

Measuring Human Comprehension from Nonverbal Behaviour using Artificial Neural Networks

Fiona J. Buckingham, Keeley A. Crockett, Zuhair A. Bandar, James D. O'Shea

The Intelligence Systems Group, School of Computing,
Mathematics and Digital Technology
Manchester Metropolitan University
Chester Street, Manchester, M1 5GD, UK
fiona.j.buckingham@stu.mmu.ac.uk

Kathleen. M. MacQueen, Mario Chen

Family Health International 360
Durham, NC 27713, USA
KMacQueen@fhi360.org

Abstract— This paper presents the adaptation and application of Silent Talker, a psychological profiling system in the measurement of human comprehension through the monitoring of multiple channels of facial nonverbal behaviour using Artificial Neural Networks (ANN). Everyday human interactions are abundant with almost unconscious nonverbal behaviours accounting for approximately 93% of communication, providing a potentially rich source of information once decoded. Existing comprehension assessments techniques are inhibited by inconsistencies, limited to the verbal communication dimension and are often time-consuming with feedback delay. Major weaknesses hinder humans as accurate decoders of nonverbal behaviour with being error prone, inconsistent and poor at simultaneously focusing on multiple channels. Furthermore, human decoders are susceptible to fatigue and require training resulting in a costly, time-consuming process. ANNs are powerful, adaptable, scalable computational models that are able to overcome human decoder and pattern classification weaknesses. Therefore, the neural networks computer-based Silent Talker system has been trained and validated in the measurement of human comprehension using videotaped participant nonverbal behaviour from an informed consent field study. A series of experiments on training backpropagation ANNs with different topologies were conducted. The results show that comprehension and non comprehension patterns exist within the monitored multichannels of facial NVB with both experiments consistently yielding classification accuracies above 80%.

Keywords- *artificial neural networks, backpropagation, human comprehension, nonverbal behaviour, silent talker.*

I. INTRODUCTION

The Human Immunodeficiency Virus (HIV) is a virus that attacks the human immune system, which gradually leads to the final stage of the infection known as the Acquired Immune Deficiency Syndrome (AIDS) where the damaged human immune system fails and is unable to fight other infections [1]. HIV can be prevented but the transmission of HIV still occurs and so the fight against reducing the global epidemic of AIDS/HIV remains an ongoing challenge. Alone, sub-Saharan Africa contains an inordinate majority of HIV infections with approximately 22.5 million people living with HIV and

approximately 1.8 million adults and children newly infected with HIV in 2009 as reported by the 2010 UNAIDS report on the Global AIDS Epidemic Update [2]. Moreover, in sub-Saharan Africa more females are infected with HIV than males and that women aged 15–24 years are up to eight times more likely to be HIV positive [2]. Therefore, sub-Saharan African females are especially vulnerable to HIV and as a result comprehension of the disease needs to be improved to help reduce further outbreaks.

HIV prevention clinical trials are often executed in developing countries such as Africa where the prevalence of HIV/AIDS is high. Within trials the informed consent process is a legal and ethical requirement, requiring each participant to voluntarily make a truly informed decision on whether to participate in the study. Informed consent requires voluntary informed consent from comprehension of adequately delivered information about the purpose of the study, any procedures involved and the effects of participation [3] before commencement of participation. Although the Nuremburg Code [4], the Declaration of Helsinki [5] and the Belmont Report [6] exist to provide guidelines to ensure that research is ethical, the quality of the assessment of participant's comprehension during the informed consent process still remains a critical area of concern. Previous studies have identified that participants have difficulties in understanding informed consent documentation [7], [8], [9], which can impair decision making. Significant differences have also been found to exist between common comprehension assessment methods [10]. Therefore, the research presented in this paper has focused upon the development of an improved automated method in the measurement of human understanding during the informed consent process in an HIV informed consent study with female, African participants.

Measurement of human comprehension has primarily been focused on verbal and written responses [11] with little attention being paid to the nonverbal dimension. Throughout everyday interaction humans are frequently exposed to nonverbal behaviours (NVB) that contain a rich source of information once deciphered e.g. facial expressions. Therefore, there is a niche for the exploration to identify whether patterns

of comprehension and non comprehension exist within NVB alone.

Artificial Neural Networks (ANN) are non-linear parallel computational models comprised of interconnected processing nodes [12]. Experiential knowledge is contained within the neural networks interconnections, which is obtained through the presentation of patterns using training and validation algorithms. ANNs have been adapted to a range of problem domains such as pattern classification [12], [13], image and speech recognition [14], financial analysis [13] and control systems [13]. Application of ANNs in the measurement of human understanding during informed consent could provide a new, more reliable technique of measuring comprehension in comparison to existing comprehension assessment methods.

