Human Perception of Laughter from Context-free Whole Body Motion Dynamic Stimuli

McKeown, G., Curran, W., Kane, D., McCahon, R., Griffin, H. J., McLoughlin, C., & Bianchi-Berthouze, N. (2013). Human Perception of Laughter from Context-free Whole Body Motion Dynamic Stimuli. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII), (pp. 306-311). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ACII.2013.57

Published in:
2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII),

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Human Perception of Laughter from Context-free Whole Body Motion Dynamic Stimuli

Gary McKeown†, William Curran*, Denise Kane*, Rebecca McMahon*  
Harry J. Griffin†, Ciaran McLoughlin†, Nadia Bianchi-Berthouze†  
*School of Psychology, Queen’s University Belfast, UK. Email: g.mckeown@qub.ac.uk; n.berthouze@ucl.ac.uk  
†UCL Interaction Centre, University College London, London, UK. Email: harry.griffin@ucl.ac.uk; n.berthouze@ucl.ac.uk

Abstract—Laughter is a ubiquitous social signal in human interactions yet it remains understudied from a scientific point of view. The need to understand laughter and its role in human interactions has become more pressing as the ability to create conversational agents capable of interacting with humans has come closer to a reality. This paper reports on three aspects of the human perception of laughter when context has been removed and only the body information from the laughter episode remains. We report on ability to categorise the laugh type and the sex of the laugh; the relationship between personality factors with laughter categorisation and perception; and finally the importance of intensity in the perception and categorisation of laughter.

Keywords: laughter, body movement, motion capture, laughter perception, personality

I. INTRODUCTION

Human laughter is one of the most intriguing yet understudied social signals. It seems primarily to function as a social bonding mechanism [1], yet it can be rapidly elicited in response to the often complex cognition required by humour. Given these features and that it occurs with common frequency in most human interactions [2] it is interesting that it has largely remained outside the sight of serious scientific scrutiny until fairly recently, presumably due to its association with the less serious elements of human behaviour.

With the technological advances in the recognition of social signals and synthesis of human behaviours [3] it has become apparent that laughter as a social signal and its role in conversational interaction must be understood if we are to produce conversational avatars that can interact naturally with humans. As part of this endeavour, a strong understanding of the component processes and social signals that contribute to a laugh response must be developed. The various aspects of a laugh response must be scrutinised to assess the contributions they make to laughter in human interaction. This information is crucial in the development of human-machine interfaces that seek to simulate natural human interactions—to inform machine learning approaches to laughter recognition and to develop naturalistic synthesis in laughter animations. These include the acoustic properties of laughter, the facial expressions associated with laughter, the body movements that result from a laughter response and the sequential dynamics of laughter within human interactions. Their has been research on robot laughter [4], however, this paper concentrates on the contributions to human perception of laughter that arise as a result of laughter associated body motion.

Here we report on three complementary studies that examined human perception of body motion associated with laughter responses. The first study examined participants’ ability to categorise laughter and recognise features of a laugh—specifically laugh sex—when presented only with the body motion information. The second study assessed certain personality factors that may be associated with the perception of laughter. A final study looks at the relationship between laugh intensity and categorisation of laughter. These aspects will now be addressed in more detail.

A. Categorisation ability

The first strand of investigation we sought to assess was the ability of people to categorise laughter. Specifically we wanted to ask whether, in the absence of its social and interactional context, people can still categorise laughter into different types related to the level of social function. Various studies have shown that acted and non-acted affective states can be categorised from simplistic faceless and genderless avatars (for a review see [5]). In this study we sought to see if people could assess laughter type based solely on body motion without any of the other factors that serve to influence the interpretation of laughter. We also sought to see if people could correctly assess the sex of a laugh from just the body motion. A number of studies have shown that sex can be discriminated on the basis of gait displayed as point light animations [6] and that the discrimination was greater when the displays were normalized to body size [7].

