What do we express without knowing? Emotion in Gesture

Socially Interactive Agents Track

Gabriel Castillo
University of California, Davis
Davis, California
gcastillo@ucdavis.edu

Michael Neff
University of California, Davis
Davis, California
mpneff@ucdavis.edu

ABSTRACT

Emotional Gestural Expression is an aspect of nonverbal communication that is critical for social agents. The way we perform gestures provides both intentional and unintentional information, leaking our emotional state. In this paper we analyze the perception of a group of base gestures under a wide set of modifications to understand how users perceive the emotional content of the movement, according to the valence and arousal dimensions of Russell's Circumplex model. From these results we are able to extract the perceived emotional quality of the base gesture forms and how the modifications shift that perception. An analysis is provided on the impact of different performance modifications of the gesture.

CCS CONCEPTS

• **Human-centered computing → Gestural input**; *HCI theory, concepts and models*; Virtual reality;

KEYWORDS

Embodied Computational Agents; Emotional Expression; Gesture Analysis

ACM Reference Format:

Gabriel Castillo and Michael Neff. 2019. What do we express without knowing? Emotion in Gesture. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 9 pages.

1 INTRODUCTION

Understanding how people perceive the emotion of another person or agent is a very complex task. There are many channels through which humans express their inner state, such as the words we use, the tone of our voice, facial expressions, posture and arm movements. Much research has been conducted on these varied channels, but relatively little attention has been paid to the role of gesture. This work helps fill that gap.

The development of Embodied Conversational Agents (ECA) has brought new attention to how gesture communicates personality and emotion. The lack of emotional expression is one of the key factors that users notice when interacting with agents and it sabotages their suspension of disbelief. While numerous attempts have been made to address the problem, it remains very challenging to develop effective solutions. This is because emotions are largely performed in an unconscious way and are displayed across a wide range of modalities. In order to progress towards agents with controllable emotional communication, this work examines the impact on the perceived emotion of both a character's gesture choice and the way that gesture is performed.

Russell's Circumplex Model, as shown in Figure 1, is a model used to represent emotions. It maps emotions into a two-dimensional space where Valence (the intrinsic attractiveness or averseness) is represented on the X axis and Arousal (level of stimulation) on the Y axis. In this work, we ask users to rate their perceptions of gesture performances on these two dimensions. The dimensional model is particularly useful for the current study as we are interested in how changes to the performance of a gesture may shift its perception, so we need to represent emotion in a continuous space. Results reveal the emotional quality of particular gestures and how this perception can be changed by altering the gesture's performance.

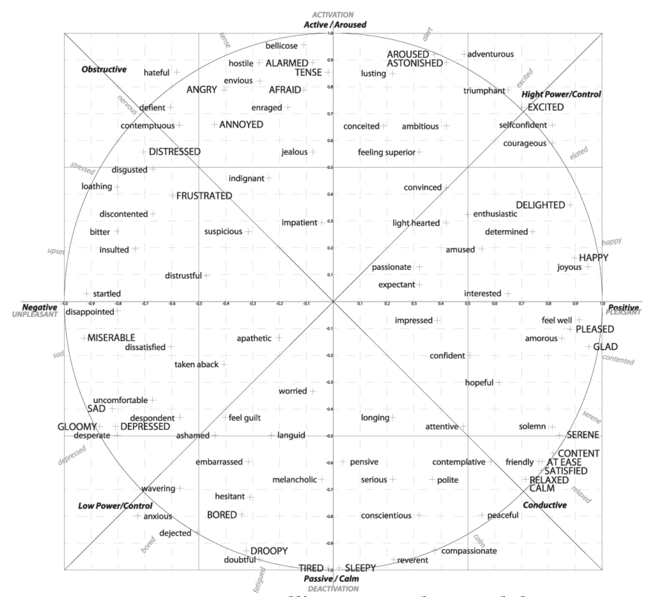


Figure 1: Russell's Circumplex Model

2 BACKGROUND

There have been several attempts to provide emotional expression in agents. Bates [1] suggested that the role of emotional expression is crucial to creating a "believable" agent. MACK ([2]) is an embodied conversational kiosk that builds on this kind of research.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada. © 2019 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

MACK worked in a mixed reality display in a kiosk format. This allowed the agent to display spatial intelligence. In [14], [8] and [7] the authors present GRETA, a general purpose ECA system aimed to create believable agents for interactive applications.

Cassell et al. [3] introduced BEAT, a toolkit that allows one to animate a human-like body using just text as input. It uses linguistic and contextual information contained in the text to control the movements of the hands, arms, face and voice intonation. All the mapping is done by a set of rules derived from nonverbal conversational behavior research. Lhommet et al. [12] describe Cerebella, a system that derives communicative functions from the text and audio of an utterance, by combining all the input available and the agent's current state to produce a communicative function (CF) that is then mapped to multimodal behavior performance.

Normoyle et al. [15] analyze how applying motion editing to some captured gestures alters the emotional content of a motion. They were able to confirm some previous findings, such as that higher velocities and longer amplitudes are related to happiness or anger, while slower joint velocities and smaller amplitudes are related to sadness. They found that after all the edits, most emotions are mainly conveyed through the upper body. They also found that the perceived intensity of an emotion can be reduced by blending with a neutral motion. Finally, they concluded that posture changes can alter the perceived emotion and intensity while changes in dynamics only alter the intensity.

The experiments performed in [5] show the importance of perception in human interaction, and how important facial expressions are to emotion transmission. People's perception of emotion leads them to make inferences about others' mental states.