Silent Talker [15] is an ANN-based psychological profiling system that monitors multiple channels of facial NVB to detect truthful and deceptive human behavioural patterns has the potential ability to be adapted to other human states in different environments. This paper discusses the adaptation and application of *Silent Talker* in the detection of human comprehension and non comprehension from multichannels of NVB using ANN. This research is centred upon addressing the measurement of comprehension during the informed consent assessment process in a HIV/AIDS prevention field study carried out in North-western Tanzania, Africa.

Overall, the aim of this research is to identify whether human comprehension can be measured from NVB in near real-time using an adapted version of the *Silent Talker* ANN-based system. If the ANNs are able to detect nonverbal comprehension patterns then it could improve and positively impact on the informed consent process in many ways. Moderators of informed consent process would be able to use the tool to identify whether participants comprehended specific features of the informed consent process such as questions in the questionnaire and be able to dynamically adapt their approach until the participant adequately understands the question. As a result, existing comprehension assessment techniques and informed consent documents should dramatically improve as problematic comprehension areas would be identified almost instantly to the moderator allowing immediate correction application. The participants would benefit from enhanced, tailored education from the moderators during the informed consent process, which would facilitate higher comprehension levels and enable execution of more informed, voluntary consent decisions. Due to better education the participants would have an improved understanding of HIV and the prevention methods available, which could potentially lead to a reduction in HIV transmission and the size of the epidemic. Furthermore, other clinical trials could also reap the benefits of comprehension assessment tool through implementation during the informed consent process.

This paper continues as follows: Sections II, III and IV review NVB, comprehension and adapting *Silent Talker*. Section V presents the methodology with the experimental results and discussion contained in Section VI. The conclusion and future work is in Sections VII and VIII.

II. NONVERBAL BEHAVIOUR

NVB, is a form of non-linguistic communication that can accompany verbal responses or standalone. Examples of NVB are facial expressions, gestures, body posture, body movement and vocal cues lacking verbal content [16]. With NVB under less conscious control [17] and forming approximately 93% of communication [18] it could provide a more accurate, reliable measure of human state than verbal responses.

The nonverbal response is available for assessment by the listener prior to the verbal response [19]. Thus, potentially leaking early signals to the listener about the sender's state whilst a verbal response is still being formulated. Even during and after the verbal response, NVB's would be available for analysis providing a large data set to use in the identification of patterns of human state. The analysis of NVB should not only be limited to the presence of NVB cues but include their absence [20] too, as it could have a significant effect upon overall interpretation of the message.

When the verbal response is available alongside the nonverbal response, it would provide additional vocal cues that could be utilised in the decoding of the expressed emotion. Sauter *et al.* [21] demonstrated that acoustic cues from nonverbal vocalisations such as amplitude and pitch could be used to predict participants rating of emotion from the vocal expression alone. Vocalisations are not considered in this research as the prime focus is upon NVB only.

With there being a large repertoire of NVB channels available [22], researchers have the opportunity to collate data on individual or collections of NVB channels to identify patterns, associated with human state. Focusing upon the most appropriate channel(s) for detecting human state e.g. NVB's associated with deceptive behaviour requires knowledge and experimentation. During face-to-face communication humans tend to primarily focus upon the face as the area contains a rich source of information. Allocating higher precedence to the face than any of the other available channels is regarded as "facial primacy" [20]. Approximately, 55% of the communicated message by a person can be obtained from facial behavioural expressions [18] thus making the face a prime area of interest in the monitoring of human states. Mehrabian [23] argued that even when facial behaviour is not consistent with speech that the listener will be most impacted by the facial emotion expressed. Furthermore, that when facial behaviour is absent such as during a telephone call that the vocal behaviour would take precedence over the words spoken during inconsistencies. Therefore, there is a great weighting toward NVB being a more reliable source than verbal responses.

To classify NVB, human judges have often been employed to code the NVB channels [24] that have been selected for monitoring. Humans are error prone, susceptible to fatigue and poor at multitasking thus inhibiting them as accurate judges in the classification of nonverbal channels. Moreover, human decoders are not always consistent, requiring time and capital to train.

The Facial Action Coding System (FACS) [25] is a common standard that has been adopted by multiple human decoders in order to aid and ensure reliability in the manual

process of categorising facial expressions to measure emotions from NVB. Encapsulated within the FACS there are Action Units (AU) that have been coded to represent a contraction of individual or groups of muscles, which can then be utilised in scoring the facial expression as an emotion. Although the FACS standard exists to ensure consistency, the manual process of encoding NVB still remains a very time-consuming task when using human decoders. More time is required as the number of monitored NVB channels increases and the quantity of video data raises, especially when taking a frame-by-frame analysis approach, resulting in great expense and delayed findings.

