Impacts of Culture on Gesture Based Interfaces:
A Case Study on Anglo-Celtics and Latin Americans

Manolya Kavakli and Karime Nasser
Department of Computing
Macquarie University
Sydney, Australia
Manolya.kavakli@mq.edu.au, Karimenasser@hotmail.com

Abstract- **This paper investigates the impacts of culture in gesture-based interface design. The aim is to identify basic cultural differences in using hand gestures between two cultures: Anglo-Celtics and Latin Americans. We videotaped and analyzed 10 participants in two groups, while describing two chairs by using their speech and hand gestures. We investigated the frequency, occurrence, and the type of gestures used by the two cultures, as well as the words most frequently used by the participants. Our findings state that there are cultural differences in hand gestures during the description of the objects. This may have implications for the development of gesture-based multimodal interfaces. Anglo-Celtics coming from a low context culture, describe objects using a larger number of gestures as well as a larger vocabulary in a longer period. On the other hand, Latin Americans coming from a high context culture, use smaller number of gestures more frequently in a shorter period. The differences in frequency of gestures may have an impact on the adoption of new technologies as well as adaptation to them. We also found that as the complexity of a task increases, so does the number and type of gestures used. Our conclusion is that the gesture vocabulary of a multimodal interface will be affected not only by the complexity of the task being performed, but also by the cultural background and the language skills of the user.**

Keywords- Gesture recognition; Cultural difference; Gesture based interface design; Gesture segmentation; Speech coding.

I. INTRODUCTION

This study builds on and extends our previous research on the "Influence of culture in multimodal interfaces" [1] and as its predecessor, aims at defining the variances in gesture behavior from one culture to another. Culture can be defined as the shared patterns of behaviors and interactions, cognitive constructs, and affective understanding that are learned through a process of socialization. These shared patterns identify the members of a cultural group while also distinguishing those of another group [2]. Our goal in this paper is to identify the impacts of culture on the frequency, occurrence, and types of hand gestures in multimodal interfaces.

Bischel et al. [3] investigated a designer describing a mechanical device to another designer. Similar to their experiments, we videotaped and analyzed the gestures of the participants from two different cultures describing two different chairs using their speech and hands. We segmented the video records using timestamps, and analyzed them using metrics such as frequency, occurrence, and the quantity of certain gesture types.

The paper is structured as follows: Section II provides a review of related literature on gesture and culture. Section III presents the experiment conducted. Section IV analyzes the data collected, and the results of frequency, occurrence and type of gesture amongst the samples. Finally, Section V discusses the experiment results and relation with culture studies, and Section VI presents our conclusions drawn from the results.

II. LITERATURE REVIEW

The aim of Human Computer Interaction (HCI) is making interactions as natural as possible, as if communicating with another human [4]. Gestures are used as a way of expression in interaction, either with or without a human. There is evidence that gestures convey information *redundant* to the information conveyed in speech [5] and gesture is a precursor to speech. Humans have an innate need to use gestures; since they complement our ideas, to such an extent that humans are known to gesture even when talking on the phone [6].

The means to communicate with computers have evolved from classic mouse input, to rich multimodal data [7]. Multimodal interfaces, that use multimodal data, have combined various user input modes beyond the known keyboard and mouse input/output [8]. They now include a wide range of interaction methods; such as hand gestures (both static and dynamic), as well as speech, and head and eye tracking. Most recently, Razer™, the world leader in high-performance gaming hardware, launched an adaptive tactile keyboard with a switchblade user interface, suggesting only the recurrent need to provide players with more options. Games and interactive entertainment industries are not, however, the only application areas for gesture based interfaces. More serious applications exist, such as The Intuitive da Vinci surgical system that is used for the capture of subtle motions of a surgeon, to teach novices complex procedures [9]. One may assume that in tasks such as the manipulation of objects, cultural implications might not be of considerable importance, but in the context of cultural and physical differences between surgeons, the topic calls for more attention [10].

Gesture-based interfaces enable freer, more intuitive, and richer digital interactions than conventional user interfaces [11]. These enhanced interactions lead to better idea generation [12]. With the decrease in the price of sensors and the growth of processor capacity, an interesting step in multimodal interface design is the creation of natural and invisible interfaces that are called ambient gestures [13]. These invisible interfaces are designed to support ubiquitous interactions with everyday computing technology. Such examples include lowering the volume of a stereo from afar using a hand gesture. This way the user does not have to leave their original activity to perform an action adding the value of ubiquity. Through appropriate gestures, these technologies allow immersion, navigation and interaction to support idea generation [14].

A. Gesture classification and segmentation

Many gesture recognition systems have been technology driven and address gesture tracking needs, rather than the requirements of human behavior [15].

When developing gesture based interfaces, programmers and designers work together to understand what types of gestures are most frequently used for interaction. Therefore, there have been many attempts to design appropriate gesture classification and segmentation "dictionaries". However, none of them provide a complete guide for what gestures are mostly used and by what group of users.

Gestures offer versatility when representing objects, or qualities of these in the scientific domain. The main problem here is that there is no common database of gestures that can be used by both developers and designers. At the top level, they are divided into representational and non-representational gestures. Representational gestures represent physical things being conceptualized by the person gesturing in an abstract or physical manner. Gesture also allows representation of movement through imitation [6]. The non-representational gestures, accompany or stress *speech*, and also involve space relations, like pointing. In a way, they are culture-specific emblems, but *how* culture-specific are gestures remains as a question to explore.

Gestures have also been classified according to their purpose. They could be goal oriented (change of position, shape), empty handed (wave, snap, point, take), for indirect manipulation (set, stop) or haptic exploration (touch, stoke, knock) [14]. The most recognized gesture classification is the one established by McNeill in 1992 [16]. McNeill classifies 4 types of gestures (See Table 1): iconic (resemble what is being talked about, e.g., flapping arms when mentioning a bird), metaphoric (abstractedly pictorial, e.g., drawing a box shape when referring to a room), beat (gestures that index a word of phrase e.g., rhythmic arm movement used to add emphasis), and deictic (gestures pointing to something, e.g., while giving directions).

