimuli.

For B1, "Hug" was chosen 15 times (37.5%), and "Big" was chosen 11 times (27.5%). The rest of the choices were split between "Size," "Free," and "Nothing." For B2, "Moving" was chosen 17 times (42.5%), and "Walking" was chosen 15 times (37.5%). They gave a broad description of the action and a more specific one.

For B3, "Falling" was chosen 14 times (35.0%), and "Sliding" was chosen 11 times (27.5%). The choices for B4 were made among the following categories: "Thanks," "Sorry," and "Praying," where the first one got the biggest percentage of responses, 14 times (35.0%). B5 had quite a scattered choice distribution: "Showing something" was chosen 12 times (30.0%), and "Explaining" was chosen 10 times (25.0%).

The Table 2 provides information on all chosen categories, and the Figure 3 presents their distribution.

Table 2. Recorded response distributions for body-gesture images.

Image	Recorded response categories, n (%)
B1	Size: 6 (15.0); Free: 4 (10.0); Hug: 15 (37.5); Nothing: 4 (10.0); Big: 11 (27.5)
B2	Going to…: 7 (17.5); Moving: 17 (42.5); Walking: 15 (37.5); Running: 1 (2.5)
B3	Falling: 14 (35.0); Sliding: 11 (27.5); Sleeping: 8 (20.0); Nothing: 4 (10.0); Escalating: 3 (7.5)
B4	Thanks: 14 (35.0); Nothing: 7 (17.5); Sorry: 10 (25.0); Praying: 9 (22.5)
B5	Showing something: 12 (30.0); Greetings: 6 (15.0); Nothing: 6 (15.0); Explaining: 10 (25.0); At the right: 6 (15.0)

Note. Each image has 40 recorded selections.

The table includes categories with positive counts in the recorded frequency distributions. Capitalization, punctuation, and obvious spelling errors have been standardized while preserving category identities.

