A Database for Facial Behavioural Analysis

Moi Hoon Yap
School of Computing, Mathematics, and Digital Technology
Manchester Metropolitan University
John Dalton Building, Chester Street
Manchester, M1 5GD, UK
M.Yap@mmu.ac.uk

Hassan Ugail
Centre for Visual Computing
University of Bradford
Bradford, BD7 1DP, UK
H.Ugail@bradford.ac.uk

Reyer Zwiggelaar
Department of Computer Science
Aberystwyth University
Aberystwyth, SY23 2AX, UK
rrz@aber.ac.uk

***Abstract*— there is substantial interest in detection of human behaviour that may reveal people with deliberate malicious intent, who are engaging in deceit. Technology exists that is able to detect changes in facial patterns of movement and thermal signatures on the face. However, there is data deficiency in the research community for further study. Therefore this project aims to overcome the data deficiency in psychology study and algorithms development. A within-subjects design experiment was conducted, using immigration as a scenario for investigate participants in control and experimental conditions. A random sample of 32 volunteers were recruited, their age group is within 18 – 33. The study design required participants to answer questions on two topics, one as themselves and one as a pre-defined character. Data regarding visible and thermal images of facial movement and behaviour were collected. A rich FACS-coded database with high quality thermal images was established. Finally, recommendations for development and subsequent implementation of the facial analysis technique were made.**

***Keywords*—Facial action units, FACS, behaviour, deception, thermal images.**

I. Introduction

An emerging theme of interest for security agencies is the detection of human behaviours that may reveal an individual as having deliberate malicious intent; for instance by attempting to deceive authorities to enter a country illegally, smuggle goods into or out of a country, being involved in a malicious act such as a terrorist bombing, or as harbouring the intention to carry out such a malicious act at a later time. Such a capability will aid in the apprehension of suspect individuals, before they are able to carry out malicious acts.

Technology exists that is able to detect changes in facial patterns and movement in both the visible and thermal fields. This motivation of this project is to exploit those capabilities by beginning development of a real-time dynamic passive profiling technique to assist security officers as a decision aid. Relevant literatures were reviewed to establish behaviours that might plausibly be used for the operational identification of malicious intent: modelling these behaviours, patterns or cues will provide a significant base for a tool for detecting suspicious individuals.

An experiment was therefore constructed to establish a baseline of the specified behaviours in honest and deceitful conditions and aimed to construct a database that can aid in answering the following questions:

- What are the typical patterns for this behaviour (in an appropriate norm group and transferable context for the application)?
- Are there any reliable differences between facial behaviour displayed (both in the visible and thermal domains) when people are known to be lying and when people are known to be telling the truth?
- Can a model be designed that is able to classify and detect these behaviours that could be used as an input to decision-making regarding who may have malicious or deceitful intent?

A rich FACS-coded (Facial Action Coding System) database with high quality thermal images was established from the baseline data to support future development of a tool for operational detection of cues to malicious intent. This will also aid the computer vision community as there is currently a data deficit in this area.

II. Background

Most people believe that they can tell when someone is lying to them. However, the evidence from psychology experiments shows that, on average, people only discriminate liars from truth tellers in about 40-60% of cases. This performance does not represent a very meaningful improvement over chance [1, 2].

Researchers [3, 4] do suggest liars behave differently from truth tellers—and so might be identifiable—because the process of lying initiates three psychological constructs: emotion [5, 6]; content complexity [6, 7]; and attempted control [7].

For example, people who are lying might be expected to experience ‘emotions’ including guilt, fear and duping delight [5]. They will also experience ‘content complexity’ due to having to ‘check their story’ to ensure its consistency and believability. This includes thinking of plausible answers to questions, avoiding contradictions, making sure lies are compatible with other available information and remembering what they have said so they can repeat it later and will increase the cognitive workload in comparison to someone telling the truth [6-8]. Liars will also be concerned about behaviours that could give them away, so need to control their actions—described as 'impression management' (Krauss, 1981, cited in Bull et al., 2002 [1]). Research shows that this often creates an over-compensation [1, 7, 9] which might be detectable, and also reinforces the increased cognitive load associated with lying. Indicators that an individual is experiencing any one of these psychological constructs might therefore indicate their attempt to deceive and so identify them for further questioning.

(a) (b) (c)

Figure 1: Experimental equipment setup: (a) facilitator briefs the participant, (b) interview session, (c) thermal camera model and visual camera model.

Moreover, it is likely that the dominance of each construct over the others will vary through the narrative of a security process. Appreciation of this variation will vastly enhance the effectiveness of any tool used to detect those with malicious intent.