Nowadays, researchers have overcome the weaknesses of manually encoding and decoding NVB through the implementation of automated computer-based solutions to enable consistent real-time processing and classification of multiple NVB channels [26]. Multiple techniques have been used to capture and monitor NVB channels, ranging from invasive to non-invasive methods such as photographs, video recordings and sensors within natural and artificial settings such as laboratories. The environment in which the assessment of NVB takes place can severely affect the natural quality of NVB produced, potentially introducing bias and affecting the reliability of the data. Photographs and videos are often used to capture NVB for analysis [27]. Photographs are weak as they only represent the static emotion expressed at a fixed point in time, are often unnatural due to being posed expressions, can miss micro-expressions, pre- and post-expressions of human state. Videos on the other hand are able to overcome the weaknesses of photographs, capturing the natural occurrences of spontaneous behaviour over time.

Overall, nonverbal communication encapsulates a large source of frequently available NVB's in a wide range of formats for analysis and overcomes the weaknesses of verbal behaviour. Interpretation of NVB's has been greatly improved through the use of computer-based technologies, overcoming human decoder weaknesses. Therefore, great potential exists in the development of computer-based programs for use in the detection of human states such as comprehension using NVB.

III. COMPREHENSION

Non comprehension is regarded as “a state of knowledge that ranges from uncertainty to complete lack of understanding of the materials under discussion” [28]; the opposite of understanding. The research in this paper is focused upon measuring human comprehension during informed consent. The informed consent process ensures that participants make a truly informed, educated decision through clear comprehension of the information delivered in the informed consent process otherwise misconceptions are liable to occur and decision related implications may arise. To ensure valid, ethical informed consent participants should receive adequate information on the nature of the research, the purpose of the study, any potential risks, participation benefits, alternatives and have a full understanding [29] of the concepts prior to consenting to participate.

Informed consent documents are often long, covering a lot of points in detail, which can inhibit comprehension through

infliction of information overloading [20]. Comparison of comprehension assessment methods: self-reports, checklists, vignettes and narratives has identified significant differences on the measurement of understanding [13]. Paris *et al.* [31] found that even after improving an informed consent document through using a working group and by enhancing the lexicosyntactic readability through sentence length reduction and shorter synonyms that neither technique significantly enhanced participant's objective comprehension score thus emphasising the need for an improvement in comprehension assessment techniques. Through early identification of low comprehension such as during the informed consent process it would provide time to adapt content delivery to facilitate understanding, thus avoiding legal and ethical implications.

Previous research on the detection of comprehension has had a dominant focus upon verbal and written responses [14]. This emphasises the need to analyse from the nonverbal angle. With NVB forming such a large proportion of communication and under less conscious control it has the potential to deliver a more accurate measure of human comprehension in comparison to verbal communication. Although NVB is susceptible to self-monitoring [32] where human behaviour is controlled in a deceptive manner to enable the person to fit in the social environment, it has been found that nonverbal difficulty displays of behaviour exhibited during problem-solving tasks were more easily distinguishable from low self-monitoring students than high self-monitors [33]. Furthermore, through observation of nonverbal videotaped behaviour of children participating in a lesson on electricity containing difficult and easy content, individual observers were able to distinguish children comprehending and not comprehending from the deliberate and spontaneous NVB behaviours alone [34]. Moreover, it was also found that the observers were able to identify high achieving children as having a greater understanding than the low achieving children in both the easy and hard environments [34]. Therefore, it can be acknowledged that comprehension and non comprehension patterns do reside within NVB alone but warrants further research as details of the NVB are often omitted and few channels are analysed. Therefore, potential distinct patterns of NVB relating to the level of understanding could be discovered.

Multiple classroom studies have identified NVB patterns associated with understanding and non comprehension [35], [36], [37], [38] with facial behaviour, general hand and body movements predominantly being monitored using human decoders. However, informed consent studies have been limited to primarily independent predictors of comprehension from known data such as age [31], educational level [29], race and ethnicity [39]. Therefore, there is the unexplored need to identify whether more complex, general trends are present within larger experimental groupings of multiple NVB channels including independent predictors.

Development of a near real-time computer-based application to monitor multiple NVBs for comprehension measurement would overcome human weaknesses and help reduce the complexities of measuring understanding.

IV. ADAPTING SILENT TALKER

Silent Talker [15] is an ANN-based software system that successfully monitors multichannels of facial NVB for psychological profiling. Through a simulated theft scenario, *Silent Talker* was able to detect truthful and deceptive NVB with 80% accuracy [15].

When using the patented [40] near real-time *Silent Talker* system, the video camera is focused so that the interviewees NVB facial cues can be captured for analysis by *Silent Talker*, operated by the interviewer. During the interview the monitored multichannels of NVB in the video frames are analysed by ANNs and at the end of the assessment of the video segment an overall deceptive/truthful classification is outputted to the interviewers display.

Within *Silent Talker* the object locators ANNs identify the location of nonverbal features such as the eye. Once located, the pattern detectors ANNs identify the state of the object such as the right eye gazing to the left. Each of the monitored NVB channels is coded, collated together for the entire answer and provided as input for the classifier ANN to determine the deceptive/truthful assessment of the response.