The iconic gestures are of particular interest to HCI as they allow accurate depiction of objects encountered by the user. An important issue here is the presence of transition movements, or junk gestures, as these are classified as meaningless, since they do not convey information.

When training a model for gesture recognition, it must be done by segmenting the individual gestures and then interpreting their representation individually or as a whole, depending on the model. A common motion segmentation technique is using distance signals to determine the type of gesture on the basis of the contour of the gesturing person's body or hand [17]. Li and Greenspan in [18] focus on how the endpoints are located. In order to do this, they had participants repeat various actions several times in order to document the variances. These variances, they claim, are useful for identifying the range of a given gesture, and therefore, provide a better identification.

The most common errors in gesture segmentation are mainly classified as three types [18]: substitution errors, deletion errors and insertion errors. Substitution errors appear when an incorrect gesture is substituted for the correct one. Deletion is where a correct gesture is omitted in the recognized sequence. Insertion is when an extra gesture was added in the sequence; hence, instead of having one gesture. Here two gestures are segmented (Gesture A and B) instead of one, because the technique recognizes part of A as a separate individual gesture, B.

Relatively less explored issues in gesture segmentation are:

- how to detect the differences between a dynamic gesture (where the path of each limb is relevant), and a static posture (where only one particular position of the limbs is relevant),
- distinction of feature descriptions (which features optimally distinguish the variety of gestures and postures from each other and make recognition of similar gestures and postures simpler), and
- gesture meaning identification (what do certain gestures mean, how can they be reliably interpreted so that the correct actions are undertaken).

B. Culture

Hooler and Beattie [19] claimed that speakers draw on gestures in order to fulfill particular communication functions. Sometimes, culture is a boundary for interpersonal communication; in the same manner

TABLE 1. GESTURE CLASSIFICATION DETAIL

Gesture	Atributes	
	Function	**Linguistic example**
Iconic	Resembles that which is being talked about	Flapping arms like wings when talking about a bird
Metaphoric	Abstractly pictorial; loosely suggests that which is being talked about	Making a box shape with hands when talking about a room.
Beat	Gestures with only two phases (up/down, in/out) indexing the word or phrase it accompanies as being significant	Rhythmic arm movement used to add emphasis
Deictic	Gestures pointing to something or somebody either in concrete or abstract	Pointing while giving directions

Source: Michael Berry, The importance of Bodily Gesture [15]

technology may enable or even hinder a communication style inherent to a culture [20]. Metrically, culture could be reflected in the interactivity, symbol variety, rehearsability and pre-processability of gestures. Therefore, the cultural background might be an influential factor in the design of gesture-based interfaces.

Technology has been conceived in 'prosthetic' terms, as an extension to the body, or support for tasks [21] and given the global diversity; cultures will perceive these tasks differently. Language and representation are critical elements in the study of culture, because humans are locked into their cultural perspectives and mindsets [22]. As defined by Hofstede [23] "Culture is the collective programming of the mind that distinguishes the members of one group or category of people from another".

1) Culture and Interfaces

Humans communicate and exchange information with a system through interfaces. The more familiar or intuitive an interface is, the higher its usability. Maximum usability of a system can be obtained through the appropriate design of support-focused interfaces.

Cultural preferences determine the type of layout, texture, pattern and color [24] in website portals. Certain colors are offensive or uncomfortable for certain cultures. For instance, red is bad luck for Koreans, therefore, Korean websites might avoid the use of red. Given the example, it is noted that to attract a targeted market, there is a need to adapt interfaces to a specific culture. Culture does not exist as a computational term in HCI, even though, as revealed, there are efforts to tailor interfaces. With any use of the technology, the success depends on the capabilities embedded in a persona who is "programmed" in a specific way. The mental "coding" of this persona will affect the usability of the system as well as its interface.

The cultural behavior is generally perceived visually, but it is not always evident until there is an interaction. Rehm, Bee, and André [25] tried to identify the culture of the user so that the behavior of an interactive system could be adapted to culture-dependent patterns of interaction. This was achieved via a Bayesian network model based on gesture expressivity and made use of metrics, such as speed, power or spatial parameters.

A study comparing North Americans to Chinese [26] explored gesture frequency of bilinguals when speaking in both languages. The study demonstrated that American monolinguals used more gestures than Chinese monolinguals and suggested that the American culture is a high-gesturing one. The study also noted that Chinese bilinguals used more gestures than Chinese monolinguals, and suggested that there was possibly a transfer in gesture frequency. The study stated that language is a medium by which culture is transmitted. When speaking in a particular language, the speaker might display the gesture pattern found in the corresponding culture. This means that bilinguals from a lower-gesturing culture and language (such as Chinese) used more gestures when talking in the secondary high-gesturing language (English). In this particular scenario, the difference in gesture frequency might be attributed to the culturally varying attitudes towards body movement.

Some research studies focused on the variation in the expressions between cultures as well as the influence of the language. Nicoladis explained that, this could be due to the fact that bilinguals might tend to produce more gestures, when speaking in a language they might feel weaker at [26]. In this case, the ability to express themselves via gestures might help the speaker to break down the preverbal, spatial-motoric information for verbalization.

In remote international collaboration, gesturing is also a concern. Here, each participant has their own symbolic, iconic and metaphoric influence on their gestures [28]. Given that the Internet is a technological, cultural, political and economic phenomenon, it produces an extraordinary volume of cultural expression [21]. There is a merge of culture amongst the latest generations because computers, interfaces and common systems increasingly assist our lives and have an influence on our gestures. There are studies stating that, since the latest generations have grown up with a mouse based interface, this has a more standard and global effect on command gestures used in their interactions [29]. This suggests the influence of interfaces on culture as opposed to the goal of having culture to influence interfaces.

As stated by Hofstede regarding communication technologies, the software of the machines may be globalized, but the software of the minds that use them is not [23]. Therefore, the dominance of technology over culture is an illusion, and differences between cultures exist.

2) Hofstede's Cultural Dimensions

Hofstede [29] has developed a set of cultural parameters that describe the way in which national societies are built and the rules by which people think, feel and act. These differences are defined as five dimensions and are measured as indexes. The higher or lower the index, more or less the culture portrays this feature.