B. Personality factors

The role of personality factors in humour and laughter has been suggested by many [8]. There are several personality factors that are likely to influence how laughter is perceived and how it is interpreted. Laughter has the interesting quality of being tightly coupled to high level cognitive processing with reflex-like responses, however there are many stages in this process where personality factors may have an influence. Laughter has important social aspects, of specific interest is its social bonding purpose; this can serve to be an inclusive social signal and also as an exclusive social signal. In this regard, laughs can be interpreted as benevolent and inclusive—laughing with a person—or, when two or more people laugh together maliciously, it can serve to bond the laughers but exclude a third person who is not included in the shared laughter—laughing at a person. The correct interpretation of laughter in these circumstances is important for efficient and successful social functioning. The importance of laughter...
interpretation is highlighted by a pathological form of laughter—gelotophobia. According to Ruch and Proyer [9], gelotophobia can be defined as “the pathological fear of appearing to social partners as a ridiculous object”, in other words the fear of being laughed at. Gelotophobes tend to be more paranoid in the presence of people laughing, more sensitive to offense, and socially withdrawn. Although gelotophobes have a lot in common with people who suffer from social withdrawal, the fundamental difference is that gelotophobes see themselves as being “ridiculous” and “weird” and expect others to laugh at them because of this. They experience shame as a result of these “presumed shortcomings”.

It is important to identify the processing stage(s) at which the perception and associated cognitive processing of laughter, and one’s response to it, is influenced by personality factors. There is a range of possibilities in this regard. It may be the case that the interpretation of a laugh as being malicious or benevolent occurs at a very early stage in the perceptual processing of laughter, or it may be the case that laughs are observed in the same manner by all participants but an interpretation on the meaning of a laugh comes at a later social cognition stage of processing—a stage that involves greater integration of the social and interactional factors associated with laughter processing. Assessing laughter using stimuli that have had all the social and interactional context factors removed permits an assessment of whether these personality factors exert their influence at an early perceptual stage. Finding a perceptual disinclination towards laughter or a lack of ability to classify laughter would suggest a deep pathology. If no low-level differences are found then this is suggestive that aversions to laughter may be more likely to occur at a social cognition stage and with higher level social interpretations of the meaning of laughter, although this would require further research to substantiate.

Aside from gelotophobia there are a number of other personality factors that may be relevant to laughter perception and cognitive processing. The related phenomena of gelotophobia, which is the the joy of being laughed at, and katagelatism, which is the the joy of laughing at others, are likely candidates to be influenced by perceptual and context-free body motion laughter. In addition we included measures of emotional contagion, the short measure of the “big five” personality factors, and measures of cheerfulness, seriousness and bad mood.

C. Intensity

Of the dimensional variables that have been associated with laughter one of the clearest in importance appears to be intensity and, in particular, the intensity of the Duchenne display [10]. The Duchenne display is typically associated with smiling, where a smile includes contraction of both the zygomatic major muscle and the orbicularis oculi; in Facial Action Coding (FACS) terms a smile that includes not only AU12 but also AU6. The classic interpretation of a Duchenne smile is that it is more genuine, whereas a smile without the Duchenne display is more likely to be fake or a polite social smile [11] (although see [12]). The Duchenne display has been argued to be an important indicator in the differentiation of laughter [13]; in particular the differentiation of hilarious laughter and social laughter can be determined by the increase in the intensity of AU6—contraction of the orbicularis oculi. Clearly this analysis of the intensity of laughter relates specifically to laughter as it is signalled through facial expression. In this study we sought to see if similar importance could be placed on the intensity of laughter as evidenced by the intensity of body motion.

This creates an important question about what constitutes intensity with respect to laughter in body motion. Obviously it is not just intensity of movement, there are many degrees of freedom of movement in body motion and only a small subset of these are related to laughter. Some suggestions have been made concerning the appropriate muscular movements associated with laughter. Mancini et al. [14] focussed on a range of key movements associated with laughter perception. In this study we do not address directly what features of the movement are important for rating the intensity of a laugh ([15] addresses features that may be involved in laughter categorisation), we assume that most humans will have a degree of expertise regarding the ability to define a laugh as intense or not. Therefore in this study we leave the interpretation of intensity of laughter up to the human “expert” and do not provide any explicit instructions regarding movements that they would expect to produce an intense laugh.

II. Methodology and Data Collection

A series of laughter stimulus events were recorded in situations in which people were made to laugh using a variety of interactive laugh induction techniques [16]. The laughs were categorised by the experimenter as either hilarious, social, awkward, fake, or non-laughter. The laughers were wearing motion capture equipment and the resulting motion capture data were turned into short animations of the body motion associated with the laugh. These animations became the focus of a forced-choice perceptual experiment which provided the results that are used in this paper.