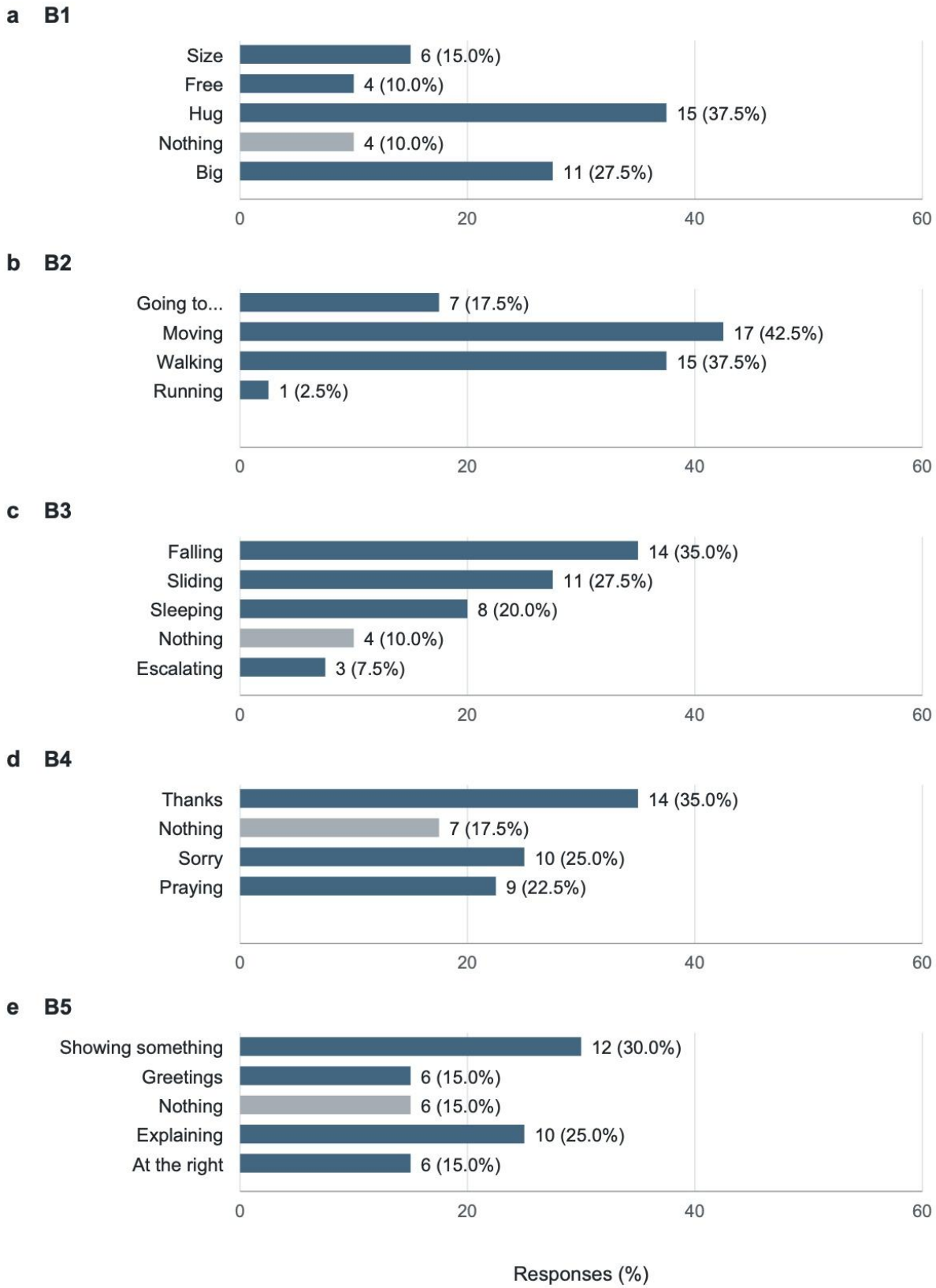

Fig. 3 Response distributions for the five body-gesture images.

Panels a–e correspond to B1–B5. Bars show the percentage of recorded selections for each response category; adjacent labels give the count and percentage. Each image has 40 selections, and all panels use the same percentage scale.

Selected meanings for hand-signal images

Hand-signal images also evoked different kinds of responses. The most common responses in the hand-signal images ranged from 27.5% to 50.0% of selections.

For H1, "Going down" was chosen by 14 respondents (35.0%), and "Nothing" by 12 respondents (30.0%). For H2, "Stop" was chosen 19 times (47.5%), while "Wait" was chosen 12 times (30.0%). Other choices included "My hands are clean" and "Greetings."

In H3, "Borrow" was chosen by 17 respondents (42.5%), followed by "Please," chosen by 10 respondents (25.0%). In H4, the lowest modal choice was shown among the ten images; "Control" was chosen by 11 respondents (27.5%), while "Rule" was chosen by 10 (25.0%). H5 was "Okay" with 20 responses (50.0%), "Good" with 12 (30.0%), "Letter O" with 5 (12.5%), and "Number zero" with 3 (7.5%). H5 is a circled finger image.

From either set of images, no recorded category took up more than half of the selections in an image. Table 3 and Figure 4 provide the hand-signal distributions.

Table 3. Recorded response distributions for hand-signal images.

Image	Recorded response categories, n (%)
H1	Nothing: 12 (30.0); Going down: 14 (35.0); Swap: 6 (15.0); Switch: 8 (20.0)
H2	My hands are clean: 7 (17.5); Wait: 12 (30.0); Greetings: 2 (5.0); Stop: 19 (47.5)
H3	Please: 10 (25.0); Borrow: 17 (42.5); Donation: 4 (10.0); Nothing: 3 (7.5); To give: 6 (15.0)
H4	Rule: 10 (25.0); Nothing: 8 (20.0); Control: 11 (27.5); Help: 4 (10.0); Work: 7 (17.5)
H5	Good: 12 (30.0); Okay: 20 (50.0); Number zero: 3 (7.5); Letter O: 5 (12.5)

Note. Each image has 40 recorded selections. The table includes categories with positive counts in the recorded frequency distributions. Percentages describe selected labels and carry no implication of objective accuracy.