We are generally not consciously aware of how we use our body in expressing our emotions, and how we read other's emotions. When we see a person expressing something we get its feeling from the whole, not just from one channel or the other. When subtle qualities are present, such as proper hand shape, fluency of movement, energy, etc. we may not think about them, but when they are missing we notice and our perception shifts. Works like Dael et al. [4] have studied the relationship between the emotional dimensions and the elements in the performance. They found that Arousal seems to be a strong determinant in the perception of the emotional expression, possibly even driving the other dimensions. Volkova et al. [18] presented a novel stimuli of intentional emotional performance and evaluated the perception of the gesture in a categorical way. They found that gesture forms' emotional expression is easily recognizable by people. They also identify features such as speed or span that most commonly are related to particular emotions. In Xu et al.[19] they study the valence of some gestures. They allowed the participants to modify the performance of a given gesture so that it expressed a particular valence. Through that, they were able to identify some body features that suggest the given emotion.

In Fourati et al. [6] they present a study over the perception of emotions and body movement across various movement tasks. In their study the user performs 2 tasks that rate the emotion perception and the body cues rating. They identified that while some emotions area properly perceived as what their intention was, some other get confused. They also worked with 8 body cues and found the presence of "emotional related" cues consistent with previous work in the literature. In this paper, we will describe a perceptual study performed to define which elements are most relied upon to evaluate the emotion of a gesture, and how some modification of these elements alter such perceptions. We will analyze each modification and define features that will help us make an informed selection when we want to modify the perception of a given gesture.

3 METHOD

To understand both the impact of gesture form and modifications to gesture performance, we selected 10 representative gesture forms and varied each along 11 different motion attributes, some of the features overlap with the cues used in [6]. In every experiment, we compared two versions of a gesture, either base vs. a modification (e.g., base vs. fist) or two extremes of the same attribute (e.g., high vs. low for *height*) as shown in Table 1. Since *base* was repeated across several factors, there were a total of 16 different clips per gesture.

ID	Factor	Factor A	Factor B
1	Height	High	Low
2	Warp	Ease-in	Ease-out
3	Disruption	Base	Disrupt
4	Lean In	Base	Lean In
5	Length	Long	Short
6	Handedness	Left	Right
7	Extension	Base	Extend
8	Shrug	Base	Shrug
9	Tension	Base	Tense
10	Fist	Base	Fist
11	Hand Spread	Base	Spread

Table 1: Perceptually compared modifications.

3.1 Participants

Participants were recruited through Amazon Mechanical Turk and completed the survey online. Every experiment recruited 40 participants, and they were rewarded with standard compensation for their time when the task concluded. The total number of participants was 880. Due to the subjective nature of the questions asked, it is difficult to identify participants that cheated or answered in a malicious way. We, therefore, decided to use a Pearson's correlation coefficient between individual's scores and mean video scores to filter outlier participants as likely dishonest survey takers, as described in [16, 17]. We rejected participants with a correlation with the mean below 0.15 and our rejection rate in the different experiments was between 5 -10%.

3.2 Material

For stimuli, we selected gestures from a database of over 250 motion captured gestures used for animating conversational characters [9]. The database includes arm, body and finger data, although the body data was not used in this study. It contains a wide range of variants on the gesture dimensions described by McNeill [13]: metaphoric, iconic, deitic and beat gestures. The gestures were classified and labeled based on the performance of the stroke, with labels derived from the survey by Kipp [10]. Our goal was to select gestures that reasonably spanned the variation in the database. Since numeric

Group	ID	Label	Description
A	CP	Cup	Starts with hands facing in above shoulder level and moves outwards in the same position in a short trajectory that ends at shoulder level.
	CX	Cup Exemplar	Starts with open hands palms facing up above the waist level and moves forward in an offering motion to end up just below shoulder level.
	CH	Cup Horizontal	Starts with open hands palms facing up above the waist level and moves sideways to end up just below shoulder level.
	PA	Pointing Abstract	Starts with both hands facing each other at waist level and moves forward and up to end with arms straight at shoulder level and palms still facing each other.
	RJ	Reject	Starts with hands overlapped at waist level with both palms facing inwards and moves out to the sides to end up with palms facing forward at shoulder level.
B	AW	Away	Starts at chest level and moves out and up with palms facing front.
	CM	Calm	Starts below chest level with palms facing forward and moves front in a push motion.
	CS	Cup Short	Starts with Palms facing each other at the mid-torso level and then hands are moved slightly outwards rotating the palms to end up facing up.
	DS	Dismiss	Starts with hands facing forward on the sides above shoulder level and moves in an inwards diagonal trajectory to end up with hands overlayed at waist level with palms facing inwards.
	SP	Short Progressive	Starts with palms facing inwards at waist level and moves outwards to the front in a circular motion to end up with the hands facing up in front just below shoulder level.

Table 2: Gestures used in the study were split into two groups, A and B.

measures of motion such as mean square error are notoriously poor at correctly capturing semantic differences, we instead rated gestures on a set of traits including height, speed, acceleration, direction, hand shape, handedness and stroke length, among some others. We then selected a final set of gestures that spanned a wide range of the variation seen across these features. The selected gestures are described in Table 3.3 and available in video form through the link listed below.