The Hofstede model of dimensions of national culture has been applied predominantly in international business; marketing and consumer behavior works [30]. The brief descriptions of Hofstede's dimensions are as follows:

Power Distance (PDI) is the acceptance and expectation of power to be distributed unequally viewed from the less powerful members of organizations and institutions (like the family). The higher the index, the more a society views an unequal distribution of power as relatively acceptable (as in Malaysia, Philippines, and Mexico). A lower index indicates that the unequal distribution of power is relatively unacceptable (Austria, Denmark, NZ, Ireland).

Uncertainty Avoidance (UAI) indicates the extent to which a member of society feels uncomfortable or comfortable in ambiguous or abnormal situations. A low UAI index indicates that the society is comfortable in unpredictable situations and has a high tolerance for ambiguity (as in Denmark, Singapore, China, and Sweden). A higher index here means that people prefer predictable situations and have a lower tolerance for ambiguity, a factor that is normally reflected in the abundance of laws in these countries (as in Belgium, Salvador, Greece, and Guatemala).

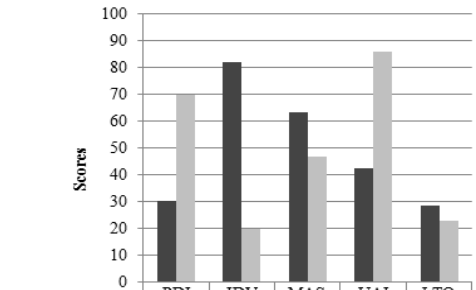
Individualism (IDV) is the extent to which individuals are merged into groups. A high IDV index means members in the society define themselves as individuals and form looser ties with their groups and immediate families (in USA, Anglo-Celtic, UK, and Canada). Countries with lower IDV have stronger bonds to their groups and extended families. Here a group membership forms a person's self identify (in Guatemala, Ecuador, Indonesia, and China).

Masculinity (MAS) refers to the distribution of emotional roles between the genders, and also serves to classify a culture as assertive and competitive (masculine) or modest and caring (feminine). A country with a high MAS index, values achievement and competitiveness, as well as acquisition of money and other material objects (as in Slovakia, Japan, Austria, and Venezuela). In low MAS cultures people value the maintenance of good relationships, are more modest and people oriented, caring for the quality of life (like in Norway, Netherlands, Sweden, and Chile).

Long-Term Orientation (LTO) considers that countries high on this index foster pragmatic virtues oriented towards future rewards, in particular saving money, persistence, and adapting to changing circumstances (as in China and India). Countries with lower LTO are short-termed and give more importance to attitudes towards national pride, tradition, and fulfilling social obligations (as in United States and Norway).

The cultures used in the experiments for this paper are Anglo-Celtic (Australian, British, Irish, New Zealanders) and Latin Americans (American countries where Spanish is primarily spoken: Argentina, Chile, Colombia, Costa Rica, Ecuador, Salvador, Guatemala, Mexico, Panama, Peru, Uruguay, and Venezuela). Fig. 1 shows a comparison of the cultural dimension indexes from both samples; an average was taken of the indexes of the countries mentioned above. The Anglo-Celtic culture had a lower PDI (30 to 70), and UA (43 to 86). On the other hand, it had a higher IDV (82 to 20), MAS (63 to 47) and LTO (29 to 23) than the Latin American one.

Therefore, one can assume that due to the greater equality (Low PDI) characteristic Anglo-Celtics have, they are more individualistic (High IDV) and can master new challenges (Low UAI) better than their fellow Latin American colleagues. Hosftede developed a solid foundation for identifying the possible complication of



	PDI	IDV	MAS	UAI	LTO
Anglo-celtic	30	82	63	43	29
Latin American	70	20	47	86	23

Figure 1. Hostedes 5D Model comparing Anglo-Celtic and Latin American countries.

cross-cultural interactions, as well as what creates cultural differences and how they would act upon this [20].

Even though Hosftede is cited by an extensive amount of sociologists and anthropologists, for the analysis in this study, it is also beneficial to analyze the context classification made by the anthropologist Edward Hall [22]. Hall identifies a culture's use context in routine communication and classifies them as High or Low. In a high context culture (including much of the Middle East, Asia, Africa, and South America), many things are left unsaid, letting the culture explain. There is more non-verbal communication, a higher use of metaphors, and more reading between the lines. In a lower context culture (including North America and much of Western Europe), the emphasis is on the spoken or written word. They have explicit messages, focus on verbal communication, and their reactions are visible, external and outward [31].

Anglo-Celtic cultures (e.g., Australian, British, Irish, and New Zealanders) categorize as low context cultures and Latin Americans (American countries where Spanish and Portuguese are primarily spoken) correspond to the high context cultures. This classification lets us make certain assumptions, like the Anglo-Celtic may predominantly use words, while the Latin Americans would use gestures.

The characteristics identified for each of the samples (IDV, UAI, MAS) will be later referred to in order to interpret their influences on gesture behavior after the experimentation.

III. EXPERIMENT

A set of experiments, following Bischel's experiment guidelines, were conducted in order to explore the influence of culture in gesture behavior. The participants were required to describe two chairs before a camera (See Fig. 2). Bischel's experiments bring language and gesture together; both of these are important in defining a culture. From this moment on, the classical chair will be referred to Chair 1 and the abstract chair will be referred to as Chair 2. Throughout this study, the observational task analysis method was used. Video analysis technique permits a

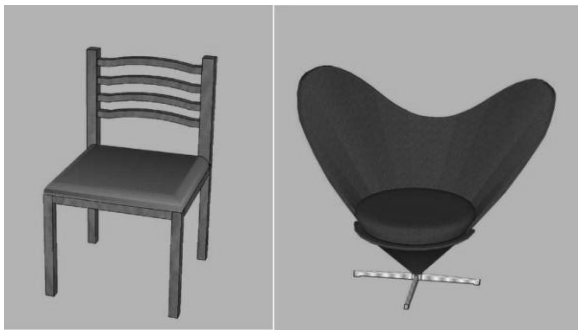


Figure 2. Classic chair (left) CHAIR1 and Abstract chair (right) CHAIR2

careful analysis of gestures occurring at certain timestamps. This is helpful in identifying individual differences in gesture behavior.