Currently, within *Silent Talker* the standard set up is 40 channels of NVB available covering the: left and right eye gaze and closure, blanching, blushing, slot (collection time), head movement, angle and rotation and known data such as the participant's gender [41].

Although, *Silent Talker* was originally, specifically designed as a "lie detector" it still has the unexplored, capability of being adapted to other unmonitored areas of human state where NVB patterns are of great interest. Adaptation to a new classification of behaviour using the existing NVB channels would only require retraining of the ANN classifiers. Furthermore, the introduction of new NVB channels and the pruning of existing NVB channels could also be implemented to extend and enhance the adaption of the ANN-based system to a new behavioural state.

Therefore, the research presented in this paper has adapted and applied the classifier ANN within *Silent Talker* through training and validation techniques to measure human comprehension from NVB in an informed consent field study. This paper shall identify whether such a complex ANN-based system can be adapted to measure human comprehension and non comprehension from NVB's.

V. METHODOLOGY

A. Field Study Background

The field study was executed in the Mwanza region of Tanzania in Africa by Family Health International 360 (FHI 360) [42] in collaboration with the National Institute for Medical Research (NIMR) [43]. Eighty female participants between the ages of eighteen and thirty five who had not participated in a clinical trial before were recruited to participate in a video recorded interview on learning task topics of high and low comprehension. A digital camera was used to record the head shot of the participant during the entire interview, generating the Audio-Video Interleaved (AVI) files

for analysis. Ethical approval was obtained from NIMR's Medical Research Coordinating Committee and from the Protection of Human Subjects Committee at FHI 360 and from Manchester Metropolitan University's (MMU) Faculty Academic Ethics Committee.

B. Interview Design

Task A was the high comprehension topic, designed to be easy to understand covering condom usage. Task B was the low comprehension topic, intended to be hard to understand on HIV mutation and treatment effectiveness. The purpose of difficult topic was to elicit high levels of distinct non comprehension NVB's and the easy topic was designed to induce high frequencies of pure comprehension NVB's. The two tasks developed by FHI 360 were necessary to enable the distinct capturing of each type of NVB to reduce noise levels in the training data set. After listening to the short learning task script the participant received the associated ten closed and open-ended questions with randomisation applied to determine which set of questions were to be encountered first. Task order was also randomised so that half of the participants completed task A followed by task B and vice versa. Implementation of randomisation was intended to enable the analysis of any potential relationships between the NVB patterns, task order and question set order to be highlighted.

The resulting participant video files from the learning interview were used to obtain the ANN training data set for use in the experiments with different training calibrations.

C. ANN Classifier Training Procedure

Normalised input vector patterns scaled from +1.0 to -1.0 for each participant's learning task interview were obtained from segments of the AVI file under experimental analysis. Each scaled value in a vector represented the state of a monitored NVB as an input for the ANN during training. For example, the left eye was coded as +1.0 when it was closed and -1.0 when open. A supervisory value was appended to the vectors, coded as +1.0 for comprehension and -1.0 for non comprehension to independently assess whether the ANN correctly classified the presented input vector pattern. The normalised vectors were randomly split in to training (Tr), validation (Va) and testing (Te) data sets.

Silent Talker's feedforward backpropagation classifier ANN was trained with randomised starting weights using incremental weight updating with the delta rule [44], [45], [46]. Presentation of the Tr, Va and Te sets during training were randomised to ensure that the neural network did not learn patterns on an order basis and to aid convergence speed.

The Tr set was used to update the weights based upon the calculated error. Va determined when to stop training with the trained weights frozen and Te determined the neural networks ability at generalising on an unseen data set. Classification accuracy (CA) was used to evaluate the performance of an ANN as a percentage of the total number of patterns that were correctly classified out of the presented desired data set. Equation (1) contains the comprehension accuracy (Comp CA) formula where CCC denotes correct comprehension

classifications and DCC denotes desired comprehension classifications.

$$Comp\ CA = (no.\ of\ CCC / no.\ of\ DCC) * 100 \quad (1)$$

The non comprehension accuracy (NonComp CA) calculation is displayed in (2) where NCC denotes correct non comprehension classifications and DNC denotes desired non comprehension classifications.

$$NonComp\ CA = (no.\ of\ NCC / no.\ of\ DNC) * 100 \quad (2)$$

To obtain the overall CA, the comprehension and non comprehension percentages from (1) and (2) were combined in (3).

$$CA = (Comp\ CA + NonComp\ CA) / 2 \quad (3)$$

n -fold cross-validation [47] was implemented where the data set is randomly split into n segments with one segment retained for testing purposes and the remaining segments used in training and validation. Training ends when all of the n segments have completed the role as the test set once.