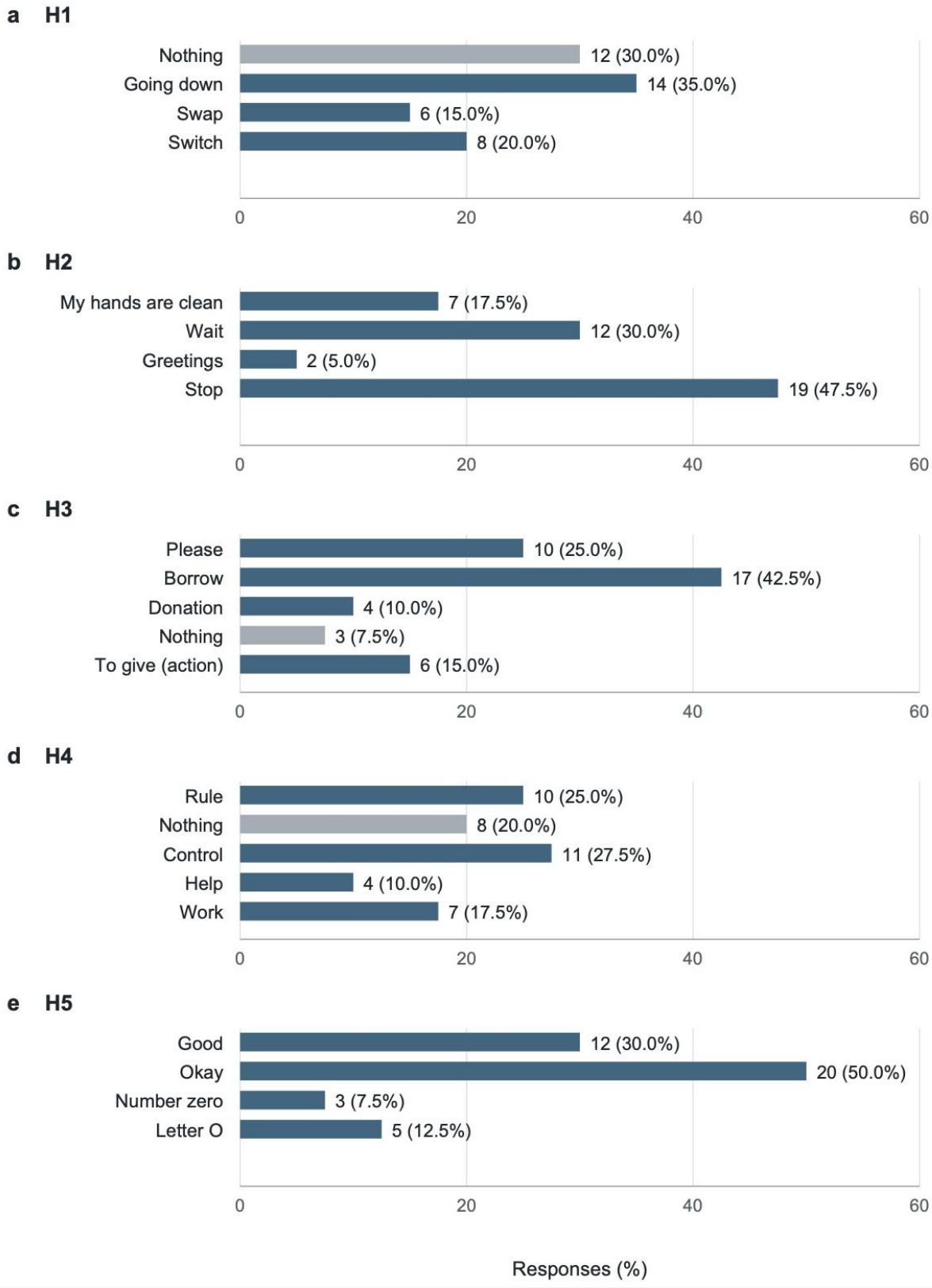

Fig. 4 Response distributions for the five hand-signal images.

Panels a–e correspond to H1–H5. Bars show the percentage of recorded selections for each response category; adjacent labels give the count and percentage. Each image has 40 selections, and all panels use the same percentage scale.

Perceived understanding across employment groups

The average image count of perceived-understanding ratings is 7.53 for students, 9.00 for workers, and 9.43 for freelancers. The corresponding standard deviation is 2.29, 1.03, and 0.54. So students had an image count of less understanding, and the variance is higher among them as compared to the other two groups.

Levene's test indicates variance heterogeneity: $F(2, 37) = 9.47$, $p < .001$. The exploratory comparison using Welch test indicates a significant difference between the three groups, $F(2, 23.8) = 5.21$, $p = .013$. It is a general result related to image counts of perceived-understanding ratings among all three groups.

There is also a difference between the three groups with respect to age; the mean age is 19.53 for students, 34.06 for workers, and 29.29 for freelancers.

Table 4. Perceived understanding and age by employment category.

Employment category	n	Perceived-understanding count, M (SD)	Age in years, M (SD)
Student	17	7.53 (2.29)	19.53 (9.91)
Worker	16	9.00 (1.03)	34.06 (6.98)
Freelancer	7	9.43 (0.54)	29.29 (10.90)

Note. Perceived understanding is the number of "yes" responses across ten images, with a possible range of 0–10.