A. Hypothesis

The paper tries to prove that: "Designers' culture may affect gesture recognition in multimodal interfaces because of variations in gesture type, gesture frequency, and gesture occurrence".

This hypothesis brings together the subjects of Section II (gesture, multimodal interfaces, gesture segmentation and culture based theories). The three metrics stated in the hypothesis are gesture type, frequency and occurrence.

- Gesture Type. The gesture type is based on McNeill's classification. Certain types of gestures could be attributed to different cultures; therefore, it is important to analyze the type of gesture that is mostly performed.
- Frequency. The frequency is measured as the number of gestures performed by a participant divided by the period of the gesture of the same participant. This way one can obtain the gestures per second, which will help assess speed of gesture performance and point out what gestures are most significant for a gesture recognition system.
- Occurrence. Occurrence measures the appearance of the gestures. This helps to identify if certain gestures are culture-oriented or task-oriented (i.e., related to the task being performed).

B. Experiment Guidelines

Participants were encouraged to use as many gestures as possible, just as in Lui and Kavakli "Temporal Relation between speech and co-verbal iconic gestures in multimodal interface design" [32]. The analysis methodology is via video analysis using a video annotation tool called Anvil (Fig. 3).

1) Procedure

In order to obtain a detailed description, and for us to derive data from video recording, the participants were required to sit in front of a camera in an enclosed setting. This framework helped avoid distractions and background noise that will help with the proper gesture classification.

The influence on gesture performance when holding a seating or standing position by the gesturer has not been assessed. The seating position was thought to simulate seating in front of a screen, where the actual interaction could take place if the user was using a virtual reality interface.

They were instructed that this was a cultural experiment that required analyzing gestures they used to describe a given object. They were encouraged to use both hands, to use as many gestures as possible, and take the time they considered adequate.

2) Participants

There were a total of 8 Latin American participants and 11 Anglo-Celtics videotaped, but only the ones with clearer hand gestures and comprehension of the task were chosen. A criterion for deselecting a video footage for analysis was either the lack of gestures, or the lack of iconic gestures, which are the focus of this study.

The final selection was 5 participants from each sample group, totaling 10 participants. For the purpose of the experiment, two samples were needed, one with English as a first language (Anglo-Celtics), and one with English as a second language (Latin Americans). For the second sample, it was important that they were sufficiently proficient and immersed in an English speaking country (Australia) at least for the past 6 months.

3) Gesture coding

For gesture analysis, the video footages were analyzed and then segmented (Fig. 3). For each occurrence, the gesture type (repetition, beat, iconic, metaphoric, deictic, and junk) that was performed by the participant was recorded. These correspond to McNeill's classification, but the repetition gesture (which is a type of deictic gesture) was coded separately because of the assumption made by Nicoladis in [27] that states that bilinguals might produce more gestures when speaking in a language they might feel weaker at.

Therefore, repetition was considered to be a potential factor that reflects culture, as uncertainty in the language, or description, could be channeled this way. Junk gestures

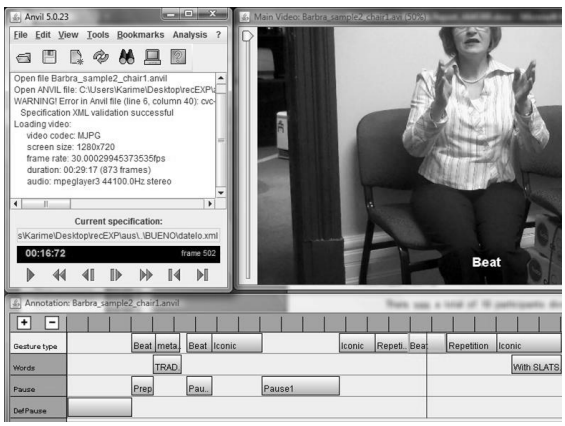


Figure 3. Anvil Snapshot of Anglo-Celtic observation

were identified as gestures without a particular meaning. This could be a gesture that the user takes the gesture back, (which is a "mistake") or made some transition movements.

Gestures are separated by pauses, and a pause is defined as a temporary stop in action or speech [33]. The purpose of this pause was to eliminate the period of inactivity at the beginning of a video, when the participant explains what he or she might do, or when the participant states that he or she has ended the gesture.

4) Speech Coding

As a result of verbal descriptions (words) used to give meaning to the participants' depiction, distinctive words were coded accompanying the gestures. These words, which were also coded on Anvil, were hand written on the track "words". The words coded on this track were identified as significant because they were especially stressed. The classification of words was as follows: adjectives, parts of the chair, verbs, positions and shapes.

The rationale behind this structure is to identify, which accompanying words were used to express gestures so that comparisons could be drawn between the two cultures. The condition where a specific gesture is expressed with a word or not, was not investigated, as the focus was limited to type, frequency, and occurrence of gestures.

IV. RESULTS

Approximately 10 minutes of monologue object descriptions in video footage was obtained. Table 2 contains the metrics referred to in this section. Seconds were used as the time measuring unit. In total, there were 595.52 seconds of video footage captured in 17754 frames. The video footage was composed of 20 individual videos, 2 videos per each participant and a total of 231 gestures.

A. Chair 1.

Chair 1 was a traditional chair, with common characteristics that all participants were able to relate to. When analyzing the data, it was found that Anglo-Celtics on average scored higher values in the gesture duration (1.84s), number of gestures (65), and gesture time per person (22.74s). It is also noticed that the Anglo-Celtics had a higher share (52%) of gestures recorded. The sample shows a considerable variation from one participant's representation to another's, as the mean was 12.8 gestures and the standard deviation (SD) was 5.63.

On the other hand, Latin Americans on average scored lower the values in the gesture duration (1.49s), number of gestures (59), and gesture time per person (17.81s). The mean was 1 gesture lower (11.8), but the variation in this sample was less (2.16) than that of the Anglo-Celtics, meaning more uniformity amongst the sample.