The final stimuli consisted of 4-5 seconds clips of complete gestures using the stroke motion data. In every clip, a modification was performed on the stroke or not, depending on what was being measured (Table 1). All the stimuli can be viewed at http://104.131.114.67/

3.3 Procedure

Ten clips were shown to the participants in every experiment in random order. Every set consisted of 5 different gesture forms, either Set A or Set B (Table). A total of 11 different modifications were evaluated, as described in Table 1. This led to 22 total experiments (11 performance modifications x 2 gesture sets). The participants were asked if the gesture conveyed any emotion in a free text field. Then in a 7-point Likert scale, they were asked to rate the valence, arousal and naturalness they perceived in the performance. Finally they were presented with 8 labels of emotions and were asked what emotion they would relate the gesture to. The free text and label data is not included in this paper. For the Likert style questions, participants were able to enter values on a pseudo-continuous scale from 10 to 70 using a slider.

3.4 Data Analysis

After the initial filtering of the data described in Sec. 3.1, it was analyzed using R. We performed within subjects 2-way ANOVAs on the results from each experiment, where the factors were *gesture form* and *performance modification*. False Discovery correction was performed to mitigate the risk of false positives given the repeated tests. Post-hoc analysis was performed using Tukey HSD.

4 RESULTS

Results for the 22 experiments are presented below, looking separately at the ratings for valence, arousal, and naturalness.

4.1 Valence

Modification	F-value	p-value	Lower Valence
Warp	F_A=5.516	p_A=0.0243	Ease-in
Disruption	F_A=7.414 F_B=4.346	p_A=0.0129 p_B=0.0439	Disrupt
Handedness	F_A=4.522 F_B=4.998	p_A=0.0536 p_B=0.0312	Left
Tension	F_A=6.519 F_B=5.137	p_A=0.0195 p_B=0.0385	Tense
Fist	F_A=7.196	p_A=0.0109	Fist

Table 3: Experiments where performance modifications had a significant main effect on Valence.

Not surprisingly, many experiments showed a main effect of the gesture form. Ranking the overall scores for each gesture form, including all modifications, shows that Away (average rating = 26.0) and Reject (26.8) were viewed more negatively, Pointing Abstract (41.9) was viewed most positively, and the remaining gestures were rated as follows : Dismiss (31.1), Cup Horizontal (32.3), Cup (32.4), Calm (33.8), Cup Exemplar (35.5), Short Progressive (36.1), Cup Short (36.4).

The cases where performance modifications had a significant main effect on valence are summarized in Table 3. Warp showed significance for Set A, Ease-in being more negative. Disruption showed significance for both sets with disruption lowering perceived valence. Handedness also showed significance for both Sets, with Left being seen as more negative than right; this was also reported in [11]. There is also a main effect for Tension, where higher tension appears to lower perceived valence. Finally, Fist shows significance against Set A, lowering the perceived valence.

The performance by gesture form interactions are shown in Figure 2. The interaction between Height and Set A was significant; Figure 2(a) shows lower gestures were generally seen as more positive, but this relationship flipped for Cup Horizontal. The Warp interactions with Sets A and B Fig. 2(b & c) show that perceived valence is only lowered by Ease-in for certain gesture forms, and in one case, Dismiss, it is actually significantly raised. Disruption presents a significant interaction for Set B, Fig. 2(d) shows that Disruption significantly lowered perception of Short Progressive, but was more neutral for other cases and showed an increase for Dismiss. For Length, only the interaction on Set A was tending towards significance and Fig. 2(e) shows that shorter gestures generally tend to be slightly more positively perceived, but the opposite trend occurs for Cup Exemplar. Handedness also shows an interaction for Set B (Fig. 2(f)). While Away is significantly more negative when performed with the left hand, it is more neutral for other gestures and made slightly more positive for Dismiss. Fist also shows interactions for both Sets. While Pointing Abstract is significantly more negative when done with a Fist, this effect tends to vary for other gestures (Figures 2(g & h)).

4.2 Arousal

In every experiment, there was a significant main effect of gesture form on perceived arousal. To understand this variation, we again ranked gestures based on their average perceptual ratings, including base and modified forms. Set A Reject (average rating = 46.9) showed the highest arousal, with Away (45.2) quite close. These are followed by Pointing Abstract (43.0) and Short Progressive (42.6). Cup Short (30.2) is perceived as very low energy. The remaining gestures lie in-between: Cup Exemplar (34.0), Calm (34.7), Cup Horizontal (36.1), Dismiss (37.7), Cup (39.2).

Modification	F-value	p-value	Higher Arousal
Height	F_A=12.771 F_B=5.885	p_A=0.00130 p_B=0.02670	High
Disruption	F_A=5.352 F_B= 23.634	p_A=0.03490 p_B=0.00003	Disrupt
Length	F_A=13.641	p_A=0.00154	Long
Handedness	F_B=11.497	p_B=0.00214	Left
Fist	F_A=5.775	p_A=0.02850	Fist

Table 4: Factors that significantly impact the Arousal

Table 4 summarizes the significant main effects of performance modifications on perceived Arousal. Gestures that are higher, longer, or use a fist are all seen as having higher arousal in at least one of the two gesture sets and they show no significant interactions for these modifications. Disruption has a significant main effect for both gesture sets where the presence of disruption was perceived as showing greater arousal. There was a main effect of Handedness on Set B; it suggests left-handed gestures are seen as showing higher arousal.