Higher values indicate that participants reported understanding more images. The employment comparison is unadjusted for age and other participant characteristics.

Discussion

This exploratory study investigates the meanings attributed to ten gesture pictures and perceptions of comprehension. Several labels were identified for each image, and the modal percentage of those ranged from 27.5% to 50.0%. The counts of perceived understanding were relatively higher among workers and freelancers and lower among students. Such results highlight two aspects of visual communication – reaching a consensus on a label and perceiving understanding of a picture.

Response variation and levels of description

Response frequencies demonstrate that selecting meaning for a picture could be based on choosing between a few options. In B2, both "Moving" and "Walking" received nearly equal amounts of attention. The labels vary in terms of their level of generality, since walking is a special case of movement. Thus, the difference in their selection could be related to the description level.

In other pictures, the alternatives pointed to different elements of visible configuration or possible communicative function. Thus, in B1, the alternatives were "Hug" and "Big," while H5 was responded with evaluations in addition to a letter and a number. It means that there are various concepts available for a depicted form in a given set of labels.

The work of Cairney et al. (2024) could serve as a good starting point for further analysis of the found results. The authors have shown the difference between label variety and semantic difference of the labels. The present findings also suggest being cautious with the translation of the label difference into communication problems. Recorded labels allow analyzing how participants chose between the offered labels. It would be necessary to define a procedure of assessing the distance between interpretations of labels.

The labels used in the task are a part of the measurement setting. The label "Moving" allows a rather broad interpretation, while "Walking" calls for a narrower interpretation. Thus, using both creates a task in which the observers respond at a different descriptive level. Further development of the instrument could consider this issue by defining the required level of description and evaluating the relationship between alternatives.

Finally, modal proportions should be understood in the right way. Selected labels mean that the most frequent response for a certain item was chosen in this sample and in this set of labels. This frequency does not serve as an independent criterion of correctness. Thus, present distributions help to find items with competing interpretations and to distinguish the acceptable response from the unique one.

Context and conventional meaning

The distribution of responses is in line with studies confirming that the interpretation of gestures

is guided by contextual information. Lascarides and Stone (2009) illustrate how interconnections between gesture, speech, and discourse facilitate resolving the underspecifications of meanings. The experimental proof of the constraints introduced by the preceding related discourse is provided by Fritz et al. (2021).

The current stimuli include the picture and word alternatives. There was little information regarding an ongoing action, intention of the speaker, or the conversation. It gives rise to various potential functions that may be attributed to the gesture observed. For instance, "Stop" and "Wait" could be applicable to the open palm gesture, suggesting different actions in this case.

Thus, it seems that context may be considered when interpreting the distributions of the responses obtained. For a more direct analysis of the context effect, it is necessary to vary the accompanying situation or discourse. Such variations may be studied based on the findings presented above.

Another perspective is related to conventionality. Gawne and Cooperrider (2024) define emblems as the forms acquiring their meanings through the community practice. Such an approach allows explaining why the recognizable configuration might have multiple interpretations depending either on conventions or on the context. Nonetheless, the current dataset does not identify the cultural or regional origins of the differences found in the recordings. The participants' background, forms depicted, and alternatives used in the task are intertwined.

The issue is especially crucial in regard to sign languages. Studies of iconicity document differences in judgments of resemblance between visual form and meaning across participants with different linguistic backgrounds (Schiefner et al., 2026; Tsang & Perniss, 2026). Thus, the current pictures can be regarded as materials for the interpretation of gestures, not as part of a natural language.

Perceived understanding and agreement on meaning

In addition to the label distribution data, the perceived-understanding measure tracks whether each participant felt he/she understood the content of each image. Across the employment categories, the participants felt that they understood between 7.53 and 9.43 images on average. The summary data for the images include distributions of the number of images chosen by each participant according to the number of labels.

The two sets of summary statistics pertain to different units of analysis. The former set refers to convergence among the viewers of each image. The latter set shows how many images were deemed understandable by the participant. A participant may believe that an image expresses a meaningful action while choosing a label that is different from another participant's choice.

The distinction is important for the assessment of the visual communication materials. A high percentage of "yes" responses shows that the participants believe in their ability to understand the respective images. It would be needed to evaluate whether they actually construct the intended interpretation. This evaluation would require establishing a communicative goal of interest and an indicator of successful performance on it.