A. Laughter Capture

Recording sessions involved volunteers interacting in pairs in each session—though this paper individuals being recorded are referred to as volunteers and those observing recordings are referred to as participants. A total of 18 volunteers were involved but only one of each pair was recorded using motion capture equipment (6 females and 3 males); their mean age was 25.7 and they were drawn from a mix of cultural backgrounds, including Western European, East Asian, North American and South Asian. An inertial motion capture suit was used to gather the body motion data (Animazoo IGS-190). Following [14] the suit had been adjusted to ensure the capture of relevant spine and shoulder movement information. Laughter was recorded while volunteers were actively engaged in the laughter-inducing tasks and in conversational periods between the tasks. At stages throughout the recording session volunteers were asked to produce fake laughter on demand.

B. Stimulus Preparation

The individual laughter animations were prepared by segmenting laughter episodes on the basis of video recordings. These were categorised by the experimenter as hilarious; social (back-channeling, polite, conversational laughter); awkward (involving a negative emotion such as embarrassment or
discomfort on another’s behalf); fake (when they had been instructed to produce a fake laugh), or non-laughter. A total of 508 laughter segments and 41 randomly located non-laughter segments, containing other behaviour such as talking, were identified. These segments were animated using the motion capture data to create a stick figure rendering of the body motion. These avatars were defined by positional co-ordinate triplets of 26 anatomical points over the whole body. The anatomical proportions were the same for all animations (Figure 1) and there was a goal of creating an androgynous figure. A standardized viewing angle placed at a slightly elevated viewpoint was used for all the animations as this has been shown to influence perception results [17], [7]; this viewpoint did not change if models walked or turned during the standing tasks. These stimuli were distributed across the five categories based on the natural frequency of laughter-types [18] which provided a total of 126 animations to be used in this study (34 hilarious, 43 social, 16 awkward, 19 fake, 14 non-laughter—mean duration = 4.1s, standard deviation = 1.8s).

An important element of the categorisation procedure depends on how we determine the “ground truth” against which we compare the categorisation of the human raters. Here we adopt the approach that the experimenter-categorized ratings are sufficient as a comparison. This is not too problematic in the categories of fake and non-laughter where the nature of the laugh is quite clear from the context, but more subjective decisions are required to differentiate between hilarious, social, and awkward laughs. Another approach would be to independently rate the laughter using a number of raters watching the video clip, however other research in laughter has suggested that independent raters may not be able to categorise laughter in this way to a sufficiently high standard [19]. There appears to be important information concerning the social setting, the mood and mannerisms of the volunteers that is missing when short video sections of laughs are provided for independent rating. Our recommendation for the analysis of “ground truth” in future would be to have rating of laughter occur live at the time of the session with a number of independent raters present to judge in real-time. In the absence of this ability we assume that the experimenter has privileged access to contextual information concerning the events at the laugh gathering session. The experimenter also knew the volunteers and their temperaments providing further social knowledge; we therefore use the experimenter ratings with the caveat that the ratings of a single rater do not provide an ideal ground truth.

C. Personality Measures

We assessed personality using a number of standard personality measures as well as the measure that assessed gelotophobia.

1) Gelotophobia: The gelotophobia scale was measured as part of a larger 45-item measure, known as “PhoPhiKat-45”. This provides scales that detect levels of gelotophobia (the fear of being laughed at), gelotophilia (the joy of being laughed at), and katagelaticism (the joy of laughing at others) [20]. The level of gelotophobia is measured on this scale as a dimension and a person is deemed to have a slight expression of gelotophobia above a threshold of 2.5 and pronounced beyond a threshold score of 3. In this study we use the 2.5 level.

2) Ten Item Personality Inventory (TIPI): This is a short 10-item questionnaire, used to measure the five factor personality model commonly known as the “big five” personality dimensions: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism, [21].

3) The State-Trait-Cheerfulness-Inventory (STCI): A 60-item questionnaire with three sub-scales used to measure participants’ cheerfulness, seriousness, and bad mood [22].

4) Emotional Contagion Scale (ECS): A 15-item measure used to assess participants’ tendency to mimic the five basic emotions: love, happiness, fear, anger and sadness [23]. It produces a sub-scale for each of these basic emotions.

D. Ratings collection

Participants were presented with the materials on a PC or laptop. They were asked to watch each of the 126 videos in turn and complete a response sheet in the form of a spreadsheet. They had to categorise the laugh they observed as either hilarious, social, awkward, fake or non-laughter; they were asked to categorise the laughter as either male or female; and finally they were asked to provide a rating of the intensity of the laugh between a low intensity level of 0 and a high intensity level of 5. The order of the laugh animation presentation was randomised for each participant and they took a 5 minute break when they had provided ratings for half of the animations. It took approximately 2 hours for each participant to complete the annotation and personality measures.