Significant interactions of the performance modifications x gesture form are shown in Figure 3. There is an interaction with Disruption for Set A. Figure 3(b) shows that the general trend holds of disruption raising arousal, except disrupted Cup Short was perceived as less aroused. Differences are significant for Dismiss. Handedness show a significant interaction. The interaction supports that left-handed gestures are seen as showing higher arousal although this only holds with certain gesture forms (Fig. 3(c)), being neutral for others. It is significant for Away.

Warp, suggests a significant interaction for Set B. Fig. 3(a) shows that Ease-in was generally perceived as showing higher arousal, but this effect flipped for Dismiss. Similarly, Shrug showed a significant interaction for Set A. Adding a shrug raised perceived arousal for some forms and lowered it for others (Fig. 3(d)).

4.3 Naturalness

In this work, naturalness is used as a control measure in order to ensure that the performance modifications do not destroy the realism of a gesture. Participants evaluated how natural or awkward a gesture appeared. Overall, naturalness levels were consistent across stimuli. The average naturalness rating for all stimuli was 46.55. It was 46.98 for the unmodified gestures and 46.35 for the modified. The one outlier is the Fist modification, which had an average naturalness of 36.00. Table 5 shows which edits significantly impacted naturalness and Fist does create a significant reduction. This may limit the usefulness of this edit in practice. This also presents a good example of how some gestures, even if performed in a physically valid way, may look abnormal to many people. The average naturalness of all the non-Fist modifications is 47.09. Aside from the Fist set, all gestures had naturalness ratings between 42.4 and 51.4. This is an acceptable range, so the modifications should be viable in practice.

It is interesting to note in Table 5 that Naturalness was significantly impacted for Height in Set A (Low mean = 46.4 and High mean = 43.0) and Handedness for Set B (Left mean = 50.19 and Right mean = 45.6). Overall, though, these naturalness values seem reasonable so do not present a serious concern.

Modification	F-value	p-value	Higher Naturalness
Fist	F_A=40.171 F_b=35.812	p_A=0.00000 p_B=0.00000	Base
Height	F_A=7.379	p_A=0.01980	Low
Handedness	F_b=24.049	p_B=0.00003	Left

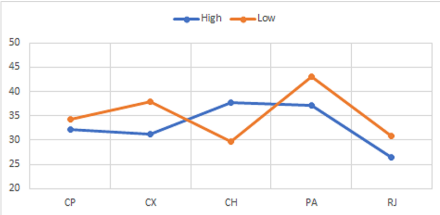
Table 5: Significant main effects of performance modification on Naturalness ratings.

5 DISCUSSION

In order to gain a deeper understanding of the relative contribution of gesture form and performance modifications, along with understanding how malleable different gesture forms are in terms of emotion, we can plot the rating data on a 2D space where the x axis represents the Valence and the y axis is the arousal, similar

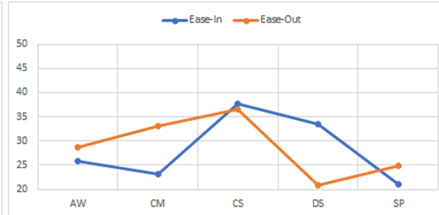
[1]**Gestures::** CP: Cup, CX: Cup Exemplar, CH: Cup Horizontal, PA: Pointing Abstract, RJ: Reject, AW: Away, CM: Calm, CS: Cup Short, DS: Dismiss, SP: Short Progressive
Modifications:: H: High, W: Low, F: Ease-in, S: Ease-out, G: Long, C: Short, L: Left, R: Right, B: Base, T: Tense, P: Fist, A: Spread, N: Lean In, E: Shrug, D: Disrupt, X: Extend

Figure 2: Significant performance x gesture form interactions for valence[1] with ANOVA values.

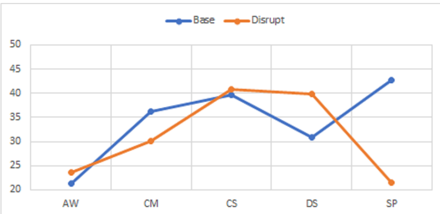


(a) Height A (F=3.9, p=0.0065)

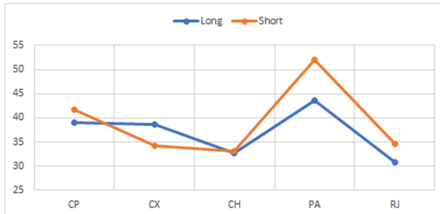
(b) Warp A (F=3.4 p=0.016)



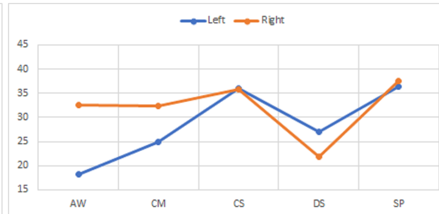
(c) Warp B (F=7.5, p<0.0001)



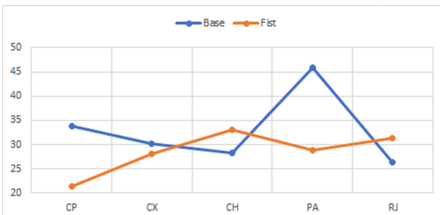
(d) Disruption B (F=14.9, p<0.0001)



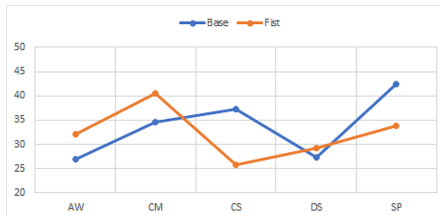
(e) Length A (F=2.5, p=0.065)