Chair 1 results reveal that Anglo-Celtics produced higher number of gestures in a longer period. Latin Americans, on the other hand, produced fewer gestures that were faster and shorter. The gesture type that was most recurrent was iconic, followed by repetition, beat, metaphoric, junk and then deictic. Anglo-Celtics had more repetition, iconic, and metaphoric gestures. Latin Americans maintained higher deictic and junk gestures. In matters of occurrence, only one Anglo-Celtic performed a deictic gesture.

Frequency might look to be proportional with the amount of gestures but in fact it was not. The overall frequency was higher with the Latin Americans, even though the amount of gestures was less when compared to the Anglo-Celtics (65 to 59). Taking the iconic gestures, for example, Anglo-Celtics on average produced more (6.2) gestures, but their frequency was lower (0.26), as they performed these in a longer time.

In regards to verbal coding, the words used to describe Chair 1 were 28 in total (See Table 3). Most of them correspond to adjectives (See Table 4) given to the chair as a whole, as well as specific parts of it. From this 9 words in total were uniquely used by Anglo-Celtics, and 13 by Latin Americans. Both shared 6 specific words: horizontal, front, seat, legs, back and square. Therefore, it can be concluded that the chair was easy to describe both physically and qualitatively.

B. Chair 2.

Chair 2 was a more complex chair, with a more elaborate structure. To an extent it is considered as "abstract". All participants mentioned some way the chair was different, therefore more difficult to describe. The data for the gesture representation for this chair is also in Table 2. The mean of gestures for the chair 2 descriptions for both samples was 10.8 (Total number of gestures per chair divided by the 10 participants) and the SD was 5.65

TABLE 2. VIDEO ANALYSIS FOR CHAIR 1 AND CHAIR 2

Chair	**Metrics**								
	Total no. gestures per chair	***Sample***	***Average gesture duration***	***Total no. of gestures per chair per sample***	***Percentage attributable to sample***	***Average gestures per person in sample***	***Standard deviation (SD)***	***Average gesture time per person in sample***	***Frequency of gesture performance per sample***
Chair 1	124	Anglo-Celtic	1.84	65	52%	12.8	5.63	22.74	0.56
Chair 1		Latin American	1.49	59	48%	11.8	2.16	17.81	0.66
Chair 2	108	Anglo-Celtic	1.73	65	60%	13	7.17	23	0.56
Chair 2		Latin American	1.67	43	40%	8.6	2.88	14.22	0.60

(Average of both SD's). This SD is higher than Chair 1, showing different approaches followed by participants to describe the chair.

The data reveal that Anglo-Celtics on average scored again higher values in gesture duration (1.73s), number of gestures (65), and gesture time per person (23s). Anglo-Celtics had higher scores in the average duration of gestures and had a higher share (60%) of gestures recorded. Latin Americans on average scored lower values in gesture duration and number of gestures, but these values were significantly lower than the time required for Chair 1.

As with gesture and frequency, the differences in chair descriptions are also reflected in the use of words. Chair 2 surpasses Chair 1 with a total of 30 words (See Table 3). Anglo-Celtic uniquely used 13 words and Latin Americans 10. Both samples used 5 words in common: Back, seat, curved, heart, and cone. This variation could be because the chair was more complex and needed a higher evaluation and explanation. It was found (Table 5) that when describing the back of the chair, the same word was used, and this word was "heart". It was also found that there were more verbs and shapes in this chair description; words used to describe the chair metaphorically and qualitatively.

Chair 2's general data confirm that Anglo-Celtics again had higher number of gestures in a longer period. This is reflected in the lower frequency achieved by the Anglo-Celtics in Fig. 4b. Latin Americans on the other hand performed fewer gestures that were faster and shorter. Table 2 shows that the Anglo-Celtics used more or less the same amount of gestures as before. This could be because they are more comfortable when describing the "abstract" concepts in their first language.

TABLE 3. WORDS DERIVED FROM ALL CHAIR RECORDINGS

Sample	Recorded Words			
	Anglo-Celtic	***Latin American***	***Both***	***Total***
Chair 1	9	13	6	28
Chair 2	13	10	5	30

TABLE 4. WORDS DERIVED FROM ALL CHAIR 1 RECORDINGS

Classification	Recorded Words
Adjectives	Traditional, Thin, Arched, Vertical, Squarish, Long, Curved, Rectangular, Straight, Round, Normal, Flat
Parts	Slats, Bottom, Front, Seat, Legs, Place, back
Verbs	Support
Position	First, left
Shapes	Holes, squares, sticks, stripes

TABLE 5. WORDS DERIVED FROM ALL CHAIR 2 RECORDINGS

Classification	Recorded Words
Adjectives	Round, Horizontal, Perpendicular, Funny, Curvy, retro, stainless, steal, shortest, curved, symmetric
Parts	Back, Bottom, Down, seat
Verbs	Focus, Spread, Crosses, extend
Position	Middle
Shapes	Heart, Cone, Square, Circle, Wing, Stick, Oval, Peak, Cross, triangle

C. Findings

After analyzing the performance of both samples, in this section, the results of the metrics stated in the hypothesis (gesture type, frequency and occurrence) are delivered. The results come as follows:

1) Frequency

Gesture frequency indicates that overall the Latin American sample performed more gestures per second; however, this evidence is not enough to state which culture was more expressive than the other. The use of gestures involves various factors, such as the comfort of a person had in front of the camera, or the confidence with the object being described, as well as the language. Chair 1 had iconic and repetition gestures with higher frequency in both samples, and this is reflected in Fig. 4.a and Fig. 4.b. Chair 2 on the other hand had an increase in junk and metaphoric gestures. The most significant gestures for gesture recognition are the iconic ones as well as repetitions, and subsequently they are the ones that convey the description of the chair more significantly.