Research on metacomprehension adds weight to the need to operationalize subjective judgments carefully. Jaeger and Fiorella (2024) found that the instructional visuals affected both the cues students use to estimate their understanding and the accuracy of their judgments. Although this research relates to learning from text and images, not to labeling gestures, the commonality is in the need to separate subjective judgment from actual performance.

Future research could evaluate the correlation between understanding judgments and accuracy of interpretations. Such research could specify the intended meaning and ask for a yes/no judgment and an action in the defined situation.

Interpreting the employment-group pattern

There was heterogeneity of the reported perceived-understanding count averages among the three employment groups. Higher counts were found among workers and freelancers, and larger dispersions among students. This is how the sample reported its self-assessments under the task conditions.

Participants within each employment category varied in age, education, language, and communicative experience. The average age of students was lower than that of workers and freelancers. Thus, there were limited reasons to explain the observed differences of means in terms of these variables in the unadjusted comparison.

The small freelancer group should also be considered. Its high mean and low standard deviation were obtained close to the upper boundary of the count based on ten images. Such a pattern may emerge if most individuals within the group claim understanding of almost all images. The described

values give a characteristic of such responses; however, estimation of occupational differences requires other samples.

Hearing status and language experience should be measured separately in future studies. Chen Pichler and Koulidobrova (2023) indicate why modality of exposure, experience, and timing of acquisition are important components of language background. The category of hearing status cannot provide this information. In this study, the categories represent sample composition, while the observed employment difference is an exploratory association.

Implications for visual communication and assessment

There are several potential practical implications of these results for the creation of gesture-based materials. First, it is important to assess an image with respect to the communicative purpose behind it. An image that is supposed to ask someone to stop doing something is not necessarily judged by the same criterion as an image meant to communicate a general request for attention. Once the purpose is specified, it will become clear whether people interpret the image with sufficient shared meaning.

Second, alternative responses have to be selected taking into account semantic similarity. The broad and narrow description of an image is equally valid depending on how one chooses to interpret the image. Thus, if a choice is needed for some kind of assessment, the options have to match the interpretive distinction that is made.

Finally, it makes sense to consider perceived understanding as a separate outcome of interest along with a performance criterion. The combination of these measures can allow distinguishing between the material that people find easy to interpret and those that really transmit the intended instruction.

All of these aspects can be considered within language education, where gesture use is integrated into instruction. Karadöller et al. (2024) examined learning with dynamic signs and meanings in terms of their production. Belío-Apaolaza & Hernández Muñoz (2024) looked at instruction through various kinds of gestures and tasks. In both cases, the learning outcomes are associated with the design of instruction and assessment.

This pilot study adds a preliminary stage of material analysis: assessing what meanings people perceive in gesture images and how understandable they find these images. This step is essential before designing the materials that will be used for testing learning or communication processes.

Limitations and future research

This convenience sample was small in size and recruited from one city. Its composition can serve as an exploratory characterization of participants included in the study, whereas generalization will need to be done with a more representative sample. Different employment groups had different sizes and ages, but the mean comparison between these groups was not adjusted for any other characteristics of the participants.

There were ten static pictures in the stimulus set. Live gestures involve movement, time, orientation, and coordination with verbal expression. These properties may affect perception and contribution to meaning of forms (Ebert, 2024; Krason et al., 2022). Current findings pertain to the displayed pictures and response conditions that were specified.

The task was a combination of picture interpretation and labeling in English. The degree of familiarity with the options offered and English language proficiency of participants may have impacted their decisions. Future research will need standardized measurements of their English proficiency, gesture familiarity, and sign language knowledge. Improved documentation of the materials used will improve reproducibility.

Perceived understanding was a task-specific measure that combined the results of ten binary responses. Psychometric properties of this measure need to be established in further studies. The bounded nature of the outcome variable and the unequal sizes of the compared groups make the mean comparison rather unstable. Also, there is no matched analysis that would compare each individual's choice of labels to an independently validated criterion.

Future research can build on current study by systematically normalizing the stimuli, improving response categorization, and using variation in the context. Using dynamic stimuli and communicative scenarios would enable evaluation of contributions of different factors to picture interpretation: visual form, contextual information, and experience of observers. In turn, it would facilitate learning outcome assessment.

Conclusion

For all subjects, the ten images of gestures and hand signals yielded several labels of responses. The most common label constituted 27.5% to 50% of responses. The mean number of responses regarding

perceived understanding was significantly greater for workers and freelancers than students. These findings reveal two different types of outcomes: consensus on a label of response and perceptions of understanding.