(f) Handedness B (F=5.7, p=0.00035)

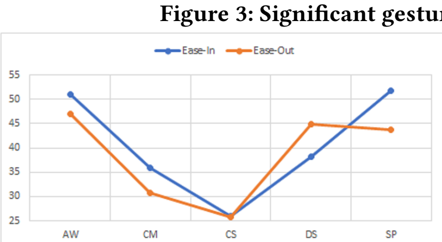


(g) Fist A (F=6.8, p<0.0001)

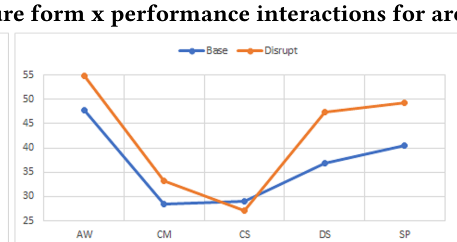


(h) Fist B (F=5.3, p=0.0011)

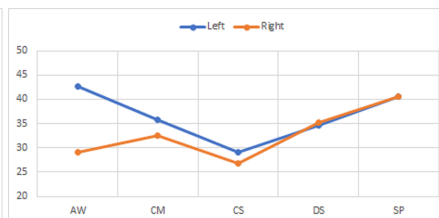
Figure 3: Significant gesture form x performance interactions for arousal[1] with ANOVA values.



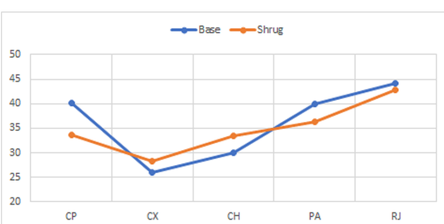
(a) Warp B (F=5.0, p=0.0011)



(b) Disruption B (F=3.2, p=0.014)



(c) Handedness B (F=4.04, p=0.0038)



(d) Shrug A (F=2.7, p=0.043)

to Russell's Model. The *origin* of this space is the point where the perception is neither positive nor negative and the arousal is between high and low. We calculate a location for the gesture based on the average ratings collected in the experiments. We explore how every modification can shift the perception and propose a simple relationship between gesture and performance modifications that should modify the gesture perception.

5.1 Gestures

The Circumplex model does not yet have an established way of stating what Valence and Arousal values are needed to describe a particular emotion. Most theories suggest the emotional space is a continuum and that emotions transition into others, but the boundaries are too fuzzy to clearly state a particular emotion's exact location. Therefore a one-to-one mapping is difficult to assign.

Figure 4 shows us where the base gestures are located in the two-dimensional space of the ratings. We will be using a simple distance measure between the location of the gestures in Figure 4 and the location of the modified gesture in the same plane, as an analysis tool to evaluate how the perception of the gesture shifts

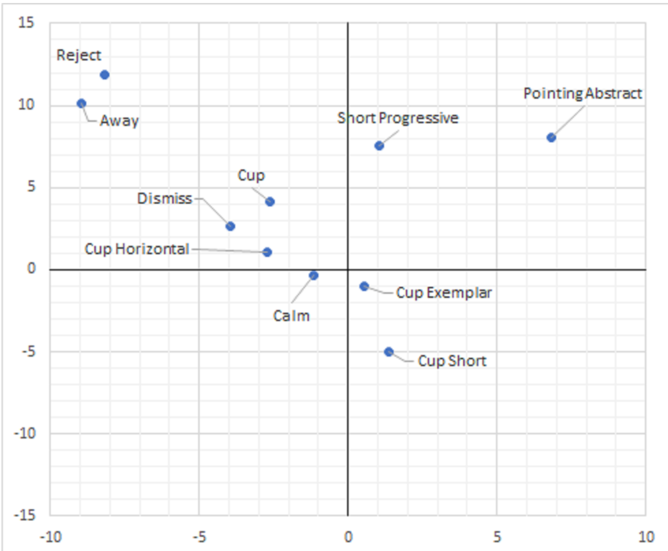


Figure 4: Plot of average Valence and Arousal gesture ratings.

when modified. The average distance the gesture perception shifts over all the modifications can tell us how variable the gesture form is in terms of our space. We call this average the **gesture malleability**. The smaller this value is, the less variation is expected of the Valence-Arousal relation in the gesture form. On the other hand, if the value is bigger, we can say that performance modifications are likely to alter the perceived quality of the gesture. In the next several subsections we will analyze in detail the values each gesture registered and how their perception was changed when they were modified.

5.2 Modifications

This section will analyze how consistent the impact of each modification is as it is applied to different gesture forms. We were able to identify the impacts of each movement factor. Taking the location of the base gesture as origin, we map the direction towards which the modification adjusted the Valence/Arousal values, calling these vectors the *constituent vectors*. With this collection of vectors, we are able to define the **effect vector** as an average of all the constituent vectors. This vector will be an aid when defining what modification is better suited to modify a gesture when we want to change its perception. For instance, if I am given a gesture that maps to the second quadrant close to the origin, what modification could be applied to shift it towards the top left of the quadrant without changing the base gesture?

The *effect vector* will be the general direction we expect the emotion conveyed to move within our model, though with an awareness that there might be expected deviations where the emotion could end up. To account for that, we define the **modification spectrum** as the arc described by the angle between two standard deviations of the *effect vector* and its *constituent vectors*. The length of the arc will tell us how consistent the modification is; the shorter the arc, the closer changes will be to the effect vector. There will be cases where the *constituent vectors* may be going in opposite directions, thus making the *modification spectrum* misleading; to check the consistency of the spectrum we will define the **spectrum containment** as the percentage of *constituent vectors* that fall within the *modification spectrum*. The combination of all these factors will give us a better idea on how this modification behaves.