VI. RESULTS AND DISCUSSION

The study yielded thirty usable videos resulting in 600 answers along with the two video segments capturing NVB during the reading of each learning task script. Fifty videos were discarded from the data set due to poor quality of the recordings.

In each of the ANN training experiments the ANNs had the same topology: forty inputs, a single hidden layer of thirty neurons and a single output neuron. The topology was chosen because it was the best performing neural network from previous experimental training configuration sessions. The inputs to the ANNs were the same as the nonverbal input channels used in the original *Silent Talker* experiments. Prior to the 10-fold cross-validation commencement the data set partitioned as Tr (60%), Va (30%) and Te (10%). The individual experiments and their results shall be discussed.

A. Experiment 1: Script Readings

The training data set for Experiment 1 came from the video segments containing the entire reading of the task A and task B scripts. Experiment 1 yielded 241,945 vectors with a near balanced data set: script A = 45% labelled comprehension and script B = 55% labelled non comprehension. During the reading of script A participants understanding was predicted to be high with high frequencies of NVB comprehension patterns. Understanding was predicted to be low with high levels of associated NVB non comprehension patterns during the reading of script B.

Table I shows the CA's for each phase of the 10-fold cross-validation iterations. Iteration 5 of the 10-fold cross-validation generated the best performing ANN with an overall testing CA of 89.29%. Table II and III display a breakdown of the

comprehension and non comprehension CA's for each training phase. All of the CA's within the Tr, Va and Te phases of training have consistently achieved CAs greater than 85%, strongly indicating that comprehension and non comprehension patterns exist within the data set.

Through examination of the participants marked responses to the open-ended questions for script A and B were found to not be purely comprehension responses for task A and not entirely non comprehension answers for task B. Thus, the two data sets were too noisy. The next experiment implemented a technique to try and overcome the latter weakness.

TABLE I. EXPERIMENT 1: CROSS-VALIDATION CA'S

Iteration	Training CA	Validation CA	Testing CA
1	88.35%	86.88%	87.15%
2	89.25%	87.69%	87.82%
3	89.46%	88.06%	88.53%
4	89.74%	88.50%	88.50%
5	90.05%	88.95%	89.29%
6	90.50%	89.06%	89.09%

TABLE II. EXPERIMENT 1: CROSS-VALIDATION COMPREHENSION CA'S BREAKDOWN

Iteration	Training CA	Validation CA	Testing CA
1	89.89%	88.55%	88.81%
2	89.73%	87.93%	87.80%
3	89.87%	88.44%	88.68%
4	89.88%	88.63%	88.69%
5	88.85%	87.41%	87.91%
6	90.72%	88.98%	89.69%

TABLE III. EXPERIMENT 1: CROSS-VALIDATION NON COMPREHENSION CA'S BREAKDOWN

Iteration	Training CA	Validation CA	Testing CA
1	86.80%	85.22%	85.49%
2	88.78%	87.45%	87.84%
3	89.06%	87.67%	88.39%
4	89.59%	88.38%	88.32%
5	91.26%	90.48%	90.67%
6	90.27%	89.13%	88.49%

B. Experiment 2: Script Reading Points Mapped to the Open-ended Questions

Experiment 2's data set consisted of the reading of the individual sentences within each of the scripts being labelled with a supervisory value through mapping the participants associated open-ended question response to the individual script point. Therefore, if the participant answered the question

correctly the corresponding script point vectors would all have a +1.0 supervisory value indicating comprehension and -1.0 when non comprehension. The close-ended questions were not selected to be mapped to the script points due to being strongly inhibited by the susceptibility to guessing reducing results to chance levels. Experiment 2 yielded a data set containing 71,787 vectors: 63.5% comprehension and 36.5% non comprehension.

The CA's for each phase of the 10-fold cross-validation iterations is displayed in Table IV. Iteration 3 of the 10-fold cross-validation generated the best performing ANN with an overall testing CA of 87.05%. Table V and VI contain the comprehension and non comprehension CA breakdown for each training phase. The ANNs consistently achieved CAs above 81%, indicating that comprehension and non comprehension patterns exist within the data set.

TABLE IV. EXPERIMENT 2: CROSS-VALIDATION CA'S

Iteration	Training CA	Validation CA	Testing CA
1	88.59%	84.25%	84.55%
2	89.74%	85.42%	85.07%
3	90.47%	85.62%	87.05%
4	89.52%	85.39%	85.87%
5	89.43%	84.83%	85.54%
6	88.33%	84.17%	85.02%

TABLE V. EXPERIMENT 2: CROSS-VALIDATION COMPREHENSION CA'S BREAKDOWN

Iteration	Training CA	Validation CA	Testing CA
1	89.52%	85.52%	86.34%
2	89.92%	87.12%	87.09%
3	90.54%	86.97%	88.27%
4	89.20%	86.05%	86.40%
5	88.28%	84.50%	84.52%
6	89.53%	87.00%	87.31%

TABLE VI. EXPERIMENT 2: CROSS-VALIDATION NON COMPREHENSION CA'S BREAKDOWN

Iteration	Training CA	Validation CA	Testing CA
1	87.67%	82.97%	82.75%
2	89.57%	83.73%	83.04%
3	90.40%	84.27%	85.84%
4	89.85%	84.73%	85.34%
5	90.57%	85.16%	86.55%
6	87.13%	81.34%	82.72%

Later inspection of the participants marked open-ended questions found that 70% of responses were correct resulting in high levels of comprehension.