E. Participants

A total of 54 participants (34 female and 20 male) rated each of the 126 laughter animations, providing approximately 6804 laugh judgments for each of the variables in the study;
Similar procedure for assessing categorisation ability we adopt
ment in the ground truth categorisation of laughter sex. Using a
categorisation of laughter type; there is no subjective judge-
issues of ground truth in the same manner that occurs in the

type.

seems to be little relationship with ability to categorise laughter
over the pathological threshold for gelotophobia and there
highlighted in green on this figure the individuals who scored
be relied upon to provide a correct categorisation. Figure 2
they are only performing at just above chance and could not
than in others; however this is not a particularly strong effect,
The ability seems to be present and better in some participants
body motion at better than chance levels

participants showed an ability to categorise based on only
features of body movement. A total of 6802 laugh judgements of
28.85 laughs correctly categorised, with a standard deviation
126 laughs produced a mean participant categorisation level of
55 participants were tested, one was excluded for not filling
out the response sheet correctly. Ages ranged from 17-69
years with a mean = 29.7 and standard deviation = 14. They
were recruited in [removed for review] and were of mixed
educational/social backgrounds. Participants were formally re-
cruited and required to give informed written consent before
participating.

III. RESULTS

A. Categorisation

categorisations in this section are made against the “ground
truth” of the experimenter’s chosen category. This differs from
[15] which uses modal laughter perceptions to investigate the
automatic recognition of laughter type based on features of

Even though performance on both of these tasks was
low we found a strong correlation between those participants
who performed well at the categorisation of laugh type task
and those who performed well at the categorisation of sex
r = .45, p = .001. Figure 4 displays a scatterplot showing
this relationship. So although people do not seem to be good
at these tasks there does seem to be a common ability that is
shared across these tasks. A further way to assess ability that
is not dependent on the subjective judgement of the ground
truth categorisation is to look for inter-rater agreement in the
categorisation (using Fleiss’ kappa). As might be expected in a
situation with a weak ability there is little inter-rater agreement
in the overall group $\kappa = 0.06$, however if we only assess the
inter-rater agreement amongst the best performing participants
in the categorisation tasks—the six participants in the top right
of Figure 4—there is a slight but present level of agreement
$\kappa = 0.12$, this is significantly different from a value of kappa
derived by subsampling 6 randomly selected raters 10,000
times which gives a mean $\kappa = 0.04$ with 99% confidence
intervals at upper $= 0.041$, lower $= 0.039$. 

Fig. 2. Performance level of the 54 participants in laugh type categorisation. The red dotted line represents chance performance while the orange line represents the group mean performance.

Fig. 3. Performance level of the 54 participants in laugh type categorisation. The red dotted line represents chance performance while the orange line represents the group mean performance.

\begin{itemize}
  \item \textbf{A. Categorisation}
  
  Categorisations in this section are made against the “ground truth” of the experimenter’s chosen category. This differs from [15] which uses modal laughter perceptions to investigate the automatic recognition of laughter type based on features of body movement. A total of 6802 laugh judgements of the 126 laughs produced a mean participant categorisation level of 28.85 laughs correctly categorised, with a standard deviation of 6.1. Using a single sample t-test against a chance level of 25.2 (20% level for each judgement times 126 judgements) participants showed an ability to categorise based on only body motion at better than chance levels $t(53) = 4.4$, $p < .001$. Mean performance was 51.4 with a standard deviation of 17.5. Actual judgement count figures for classification are: actually female: 5561; actually male: 1241; categorised as female: 2449; categorised as male: 4350. Despite the goal of creating an androgynous stick figure it seems that it is much more likely to be categorised as male. It may be the case that use of a point light display [6] induces less gender bias than the stick figures used in this experiment, this is an issue for future research. Figure 3 displays the performance of individuals on this task. Again we have highlighted the individuals who scored over the pathological threshold for gelotophobia and again there seems to be little relationship with ability to categorise laugh type.