Alongside these three constructs, there are other necessary considerations. Cues related to anxiety, for example, may be more difficult to detect in less trait-anxious individuals [10], or those who are experienced at deception. Furthermore, innocent individuals may display signs of anxiety since emotions are likely to always 'run high' in security settings, for a variety of reasons. The difference between 'state' and 'trait' anxiety therefore becomes pertinent. State anxiety is a temporary feeling of anxiety experienced as a result of an external influence whereas trait anxiety is the individual's general tendency to respond with anxiety to perceived threats: the 'individual differences' between people in terms of their experience of state and trait anxiety will impact on their behaviour in security settings. These points suggest that the cues that indicate a high cognitive load or attempts at control may be more promising as operational indicators of deception since they are less likely to appear in innocent subjects.

In terms of emotion expression within the face, some researchers believe there are different elements of specific expressions corresponding with specific emotions [11]. Others argue for a more general dimensionality [12]. Cultural display rules affect the relationship between feeling and display, people can exaggerate or hide expressions to conform to accepted patterns [5], and there are questions about whether emotions can be expected to have basic links to expressions, or whether the face is simply a tool for communicating intentions [13, 14]. There may be common clues to 'abnormal' behaviour, or to attempts to conceal feelings, as they will not always (depending on the skill of the individual) appear the same as natural, unchecked expressions.

Therefore a baseline was sought to understand facial behaviour in honest and deceptive scenarios, to enable development of a suitable decision-aid tool.

III. Methodology

The experiment was constructed as two interview scenarios. *Participants* were interviewed by an *Examiner* who was introduced by the *Facilitator* as having recently trained in lie detection techniques. *Participants* were told it was important that they appear truthful throughout to convince the *Examiner*. For one session, they were asked to answer questions as themselves. For the other, they were given a character profile to learn and were asked to answer the questions as if they were the character in the profile. Some questions went beyond the information in the profile, requiring participants to create plausible answers.

Each session consisted a period of introduction by a series of five introductory questions (for example, 'what is your name?') asked by the *Facilitator*. Then followed by an interview with the second experimenter: the *Examiner* who asked 10 questions on the relevant topic. Throughout the experiment, the data regarding visible images of facial behaviours were recorded and coded by certified FACS coders.

A. Experimental Design

A within subjects approach was deployed with two independent variables: interview topic (university study and career; dwelling, hobbies, personality and family) and truthfulness (truthful state, deceitful state). Condition orders were counterbalanced, as shown in *Table I*, and the *Examiner* was blind to the condition to prevent bias. *Participants* were invited for two interview sessions in the same day: one in the morning and one in the afternoon. This provided separation between the two topics: the truthful condition and deceitful condition.

The questions were designed to elicit answers of 2 to 10 seconds in the majority of questions. It was anticipated that this would be sufficient, combined with measurement of facial behaviour during the question period, to represent the range of facial behaviour satisfactorily. In the next section, we provide further description for the equipment setup.

B. Equipment Setup

The experiment was conducted in a dark studio room with controlled lighting condition. Figure 1(a) illustrates the facilitation session, while the *Facilitator* was giving the instructions to the *Participant*. Figure 1(b) illustrates the position of *Examiner* and the *Participant* during the interview session. The *Participant*'s facial activitities were recorded by using a high definition visual camera and a thermal camera, as illustrated in figure 1(c). The model of high definition camera used in this experiment is JVC-GY-HM100E, we set the resolution to 1280 by 720. The model of thermal camera used in this experiment is FLIR SC7600, 14bits, with resolution of 640 by 512.

C. Examiner

During the interview, the *Examiner* dressed formally to reinforce the impression of authority. The *Examiner* was blind in that he did not know about the design of the study or which condition a *Participant* would be in. He was not involved in the day to day running of project. To enable rewards to be given to participants as an incentive, the *Examiner* recorded his judgment as to whether each participant was telling the truth but was not told whether his judgment was correct.

Although not the focus of the experiment, it may be noteworthy that the *Examiner* who took part in the study is an expert in crime scene reconstruction and forensic science.

D. Facilitator

The experiment was fully facilitated using scripted *Participant* introduction and instructions. The *Facilitator* mentioned the *Examiner* and informed the *Participant* that the *Examiner* has been trained in techniques for detecting lies.

The *Facilitator* explained that the *Examiner* would interview the participant on two topics and informed the *Participant* that the trial is designed to investigate methods for detecting when someone is lying.

Finally, the *Facilitator* reminded the *Participant* of the importance of presenting themselves as truthful throughout the entire interview, and, it is appropriate to stay consistent and in character for the relevant topic. The *Participant* was informed by the *Facilitator* that there was a small reward available for those *Participants* who convince the *Examiner* that they are truthful throughout the interview.