C. Summary of Experiments

Both experiments consistently attained CA's grouped above 80% in the Tr, Va and Te phases of the cross-validation from randomised starting weights and vector presentation. Therefore, the latter results strongly indicate that multichannel NVB patterns of comprehension and non comprehension do exist within the data set.

Silent Talker performed well at classifying the previously unencountered ethnic data set of Tanzanian women. However, retraining of the pattern detector and object locator ANN's with the unseen ethnic data set would enhance *Silent Talker*'s ability at locating the individual facial features and their state, which would increase the CAs. Furthermore, improvements in the quality of recordings would lead to a larger training data set with lower levels of discarded videos thus potentially increasing the applicability of the results.

From executing the ANN training experiments, it has identified key elements of the field study that can be improved and applied to future work in order to enhance the quality of the data set. Firstly, the style of the open-ended questions and the learning task scripts should utilise techniques to increase the difficulty of task content to raise participant's non comprehension levels to ensure that a more balanced data set is obtained. For instance, question six in task B was "Please name at least one of the two major HIV viral types." and the corresponding script point was "These include two major viral types (HIV-1 and HIV-2)". Therefore, increasing the complexity of the tasks would make recall and guessing more difficult for participants, ensuring clearer comprehension and non comprehension data is collated. Introduction of a time delay between script reading and questioning could help to minimise the lack of deeper understanding through the production of recalled script keywords thus enhancing the quality of the responses. Rushing points, overloading the receiver with complex information and using jargon have been found as causes of patient misunderstanding during clinician and patient communication [48] thus providing additional techniques that could be used to impair comprehension. Application of the improved techniques should provide a more balanced data set with low noise levels resulting in clearer comprehension and non comprehension data sets for experimentation.

VII. CONCLUSION

Application of *Silent Talker* to a new environment in the role as a "comprehension detector" during the informed consent process has proven that it is possible to adapt such a complex ANN-based system to a different human state in a new environment. The experimental results strongly indicated that detectable patterns of comprehension and miscomprehension exist within the monitored facial NVB multichannels of the data set as the cross-validation consistently attained CA's grouped above 80% from randomised weights and vector presentation. The results are limited to African women and so further experiments with participants of different ethnicities, ages and genders are necessary to enhance the reliability of the results and to confirm whether general patterns of comprehension exist

outside of the demographics contained in the field study population.

Through adapting *Silent Talker* in the measurement of human understanding during informed consent it has also identified features of the field study and components of *Silent Talker* that can be enhanced in future work to improve the accuracy of comprehension measurement and the quality of the training data set.

VIII. FURTHER WORK

This paper presents a potential comprehension detector prototype that could be used as a proxy tool alongside existing comprehension assessment techniques to aid understanding during the informed consent process in near real-time such as identifying problematic questions that need improving. Future work includes the introduction of new NVB channels and the pruning of existing NVB channels to extend and enhance the adaption of the ANN-based system to the new behavioural state.

Further applications of a comprehension monitoring system are numerous. In education, it would provide teachers with a quick and easy method of gauging the student's level of understanding, allowing them to change the lesson content to suit the needs of the individuals without great delay thus saving time for further progression, which would have not been possible before. Within the medical environment it could be used to ensure that misconceptions during diagnosis are rectified and to make sure that patients fully understand treatments and their prescriptions.

ACKNOWLEDGMENT

We wish to thank the women in Tanzania who participated in this research and especially for their willingness to be video-recorded. We also wish to thank our collaborators at NIMR in Tanzania, including Bahati Andrew, Tusa Erio, Catherine Bunga, and Gerald Lumanyika.