2) Occurrence

There are no junk and deictic gestures in the description

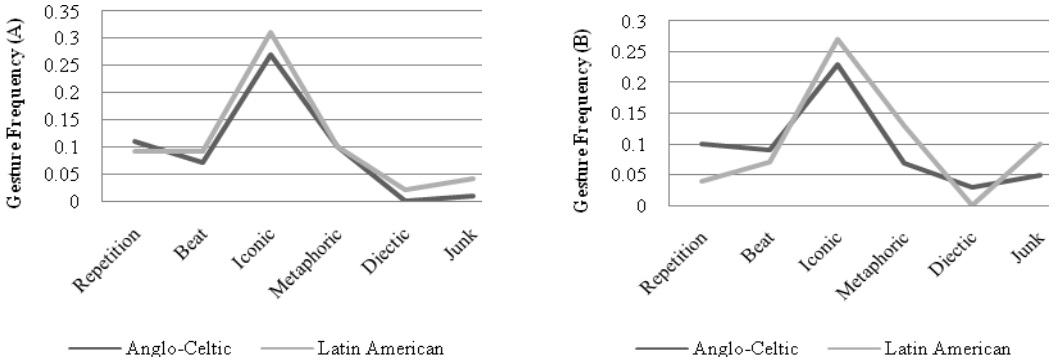


Figure 4. Gesture Frequency of gestures by chair and sample:
Gesture Frequency Chair 1 (a)
Gesture Frequency Chair 2 (b)

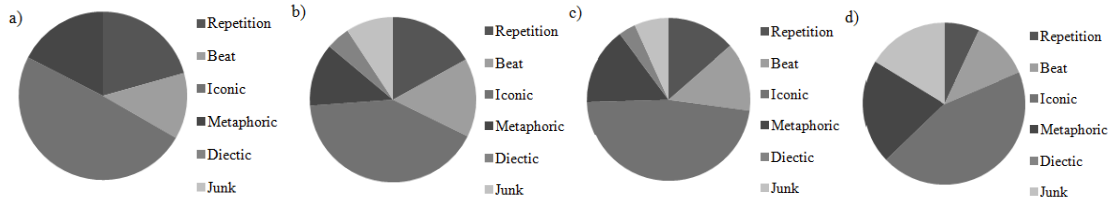


Figure 5. Gesture types by chair and sample:
Chair 1- Anglo-Celtic (a) Chair 1- Latin American (b)
Chair 2- Anglo-Celtic (c) Chair 2- Latin American (d)

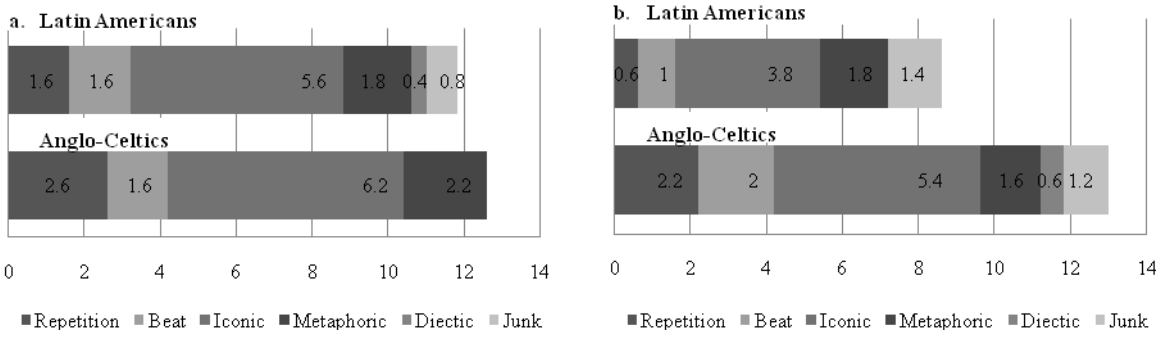


Figure 6. Average Gesture type by sample in
Chair 1 (a) and Chair 2 (b)

of Chair 1 for the Anglo-Celtic sample, but they do appear in Chair 2 when the number of gestures increases. This means that the occurrence of gestures was related to the task, not to the culture. This may be because Chair 2 was more complex and there was a need for more explanation by the user. This explanation required more cognition and the extra gestures reflected this process.

3) Gesture Type

For Chair 1, the iconic gestures were close to 50% in both sample groups (See Fig. 5a and Fig. 5b), showing equality. For Chair 2, the iconic gestures diminish and metaphoric gestures increase for the Latin American sample group (See Fig. 5c and Fig. 5d). Again, this may be related to the relative complexity of the chairs. Regarding verbal depictions, each iconic and metaphoric gesture was related to at least one word, reflecting the participants' cognition.

D. Comparisons

In this section, we summarize the important points described in Section C. Numerically; Anglo-Celtics did not display too much variation between chair descriptions, even though they all encountered a new and different chair that required a lot more extraordinary description.

Chair 2 had more gestures on average by participant in the Anglo-Celtic sample. On the contrary, in Chair 2, Latin Americans performed fewer gestures on average by participant. The reason could be the degree of comfort Anglo-Celtics had when describing an abstract chair.

The SD was again higher with the Anglo-Celtics. This made it hard to identify a pattern. On the other hand, Latin Americans had a smaller SD and more frequent gestures, meaning shorter, more concise and common gestures by most of the participants.

The gesture frequency was higher in Chair 1, and it increased with the Latin Americans. This could be partly because; they scored higher values in junk gestures in the description of Chair 2. Latin Americans had more frequent gestures in both chairs. This means that they performed more gestures per second, even though they had fewer gestures in total. The smaller count of gestures by Latin Americans is justified by shorter time frame in which they performed the gestures.

Deictic and junk gesture occurrence was significantly less than the other gestures. Therefore, the idea of Latin American's being more explanatory with hands is disproved. A potential explanation in this case, is perhaps the implication of having to speak in a second language.

Given the distribution of gestures (Fig. 5 and Fig. 6), it is identified, that in general, iconic gestures decrease with Chair 2, as well as the repetition gestures. In contrast, junk and deictic gestures appear more; the average of these gestures therefore, increases from one chair to another (Fig. 6a and Fig. 6b).

Latin Americans used more distinctive words for Chair 1 (13) and less in the Chair 2 (10) (Table 3). Using less gestures and words to describe Chair 2 could probably mean a better selection of words and gestures, or the lack of vocabulary. The higher word count for Chair 1 may refer to either a higher degree of confidence, or more predictable and well structured ideas.