This study contributes preliminary findings to guide improvements of gesture-interpretation exercises and visual communication materials. Specifications of context, response options, and communicative purposes may facilitate a better evaluation of interpretation of gesture images by viewers.

Statements and declarations

Funding

No funding was received for this study.

Competing interests

The author declares no competing interests.

Ethics approval

This study was approved by the Ethics Committee of the College of Foreign Languages, Tarim University.

Consent to participate

Written informed consent was obtained from all participants before participation.

Data availability

The aggregate response frequencies and descriptive summaries supporting the findings are included in this article. Participant-level data are available from the corresponding author on reasonable request.

Author contributions

Manuel Alejandro Romero Hung was responsible for the study and manuscript preparation. Yanfei Li was responsible for conceptualization, methodology, formal analysis, supervision, review & editing.

References

Akers, E. M., Midgley, K. J., Holcomb, P. J., & Emmorey, K. (2025). Semantic processing of iconic signs is not automatic: Neural evidence from hearing non-signers. *Bilingualism: Language and Cognition, 28*, 1281–1289. https://doi.org/10.1017/S1366728924001093

Belío-Apaolaza, H. S., & Hernández Muñoz, N. (2024). Emblematic gestures learning in Spanish as L2/FL: Interactions between types of gestures and tasks. *Language Teaching Research, 28*(2), 599–631. https://doi.org/10.1177/13621688211006880

Cairney, B. E., West, S. H., Haebig, E., Cox, C. R., & Lucas, H. D. (2024). Interpretations of meaningful and ambiguous hand gestures in autistic and non-autistic adults: A norming study. *Behavior Research Methods, 56*, 5232–5245. https://doi.org/10.3758/s13428-023-02268-1

Chen Pichler, D., & Koulidobrova, E. (2023). The role of modality in L2 learning: The importance of learners acquiring a second sign language (M2L2 and M1L2 learners). *Language Learning, 73*(S1), 197–233. https://doi.org/10.1111/lang.12607

Ebert, C. (2024). Semantics of gesture. *Annual Review of Linguistics, 10*, 169–189. https://doi.org/10.1146/annurev-linguistics-022421-063057

Fritz, I., Kita, S., Littlemore, J., & Krott, A. (2021). Multimodal language processing: How preceding discourse constrains gesture interpretation and affects gesture integration when gestures do not synchronise with semantic affiliates. *Journal of Memory and Language, 117*, Article 104191. https://doi.org/10.1016/j.jml.2020.104191

Gawne, L., & Cooperrider, K. (2024). Emblems: Meaning at the interface of language and gesture. *Glossa: A Journal of General Linguistics, 9*(1), 1–39. https://doi.org/10.16995/glossa.9705

Jaeger, A. J., & Fiorella, L. (2024). Metacognitive effects of instructional visuals: The role of cue

use and judgment type. *Metacognition and Learning, 19*, 249–291. https://doi.org/10.1007/s11409-023-09370-x

Karadöller, D. Z., Peeters, D., Manhardt, F., Özyürek, A., & Ortega, G. (2024). Iconicity and gesture jointly facilitate learning of second language signs at first exposure in hearing nonsigners. *Language Learning, 74*(4), 781–813. https://doi.org/10.1111/lang.12636

Krason, A., Fenton, R., Varley, R., & Vigliocco, G. (2022). The role of iconic gestures and mouth movements in face-to-face communication. *Psychonomic Bulletin & Review, 29*, 600–612. https://doi.org/10.3758/s13423-021-02009-5

Lascarides, A., & Stone, M. (2009). Discourse coherence and gesture interpretation. *Gesture, 9*(2), 147–180. https://doi.org/10.1075/gest.9.2.01las

Schiefner, A., Roelofsen, F., & Sümer, B. (2026). Iconicity and transparency in Dutch Sign Language. *Behavior Research Methods, 58*, Article 87. https://doi.org/10.3758/s13428-026-02969-3

Tsang, V. W. Y., & Perniss, P. (2026). What's seen as iconic in a sign and what's not? Evidence from iconicity ratings by signers and non-signers on the HKSL lexicon. *Language and Cognition, 18*, Article e53. https://doi.org/10.1017/langcog.2026.10103

How to Cite: Romero Hung, M. A., & Li, Y. (2026). Interpreting Body Gestures and Hand Signals: An Exploratory Study of Selected Meanings and Perceived Understanding. *Contemporary Education and Teaching Research*, 7(11), 381-394.
https://doi.org/ 10.61360/BoniCETR262021001103