Even though performance on both of these tasks was low we found a strong correlation between those participants who performed well at the categorisation of laugh type task and those who performed well at the categorisation of sex $r = .45$, $p = .001$. Figure 4 displays a scatterplot showing this relationship. So although people do not seem to be good at these tasks there does seem to be a common ability that is shared across these tasks. A further way to assess ability that is not dependent on the subjective judgement of the ground truth categorisation is to look for inter-rater agreement in the categorisation (using Fleiss’ kappa). As might be expected in a situation with a weak ability there is little inter-rater agreement in the overall group $\kappa = 0.06$, however if we only assess the inter-rater agreement amongst the best performing participants in the categorisation tasks—the six participants in the top right of Figure 4—there is a slight but present level of agreement $\kappa = 0.12$, this is significantly different from a value of kappa derived by subsampling 6 randomly selected raters 10,000 times which gives a mean $\kappa = 0.04$ with 99% confidence intervals at upper $= 0.041$, lower $= 0.039$. 
\end{itemize}
stages when social cognition and social interpretation factors become important. There are a number of caveats here. This is only the perception of body motion, other social signals such as facial expression and acoustic signals may carry more perceptual weight. Alternatively it may be that there is not enough information in any one signal alone but if they are provided in combination perceptual interpretation may differ.

C. Intensity

There is no relationship between the rated intensity of a laugh and the ability to categorise the sex of the laughter. Logistic regression and generalized additive mixed models were used to assess the relationship between rated intensity and the ability to categorise the type of laugh. This showed a small but significant effect, however the effect size was so small that we do not report the results of these analyses.

If we set aside the relationship with a ground truth categorisation and examine only the perceived categorisation and its relationship with rated intensity we find a very strong relationship; suggesting that when there is only body motion information regarding a laugh, raters rely heavily on intensity information to make their judgements.

An analysis of variance (ANOVA) shows a strong effect of difference between categories when intensity is the dependent variable F(4,5846)=1444.95, p<.0001, with a large effect size $\eta^2 = 0.497$. This means that approximately 50% of the variance in intensity is explained by the categorisation. This has no relationship with ability to correctly categorise the laughs, this simply assesses the relationship between rated intensity and the categorisation the rater chose. Figure 5 displays the relationship between category and intensity (error bars represent 95% confidence intervals).

Again we make no claims in this paper to know what the specific movements are that the participants are using to rate intensity; however, [15] provides insights concerning the specific movements that can be used to categorise laughter type, using a different set of raters; these movements may be candidates for informing the perception of laughter intensity.

IV. Conclusion

The pattern of results that we find in the human perception of context-free stimuli of whole body motion data of laughter shows a slight ability to categorise laughter in the same way that they are categorised by a rater in possession of the full social context, auditory, and visual cues associated with the laughter. The ability to correctly identify the sex of the laugh is also weak and there was a strong bias towards identifying the stick figure stimulus used in these experiments as male rather than female. Despite the weakness of the two categorisation effects there was a strong relationship between performance on the two categorisation tasks. This pattern of converging evidence suggests that there is a weak but real ability to use laughter body motion information informatively, at least in some of the raters.

The question of personality factors was more straightforward, as there was largely no influence of personality factors on either categorisation ability or the chosen category. The exceptions were gelotophiles—who were more likely to
correctly classify laugh sex—and also that positive emotional contagion traits were positively related to a likelihood of rating a laugh stimulus as hilarious and, to some extent, not rating it as social. These are perhaps not surprising results. The surprise was that there seemed to be no relationship in either classification ability or preferred categorisation type for gelotophobes. We can draw few conclusions here concerning gelotophobes: other than that they do not seem to be strongly influenced by body motion laughter—some performed at the higher end of the distribution in the classification task and they showed no preference for classifying laughter as one type or another. The evidence from this paper suggests that decisions concerning the nature of a laugh as malicious or benevolent does not involve much information from body movement. Whether this suggests an interpretation that decisions of these nature involve social cognition and not perceptual level features would require analysis of the other important perceptual components in laughter, the acoustic features and facial expressions.

The clearest finding of these studies comes from the relationship between categorisation and ratings of intensity, when the issue of comparison with a ground truth is ignored. There is a strong relationship between the ratings provided for intensity and the chosen categorisation. This would suggest that, in circumstances in which there is a reduction in information, the participants use the intensity of movement information to make their categorisation decision. Alternatively this could be viewed as categorisation decision strongly influencing the rating provided for intensity. The experimental design does not allow these two options to be distinguished. This also does not tell us what kind of movement information raters are using to make these judgments, and further research is required to assess the contributions of various types of movement to an intensity rating and to categorisation of laugh type.

**Fig. 5.** The relationship between intensity ratings and categorisation of laughter type. Error bars represent 95% confidence intervals. The NA category includes the instances where an intensity rating was provided without a laugh type categorisation.

**References**