The length of the acting *effect vector* would be dependent on the interaction of the modification and the gesture. We can calculate the estimated shift of the modification by getting a simple average of the length of the *constituent vectors*; we will call this the **modification strength**. The smaller this value is, the weaker the expected effect will be.

All these defined values help describe and classify the manner in which modifications interact with gesture forms to impact the perceived emotion. Figure 5 summarizes this data graphically, visualizing both the type of perceptual changes possible and their consistency. This serves as both an analysis tool and a potential set of heuristics for modifying agent behavior to achieve particular perceptions.

Table 6 shows the modifications that most affected the gestures in the study and showed statistical significance. It shows the base perception of the gesture, how it was shifted and what the new perception is. In Table 7 we summarize the *effect vector*, *modification spectrum*, *containment* and the *modification strength* of the modifications that were determined in this study. From these observations, we can conclude, for instance, that the strength of the Height factor is rather low, but the largely opposing directions of the High and Low adjustment do provide an intuitive relationship between both ends of the scale.

The Handedness adjustment shows an increase in negativity for gestures performed with the left hand, providing further evidence of something previously noted in an observational study of actor performances [11]. This adjustment has high Strength, and although the containment in Right drops, their modifications are noticeably different than the ones from Left.

Shortening the length of the stroke shows an interesting interaction with gesture form. The impact is largely to change the valence, but it can do so in both directions depending on the gesture form, with changes running bidirectionally along the effect vector. We refer to this type of modification as **split spectrum**. The Warp effect is another case of a *split spectrum* modification, where the effect vector describes an axis rather than a direction.

Tension, Lean In and Extend are the strongest independent modifications found in this work, not only considering their strength but also their spectrum and containment. Fist is the strongest modification based on the strength value, but with the caveat that it lowers the evaluation of Naturalness.

5.3 What carries meaning?

This study reveals that in terms of conveying emotion, the relationship between gesture form and gesture performance is complicated. In some cases, such as the impact of height, stroke length or the use of fist on Arousal, modifications act in a consistent manner. In other cases, there are interactions between performance modifications and gesture form. It is clear, at this point, that the gesture form sets an initial point in emotional space, but the impact of performance

Gesture	Modification	Δ	Valence	Arousal
Away Malleability: 5.38 (-0.90,1.01)	Left	0.82	-1.68	0.77
	Right	1.73	-0.25	-0.59
	Disrupt	0.99	-1.14	1.99
	Fist*	0.73	-0.29	0.61
	Ease-Out	0.32	-0.64	1.20
	Ease-In	0.59	-0.92	1.61
Calm Malleability: 5.40 (-0.12,-0.03)	Ease-In	1.07	-1.18	0.08
	Ease-Out	0.41	-0.19	-0.43
	Left	0.90	-1.01	0.07
	Right	0.27	-0.27	-0.26
	Fist*	0.69	0.56	0.09
Cup Malleability: 6.27 (-0.26,0.42)	Fist*	1.21	-1.36	0.93
	Short	0.93	0.66	0.50
	Long	0.78	0.40	0.83
	Ease-Out	0.82	0.47	0.80
	Ease-In	0.56	0.27	0.60
	High	0.18	-0.28	0.59
	Low	0.65	-0.08	-0.21
Cup Horizontal Malleability: 4.79 (-0.27,0.11)	Long	0.94	-0.23	1.05
	Short	0.14	-0.19	0.23
	Shrug	0.72	-0.94	-0.16
	Ease-Out	0.57	0.23	0.38
	Ease-In	0.40	-0.63	0.30
	High	0.75	0.26	0.64
	Low	0.51	-0.52	-0.34
Cup Short Malleability: 3.88 (0.14,-0.50)	Fist*	1.09	-0.91	-0.20
	Disruption	0.52	0.57	-0.78
	Ease-Out	0.43	0.16	-0.92
	Ease-In	0.43	0.27	-0.91
	Right	0.33	0.09	-0.82
	Left	0.11	0.10	-0.60

Gesture	Modification	Δ	Valence	Arousal
Cup Exemplar Malleability: 5.48 (0.05,-0.10)	Long	0.99	0.37	0.84
	Short	0.27	-0.08	0.14
	Fist*	0.91	-0.70	0.42
	Shrug	0.65	0.36	-0.67
Dismiss Malleability: 6.10 (-0.40,0.27)	Disruption	1.31	-0.49	1.23
	Ease-Out	1.24	-1.40	1.00
	Ease-In	0.23	-0.15	0.33
	Right	0.96	-1.32	0.02
	Left	0.52	-0.81	-0.05
	Fist*	0.21	-0.58	0.37
Pointing Abstract Malleability: 6.39 (0.69,0.80)	Fist*	1.30	-0.62	0.72
	Short	1.03	1.70	1.00
	Long	0.75	0.85	1.54
	Ease-Out	0.69	1.36	0.92
	Ease-In	0.24	0.46	0.90
	Shrug	0.68	0.77	0.13
	High	0.54	0.22	1.10
	Low	0.12	0.80	0.83
Reject Malleability: 6.93 (-0.82,1.12)	Ease-In	1.01	0.18	1.31
	Ease-Out	0.84	0.02	1.13
	Short	0.80	-0.03	1.02
	Long	0.39	-0.42	1.24
	Low	0.41	-0.42	1.31
	High	0.12	-0.85	1.53
	Fist*	0.45	-0.37	1.22
Short Progressive Malleability: 6.42 (0.10,0.76)	Ease-In	1.76	-1.39	1.67
	Ease-Out	1.13	-1.02	0.87
	Disrupt	1.61	-1.36	1.43
	Fist*	0.27	-0.11	0.59
	Left	0.21	0.14	0.55
	Right	0.24	0.25	0.56