REFERENCES

- [1] WHO, 2011, HIV/AIDS Fact Sheet [Online] Available: <http://www.who.int/mediacentre/factsheets/fs360/en/index.html>
- [2] UNAIDS, 2010, Global report: UNAIDS report on the global AIDS epidemic 2010 [Online] Available: http://www.unaids.org/globalreport/documents/20101123_GlobalReport_full_en.pdf
- [3] G. Lindegger, "Informed consent in HIV vaccine trials" In P. Kahn, AIDS Vaccine Handbook: Global Perspectives, 2nd Ed., New York: AIDS Vaccine Advocacy Coalition, pp.109-116, 2005.
- [4] Nuremberg Code. In: Trials of war criminals before the Nuremberg military tribunals under control council law, vol. 11. Washington, DC: U.S. Government Printing Office, No. 10, 1949. p. 181-2.
- [5] World Medical Association (WMA). Declaration of Helsinki: ethical principles for medical research involving human subjects, as amended by the 52nd WMA General Assembly, Edinburgh, Scotland; 2000.
- [6] National Commission for the Protection of Human Subjects of Biomedical, Behavioral Research. The Belmont report: ethical principles and guidelines for the protection of human subjects of research. Washington, DC: U.S. Government Printing Office; 1979.
- [7] D. W. Fitzgerald, C. Marotte, R. I. Verdier, W. D. Jr. Johnson and J. W. Paper, "Comprehension During Informed Consent in a Less-developed Country", *Lancet*, vol. 360, pp. 1301-1302, 2002.
- [8] J. Flory and E. Emanuel, "Interventions to Improve Research Participants' Understanding in Informed Consent for Research: A Systematic Review", *JAMA*, vol. 292(13), pp. 1593-1601, 2004.
- [9] A. R. Oduro, R. A. Aborigo, D. Amugsi, F. Anto, T. Anyorigiya, F. Atuguba, A. Hodgson and K. A. Koram, "Understanding and Retention of the Informed Consent Process Among Parents in Rural Northern Ghana", *BMC Medical Ethics*, vol. 9(12), 2008.
- [10] G. Lindegger, C. Milford, C. Slack, M. Quayle, X. Xaba and E. Vardas, "Beyond the Checklist: Assessing for Understanding for HIV Vaccine Trial Participation in South Africa", *Acquired Immune Deficiency Syndrome*, vol. 43(5), pp. 560-566, 2006.
- [11] M.E. Falagas, I.P. Korbila, K.P. Giannopoulou, B.K. Kondilis, and G. Peppas, "Informed consent: how much and what do patients understand?," *American Journal of Surgery*, vol. 198(3), pp. 420-435, 2009.
- [12] M. H. Hassoun, *Fundamentals of Artificial Neural Networks*, London: MIT Press, 1995.
- [13] B. Widrow, D. E. Rumelhart and M. A. Lehr, "Neural networks: applications in industry", *Business and Science. Commun. ACM*, vol. 37(3), pp. 93-105, 1994.
- [14] K. Gurney, *An Introduction to Neural Networks*, London: UCL Press, 1997.
- [15] J. Rothwell, Z. Bandar, J. O'Shea and D. McLean, "Silent Talker: A New Computer-based System for the Analysis of Facial Cues to Deception", *Applied Cognitive Psychology*, vol. 20, pp. 757-777, 2006.
- [16] E. Babad, "Teaching and Nonverbal Behavior in the Classroom" In L. J. Saha, and A. G. Dworkin (Eds.) *International Handbook of Research on Teachers and Teaching*, Boston, Massachusetts: Springer US, pp. 817-827, 2009.
- [17] T. P. Mottet and V. P. Richmond, "Student Nonverbal Communication and its Influence on Teachers and Teaching" In J. L. Chesebro & J. C. McCroskey (Eds.), *Communication for Teachers*, Needham Heights, Massachusetts: Allyn and Bacon, pp. 47-61, 2002.
- [18] A. Mehrabian, "Communication Without Words", *Psychology Today*, vol. 2(4), pp. 53-56, 1968.
- [19] V. Manusov and A. R. Trees, "Are You Kidding Me?": The Role of Nonverbal Cues in the Verbal Accounting Process", *Journal of Communication*, vol. 52(3), pp. 640-656, 2002.
- [20] M. L. Knapp and J. A. Hall, *Nonverbal Communication in Human Interaction*, 3rd Ed, Fort Worth: Harcourt Brace, 1992.
- [21] D. A. Sauter, F. Eisner, A. J. Calder and S. K. Scott, "Perceptual cues in nonverbal vocal expressions of emotion", *The Quarterly Journal of Experimental Psychology*, vol. 63(11), pp. 2251-2272, 2010.
- [22] P. Ekman and W. V. Friesen, "The Repertoire of Nonverbal Behavior Categories, Origins, Usage, and Coding", *Semiotica*, vol. 1, pp. 49-98, 1969.
- [23] A. Mehrabian, *Silent Messages*, 5th Ed., California: Wadsworth Publishing Company, 1971.
- [24] B. McDaniel, S. D'Mello, B. King, P. Chipman, K. Tapp and A. Graesser, *Facial Features for Affective State Detection in Learning Environments*, Proc. 29th Ann. Meeting of the Cognitive Science Soc, 2007.
- [25] P. Ekman and V. W. Friesen, *The Facial Action Coding System (FACS)*, Consulting Psychologists Press, Palo Alto, CA, US, 1978.
- [26] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan and A. Bigdeli "How do you know that I don't understand?" A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24(4), pp. 1342-1363, 2008.
- [27] Z. Zeng, M. Pantic, G. I. Roisman and T. H. Huang, "A survey of affect recognition methods: audio, visual and spontaneous expressions" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31(1), pp. 39-58, 2009.
- [28] H. Z. Waring, "Expressing noncomprehension in a US graduate seminar", *Journal of Pragmatics*, vol. 34(12), pp. 1711-1731, 2002.
- [29] J. T. Krankl, S. Shaykevich, S. Lipsitz and L. S. Lehmann, "Patient Predictors of Colposcopy Comprehension of Consent Among English- and Spanish-speaking Women", *Women's Health Issues*, vol. 21(1), pp.80-85, 2011.