V. DISCUSSION

Now the relation between the gesture metrics and the cultural attributions made by both Hofstede and Hall (Section II b) is presented. As Anglo-Celtics are low context cultures, they used more words and gestures in longer time, since they took time to explain the chair in detail. On the other hand, Latin Americans, which represent the high context culture, performed fewer gestures, in shorter time and used fewer words. The element that calls for attention is the higher use of metaphoric gestures, a trait of a high context cultures. This exemplifies a characteristic of a society that relies on reading between the lines and letting nonverbal cues explain the meaning.

The cultural analysis now continues by relating gesture performance with Hofstede's cultural dimensions. A clear integration between the culture literature cited and the experiment conducted is exposed in Table 6. As mentioned before, the traits that are mostly reflected are IDV, UAI, and MAS.

- IDV. This trait could be related to the fact that the SD between samples is higher with the Anglo-Celtic cultures reflecting the societies high individualism index (IDV, 82). On the other hand, the low SD with the Latin Americans shows the low individualism index (IDV, 20).
- UAI. This trait could be reflected in the overall impression of Chair 2. The Anglo-Celtic sample did not vary too much in gesture means and time from one chair to another, showing greater comfort with adverse situations (UAI, 43). It is possible to say that Latin Americans showed their high uncertainty avoidance (UAI, 86) since they use less time and limited gestures, possibly sticking to "what they knew" instead of managing the abstract.
- MAS. This trait could be related to the fact that the Anglo-Celtics as a low context culture are more masculine and assertive (MAS, 63), in comparison to the Latin Americans that are more human-oriented and feminine. Therefore there is a higher use of metaphors (MAS, 47) in their descriptions.

It is important to remark that the Latin Americans in this sample have more of an advantage with the language compared to "at home" Latin Americans, as they have been immersed in a different culture and language for the past 6 months. Regardless of that, they still performed fewer gestures and chose different words.

VI. CONCLUSION AND FUTURE WORK

The goal of this paper was to explore if gesture-based interfaces could be affected by a user's culture. The literature review agrees that any interaction is a result of user, task and input. Aside from performance or stability issues, gesture-based interfaces are subject to a context problem. In the international scene, depending on where participants are from, their style of communication will vary. This analysis arrived to the conclusion that as the complexity of a task increases, so does the use and type of gestures. The metrics stated in the hypothesis influence multimodal interfaces and their performance in the following ways:

- Frequency may affect the recognition rate because of the need for faster and more efficient algorithms.
- Occurrence also affects interaction due to the possibility of absence (zero occurrences) of certain gestures that may convey functionality (i.e., iconic).
- Gesture type, as well as occurrence, also affects the goal that the user wishes to attain. Identifying and classifying certain gestures due to their use during trials would permit the identification of type tendencies and will assist in embedding differences in the development of the gesture recognition tools.

Due to the "freedom" that hand gestures provide, gesture based interfaces gain popularity. The aim of HCI is to have users adopt the new technologies for interaction because their usability is better. Studies have shown that culture influences a user's acceptance of the technology. A more conservative or traditional culture, such as the Latin American culture, could take more time to adapt. This was visible with the differences in frequency rates between the classic and abstract chairs. The performance of a gesture based interface will not only be affected by the task being performed, but by the cultural background and language skills of the user. Therefore, the design of gesture-based interfaces not only requires a multidisciplinary approach, but also a culturally sensitive one. The work conducted in order to develop applications that are consistent with the perception, use, and understanding of gestures by users, still continues.

It is acknowledged that future studies need a larger sample size. Similarly, future studies need to focus on the consistency of the annotations by having more than one coder to increase the objectivity of analysis. The results might have significant variations if the experiment were to

TABLE 6. INTEGRATION OF EXPERIMENT AND CULTURE

Sample	**Metrics**			
	Context	***Predominant culture trait by Hofstede***	***Metric Evidence***	***Predominant Gesture Type***
Anglo-Celtic	Low context (assertive, rely con words)	Individualism Masculinity	High SD Constant gestures beween chairs More gestures and more time	Iconic
Latin American	High context (rely heavily on non verbal communication)	High Uncertainty Avoidance Collectivism	Low SD Fewer Gestures in the second Chair Fewer gestures in less time	Metaphoric Repetition

be carried out in Spanish, the native language of the Latin American culture. Further research studies may also attempt to investigate the effects of gender on gesture behavior.

ACKNOWLEDGMENT

This research is supported by the Australian Research Council Discovery grant DP0988088 to Kavakli, titled "A Gesture-Based Interface for Designing in Virtual Reality". Also, the experiment on which this project was based, is inspired by Jing Liu's previous work titled "Temporal Relation between speech and co-verbal iconic gestures in multimodal interface design". Authors are grateful to her for her support in sharing data.

REFERENCES

[1] Kavakli, M. and Nasser, K. 2012. Influence of Culture in Multimodal Interfaces: Differences in Gesture behavior between Anglo-Celtics and Latin Americans, *MMEDIA 2012: the fourth International Conferences on Advances in Multimedia*. Pp 148-154 ISBN: 978-1-61208-195-3

[2] CARLA. 2012. What is Culture?. The Center for Advanced Research on Language Acquisition, http://www.carla.umn.edu/culture/definitions.html [Accessed: December 15, 2012]

[3] Visser, W. and Maher, M.L. 2011. The Role of gesture in designing. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 25, pp. 213-220. Cambridge University Press, doi: 10.1017/S08900604/11

[4] Wachs, P., Kölsch, M., Stern H., and Edan Y. 2011. Vision-Based Hand-Gesture Applications. *Communications of the ACM*, 54 (2), pp. 60-71, doi:10.1145/1897816.1897838

[5] Kita, S. (ed.). 2003. Pointing and Placing. Pointing: Where Language, Culture and Cognition Meet, http://www-psych.stanford.edu/~herb/2000s/Clark.Pointing.placing.03.pdf [Accessed: December 15, 2012]

[6] Bavelas, J., Gerwing, J., Sutton, Ch., and Prevost, D. 2008. Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58 (2), pp. 495-520 ISSN: 0749596X, doi: 10.1016/j.jml.2007.02.004