Table 6: Perception of Gestures after Modification

Modification	Effect Vector		Spectrum (°)	Containment	Strength
	Valence	Arousal			
High	-0.14	0.99	179	0.8	3.96
Low	0.75	-0.66	155	0.7	3.24
Left	-0.88	-0.48	149	0.8	5.86
Right	-0.01	-1.00	127	0.6	6.97
Long	0.32	0.95	232	0.7	6.49
Short	0.96	0.26	275	0.6	5.90
Ease-in	-0.56	0.83	202	0.6	6.71
Ease-out	0.61	0.79	218	0.6	6.87
Tense	-0.86	-0.51	166	0.9	5.41
Fist	-0.93	0.36	256	0.7	7.03
Spread	-0.57	0.82	212	0.7	3.84
Lean In	0.90	0.43	60	0.8	6.07
Shrug	-0.26	-0.97	218	0.8	5.14
Disrupt	-0.48	0.88	160	0.7	7.31
Extended	0.52	-0.85	94	0.6	4.77

Table 7: Values of features by modification

changes do indicate that the perception can be modified, sometimes in very important ways. Some gesture forms are less emotionally malleable, while other can be changed substantially depending on how their performance is altered.

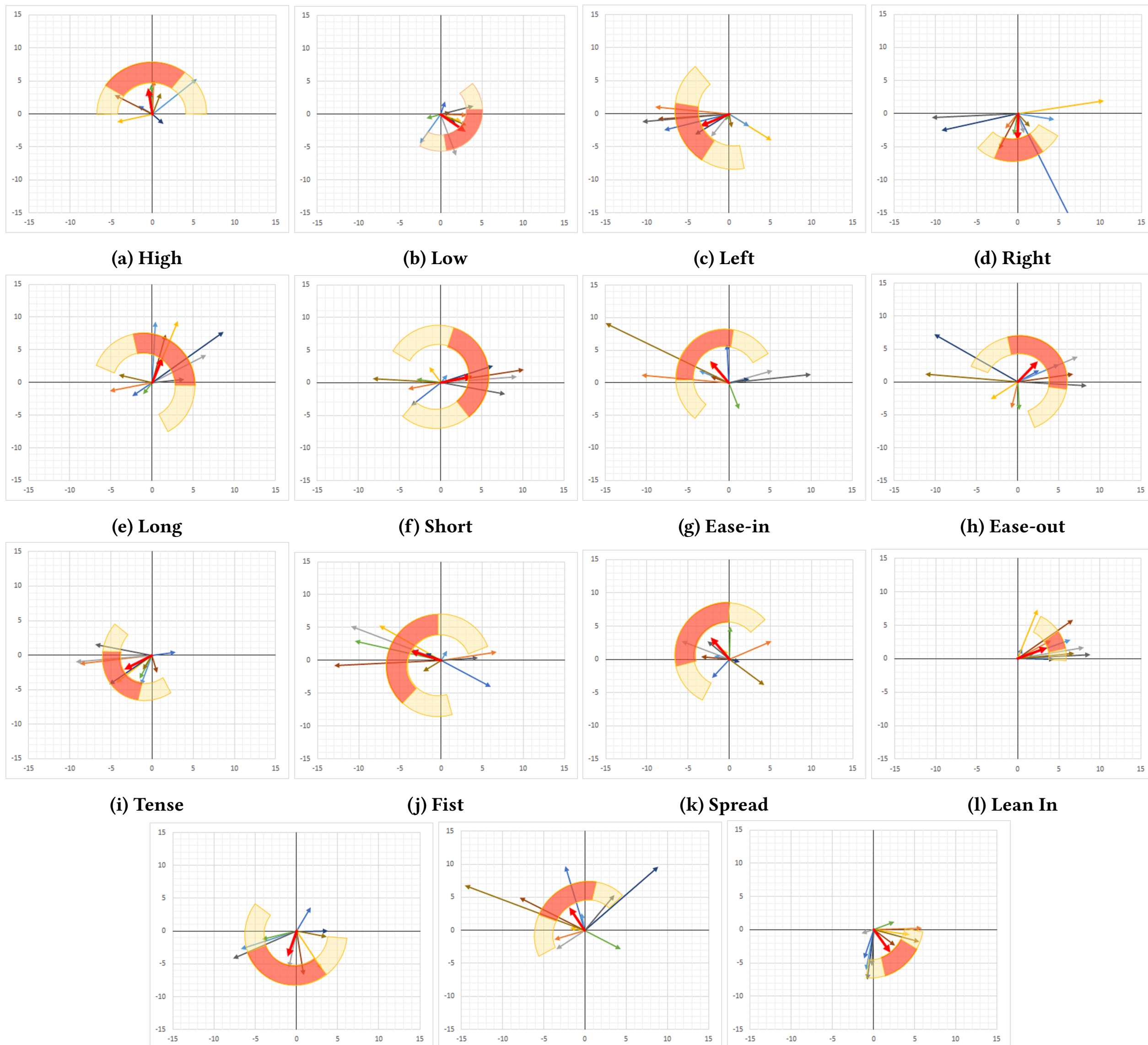
The interactions between form and performance suggest that performance modifications on their own cannot guarantee a particular meaning, but in this study, we have seen that some of them do indicate a very strong consistency in the direction and distance they can shift any given gesture. Other modifications, although not as strong, do give at least a good probability to predict where the gesture will shift with the modification. The third group of modifications does not show any evident pattern and may benefit from further observational data.