- [30] M. Hochhauser, 2004, Informed Consent: Reading and Understanding Are Not the Same. *Applied Clinical Trials* [Online] Available: <http://appliedclinicaltrialsonline.findpharma.com/appliedclinicaltrials/article/articleDetail.jsp?id=90594>.
- [31] A. Paris, C. Brandt, C. Cornu, P. Maison, C. Thalamas and J. Cracowski, "Informed consent document improvement does not increase patients' comprehension in biomedical research", *British Journal of Clinical Pharmacology*, vol. 69(3), pp. 231-237, 2010.
- [32] M. Snyder, "Self-monitoring of Expressive Behavior", *Journal of Personality and Social Psychology*, vol. 30, pp. 526-537, 1974.
- [33] D. Hrubes and R. S. Feldman, "Nonverbal Displays as Indicators of Task Difficulty", *Contemporary Educational Psychology*, vol. 26, pp. 267-276, 2001.
- [34] V. L. Allen and M. L. Atkinson, "Encoding of nonverbal behavior by high-achieving and low-achieving children", *Journal of Educational Psychology*, vol. 70(3), pp. 298-305, 1978.
- [35] J. D. Jecker, N. MacCoby, and H. S. Breitrose, "Improving Accuracy in Interpreting Non-verbal Cues of Comprehension", *Psychology in the Schools*, vol. 2(3), pp. 239-244, 1965.
- [36] S. Machida, "Teacher Accuracy in Decoding Nonverbal Indicators of Comprehension and Noncomprehension in Anglo- and Mexican-American Children", *Journal of Educational Psychology*, vol. 78(6), pp. 454-464, 1986.
- [37] C. J. Patterson, M. J. Cosgrove and R. G. O'Brien, "Nonverbal Indicators of Comprehension and Noncomprehension in Children", *Developmental Psychology*, vol. 16(1), pp. 38-48, 1980.
- [38] E. Skarakis-Doyle, N. MacLellan and K. Mullin, "Nonverbal Indicators of Comprehension Monitoring in Language-disordered Children", *Journal of Speech and Hearing Disorders*, vol. 55(3), pp. 461-467, 1990.
- [39] A. S. Fink, A. V. Prochazka, W. G. Henderson, D. Bartenfeld, C. Nyirenda, A. Webb, D. H. Berger, K. Itani, T. Whitehill, J. Edwards, M. Wilson, C. Karson, and P. Parmelee, "Predictors of Comprehension during Surgical Informed Consent" *Journal of the American College of Surgeons*, vol. 210(6), pp. 919-926, 2010.
- [40] Z. Bandar, D. A. McLean, J. D. O'Shea and J. A. Rothwell, International Patent Number WO02087443. Geneva, Switzerland: World Intellectual Property Organization, 2002.
- [41] J. Rothwell, Z. Bandar, J. O'Shea and D. McLean, "Charting the Behavioural State of a Person Using a Backpropagation Neural Network", *Neural Computing & Applications*, vol. 16, pp. 327-339, 2007.
- [42] Family Health International 360, 2011, Who We Are. [Online] Available: <http://www.fhi.org/en/AboutFHI/index.htm>.
- [43] National Institute for Medical Research Tanzania, 2011, About Us [Online] Available: <http://www.nimr.or.tz/>.
- [44] P. Werbos, *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*, Wiley-Interscience New York, NY, USA, 1994.
- [45] P. Werbos, "Backpropagation through time: what it does and how to do it", *Proceedings of the IEEE*, vol. 78(10), pp. 1550-1560, 1990.
- [46] P. Werbos, "Beyond Regression: New tools for prediction and analysis in the behavioural sciences", Doctoral Dissertation, Appl. Math., Harvard University, Cambridge, MA, USA, 1974.
- [47] M. Stone, "Cross-validators Choice and Assessment of Statistical Predictions", *Journal of the Royal Statistical Society*, vol. 36(2), pp. 111-147, 1974.
- [48] H. F. West and W. F. Baile, "'Tell me what you understand': the importance of checking for patient understanding", *The Journal of Supportive Oncology*, vol. 8(5), pp. 216-218, 2010.