[7] Gullberg, M. 2010. Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research*, 26 (1), pp. 75–102. doi: 10.1177/0267658309337639

[8] Karray, F., Alemzadeh, M., Saleh, J., and Arab, N. (2008) Human-Computer Interaction: State of the Art. *International journal on smart sensing and intelligent systems*, 1(1), http://www.s2is.org/Issues/v1/n1/papers/paper9.pdf [Accessed: December 15, 2012]

[9] Lanfranco, A., Castellanos, A., Desai, J., and Meyers, W. 2004. Robotic Surgery. *A Current Perspective Ann Surg*, 239 (1), pp. 14–21. doi: 10.1097/01.sla.0000103020.19595.7d

[10] Purdue University. 2011. Future surgeons may use robotic nurse, 'gesture recognition'. *ScienceDaily*, http://www.sciencedaily.com/releases/2011/02/110203152548.htm [Accessed: December 15, 2012]

[11] Van den Hoven, E. and Mazealek, A. 2010. Grasping Gestures: Gesturing with physical artifacts. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 25, pp. 255-271, doi:10.1017/S0890060411000072

[12] Kim, M.J. and Maher, M.L. 2008. The impact of tangible user interfaces on spatial cognition during collaborative design. *Design Studies*, 29 (3), pp. 222–253, doi: 10.1016/j.bbr.2011.03.031

[13] Karam, M., Hare, J., and Schraefel, M. n.d. Ambient Gestures, http://users.ecs.soton.ac.uk/amrk03r/AmbientGestures.pdf [Accessed: December 15, 2012]

[14] Kavakli, M., Taylor, M., and Trapeznikov, A. 2007. Designing in virtual reality (DesIRe): a gesture-based interface. In *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts* (DIMEA '07). ACM, New York, NY, USA, 131-136, doi:10.1145/1306813.1306842

[15] Kavakli, M. 2008. Gesture Recognition in Virtual Reality, *Special Issue on: "Immersive Virtual, Mixed, or Augmented Reality Art" of The International Journal of Arts and Technology* (IJART), 1 (2), pp. 215-229. doi: 10.1504/IJART.2008.021928.

[16] Berry, M. 2009. The Importance of Bodily Gesture Example 1. McNeill's classification of gesture types, http://www.mtosmt.org/issues/mto.09.15.5/berry_ex1.html [Accessed: December 15, 2012]

[17] Kang, H., Lee, C., and Jung, K. 2004. Recognition-based gesture spotting in video games. *Pattern Recognition Letters*, 25 (15), pp. 1701-1714, doi:10.1016/j.patrec.2004.06.016

[18] Li, H. ad Greenspan, M. 2011. Model-based segmentation and recognition of dynamic gestures in continous video streams. *Pattern Recognition*, 44, pp. 1614-1628, doi:10.1016/j.patcog.2012.12.014

[19] Holler, J. and Beatiie G. 2004. The interaction of iconic gesture and speech in talk. *Gesture Based Communication in Human Computer Interaction*, 2915, pp. 3853-3885, doi: 10.1007/978-3-540-24598-8_6

[20] Massey, A., Hung, C., Ramesh V., and Montoya-weiss, M. 2001. Cultural perceptions of task-technology. *Communications of the ACM*, 44 (12), doi:10.1145/501347.501353

[21] Murphie, A. and Potts, J. 2003. Culture and Technology. N.Y: Palgrave Mcmillan

[22] Hall, S. 2007. Review: Representation: Cultural Representations and Signifying Practices (Culture, Media and Identities Series) London: SAGE Publications Limited

[23] Hofstede, G. H., Hofstede G. J., and Minkov, M. 2010. Cultures and organizations: software of the mind : intercultural cooperation and its importance for survival. (3rd ed) McGraw-Hill Professional ISBN 0071664181, 9780071664189

[24] Aaron Marcus and Associates, Inc . 2010. Cross-Cultural User-Interface Design for Work, Home, and On the Way: Accounting for Cultural Preferences, Acceptance, and Constraints. ACM SIGGRAPH ASIA 2010 Courses, Doi: 10.1145/1900520.1900525

[25] Rehm, M., Bee, N., and André, E. 2008. Wave like an Egyptian — Acceleration based gesture recognition for culture-specific interactions. In *Proceedings of HCI 2008*

Culture, Creativity, Interaction, 1, pp. 13–22, ISBN: 978-1-906124-04-5

[26] Wing Chee, S. 2010. "Cross-cultural transfer in gesture frequency in Chinese–English bilinguals". *Language and Cognitive Processes*, 25 (10), pp. 1335-1353, doi: 10.1080/01690961003694268

[27] Nicoladis, E. 2007. The effect of billingualiam on the use of manual gestures. *Applied Psycholinguistics*, 28, pp. 514-526.

[28] Yammiyavar, P., Clemmensen, T., and Kumar, J. (2008). Influence of cultural background on non-verbal communication in a usability testing situation. *International Journal of Design*, 2(2), pp. 31-40, http://www.ijdesign.org/ojs/index.php/IJDesign/article/view/313/164 [Accessed: December 15, 2012]

[29] Fikkert, W., Vna der Vet, P., Van der Ver, G., and Nijholt, A. 2009. Gestures for large display control . *Computer Interaction: 8th International Gesture Workshop*, Revised Selected Papers, pp. 245-256. Germany: Springer

[30] Hofstede, G. 2011. Dimensions of national Cultures., http://www.geerthofstede.nl/dimensions-of-national-cultures [Accessed: December 15, 2012]

[31] Hall, E. T. 1976. *Beyond Culture*. Knopf Doubleday Publishing Group. ISBN 0385124740, 9780385124744

[32] Liu, J. and Kavakli, M. 2011. Temporal Relation between speech and co-verbal iconic gestures in multimodal interface design, http://coral2.spectrum.uni-bielefeld.de/gespin2011/final/Liu.pdf [Accessed: December 15, 2012]

[33] Pause .2011. In *Oxford Dictionaries*, http://oxforddictionaries.com/definition/pause [Accessed: December 15, 2012]