6 FUTURE WORK

A number of issues related to the impact of performance modifications would benefit from further evaluation. For instance, can the combination of modifications drive a gesture to a desired perception? What is the best way of looking at the wide spectrum modifications that could allow us to use them in a more informed way? How do the modifications affect the perception when the modifications are used in a series of gestures? Is the perception reinforced or diluted? Are there complementary modifications?

There is also more to learn about the role of gesture form. How is it that the interaction of the modifications and the gesture really

Figure 5: Modification graph



(a) High (b) Low (c) Left (d) Right (e) Long (f) Short (g) Ease-in (h) Ease-out (i) Tense (j) Fist (k) Spread (l) Lean In (m) Shrug (n) Disrupt (o) Extend

affect the perception? Our data suggest that there are indeed very malleable gestures, but are there any gestures that are not malleable at all? Can we better classify the gestures in a manner that will lead to a more deterministic way of implementing modifications?

The relationship between form and performance is quite complex and there is still a lot to be observed and measured. The set of the gestures in this study was limited, but it can be easily expanded and the experiments replicated in order to strengthen the conclusions observed in this work.

ACKNOWLEDGMENTS

The work is supported by the UC MEXUS - CONACYT under Grant No.: 220351

REFERENCES

[1] Joseph Bates. 1994. The role of emotion in believable agents. *Commun. ACM* 37, 7 (1994), 122–125.

[2] Justine Cassell, Tom Stocky, Tim Bickmore, Yang Gao, Yukiko Nakano, Kimiko Ryokai, Dona Tversky, Catherine Vaucelle, and Hannes Vilhjálmsson. 2002. Mack: Media lab autonomous conversational kiosk. In *Proc. of Imagina*, Vol. 2. 12–15.

[3] Justine Cassell, Hannes Högni Vilhjálmsson, and Timothy Bickmore. 2004. Beat: the behavior expression animation toolkit. In *Life-Like Characters*. Springer, 163–185.

[4] Nele Dael, Martijn Goudbeek, and K R Sherer. 2013. Perceived gesture dynamics in nonverbal expression of emotion. *Perception* 42 (2013), 642–657.

[5] Celso M de Melo, Peter J Carnevale, Stephen J Read, and Jonathan Gratch. 2014. Reading people's minds from emotion expressions in interdependent decision making. *Journal of personality and social psychology* 106, 1 (2014), 73.

[6] Nesrine Fourati and Catherine Pelachaud. 2018. Perception of emotions and body movement in the emilya database. *IEEE Transactions on Affective Computing* 9, 1 (2018), 90–101.

[7] Björn Hartmann, Maurizio Mancini, Stéphanie Buisine, and Catherine Pelachaud. 2005. Design and evaluation of expressive gesture synthesis for embodied conversational agents. In *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*. ACM, 1095–1096.

[8] Björn Hartmann, Maurizio Mancini, and Catherine Pelachaud. 2005. Implementing expressive gesture synthesis for embodied conversational agents. In *International Gesture Workshop*. Springer, 188–199.

[9] Zhichao Hu, Michelle Dick, Chung-Ning Chang, Kevin Bowden, Michael Neff, Jean E Fox Tree, and Marilyn A Walker. 2016. A Corpus of Gesture-Annotated Dialogues for Monologue-to-Dialogue Generation from Personal Narratives.. In *LREC*.

[10] Michael Kipp. 2005. *Gesture generation by imitation: From human behavior to computer character animation*. Universal-Publishers.

[11] Michael Kipp and Jean-Claude Martin. 2009. Gesture and emotion: Can basic gestural form features discriminate emotions?. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*. IEEE, 1–8.

[12] Margaux Lhommet and Stacy C Marsella. 2013. Gesture with meaning. In *International Workshop on Intelligent Virtual Agents*. Springer, 303–312.

[13] David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.

[14] Radoslaw Niewiadomski, Elisabetta Bevacqua, Maurizio Mancini, and Catherine Pelachaud. 2009. Greta: an interactive expressive ECA system. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 1399–1400.

[15] Aline Normoyle, Fannie Liu, Mubbasir Kapadia, Norman I Badler, and Sophie Jörg. 2013. The effect of posture and dynamics on the perception of emotion. In *Proceedings of the ACM Symposium on Applied Perception*. ACM, 91–98.

[16] Harrison Jesse Smith and Michael Neff. 2017. Understanding the impact of animated gesture performance on personality perceptions. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 49.

[17] Jon Sprouse. 2011. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior research methods* 43, 1 (2011), 155–167.

[18] Ekaterina P. Volkova, Betty J. Mohler, Trevor J. Dodds, Tesch Joachim, and Heinrich H. Bulthoff. 2014. Emotion categorization of body expressions in narrative scenarios. *Frontiers in Psychology* 5 (2014).

[19] Junchao Xu, Joost Broekens, Koen Hindriks, and Mark A Neerincx. 2013. *Mood expression through parameterized functional behavior of robots*. IEEE.