E. Participants

32 volunteer undergraduate students and research assistants were took part in the study. 27 were male and 5 were female. They ranged from 18 years to 33 years.

F. Self-report

At the end of each session, the *Participant* was asked to confirm whether they had followed the instructions correctly and answered as themselves or as the character (as appropriate) for each question.

The *Facilitator* also thanked the *Participant* for their participation, informed the *Participant* of the *Examiner*'s judgment and provided a small reward if the *Participant* was successful in convincing the *Examiner* that they were truthful throughout the interview.

TABLE I. PARTICIPANT ORDERING AND TOPIC ORDERING

Subject	First session	Second session
1	Topic A – lie	Topic B - truth
2	Topic A – truth	Topic B – lie
3	Topic B – lie	Topic A – truth
4	Topic B – truth	Topic A – lie
5	Topic A – lie	Topic B - truth
6	Topic A – truth	Topic B – lie
7	Topic B – lie	Topic A – truth
8	Topic B – truth	Topic A – lie
9	Topic A – lie	Topic B - truth
10	Topic A – truth	Topic B – lie
11	Topic B – lie	Topic A – truth
12	Topic B – truth	Topic A – lie
13	Topic A – lie	Topic B - truth
14	Topic A – truth	Topic B – lie
15	Topic B – lie	Topic A – truth
16	Topic B – truth	Topic A – lie
17	Topic A – lie	Topic B - truth
18	Topic A – truth	Topic B – lie
19	Topic B – lie	Topic A – truth
20	Topic B – truth	Topic A – lie
21	Topic A – lie	Topic B - truth
22	Topic A – truth	Topic B – lie
23	Topic B – lie	Topic A – truth
24	Topic B – truth	Topic A – lie
25	Topic A – lie	Topic B - truth
26	Topic A – truth	Topic B – lie
27	Topic B – lie	Topic A – truth
28	Topic B – truth	Topic A – lie
29	Topic A – lie	Topic B - truth
30	Topic A – truth	Topic B – lie
31	Topic B – lie	Topic A – truth
32	Topic B – truth	Topic A – lie

G. Analysis

Facial behaviours were measured throughout the interview sessions, during both the introductory sessions, and the interview sessions with the *Examiner*. Facial indicators are likely to occur throughout listening and preparation of an answer; therefore *Participant* behaviours were analyzed for both question and response periods. The measures of facial behaviours were done manually by the FACS coders. To avoid bias scores, the FACS coders did not know the condition of the coding or the meaning of the cues.

IV. Result and Analysis

We discuss the results from two perspectives: first analysis from human judgment (*Examiner*'s judgment) based on verbal and non-verbal cues; and second is to explain the process in the database preparation as a contribution to the computer vision community for future research. From 32 subjects, we filtered out the subjects whom confused with the instructions and uncertain about their own intention in the interview sessions. After filtering, there were only 28 healthy subjects available for analysis.

A. Analysis on Examiner's score

The *Examiner*'s judgment provided a means to incentivize and reward *Participants*; it was not the focus of this research. Research showed that average person spots liars at approximately 54% accuracy, while the specialized groups (trained psychologist, police etc.) score approximately 60% accuracy in identifying deception [15].

The confusion matrix of the *Examiner*'s score in detecting deception is presented in Table 2, which shows that the *Examiner* achieved 57.13% accuracy in detecting truth tellers and 57.13% in detecting deceit. The sensitivity and specificity of 57.13% revealed the weakness of human in lie detection.

The next section presents discussion of the analysis of facial action units in evaluating truthfulness.

TABLE II. Confusion Matrix on Examiner's Score

		Predicted Class		
		Lie	**Truth**	**Total**
Actual Class	**Lie**	16	12	28
	Truth	12	16	28
	Total	28	28	56

B. FACS Coding Annotation

The Facial Action Units (AUs) were coded using FACS [16]. FACS provides comprehensive and objective way to analyze expressions into elementary components. It has been used widely in behavioural sciences. All the AUs were coded by certified FACS coders, i.e. by human experts in AUs reading. In our investigation, the duration of an AU is the total time taken from onset, apex, and offset. Please note that this analysis is not targeted on micro facial expressions, as the recommended setup for micro-expressions detection is a high-speed camera, which was not available in this particular study.

We used the Language Archiving Technology (ELAN) [17, 18] in FACS annotation. Figure 2 illustrates the annotation software, with a video of a subject on the top left corner, and the coded AUs displayed below the video.

After annotation, the data was exported to an excel spreadsheet as shown in Figure 3. This provides a rich FACS-coded database, which is available for researchers in further investigation and study on the deceptive facial behavioural analysis.

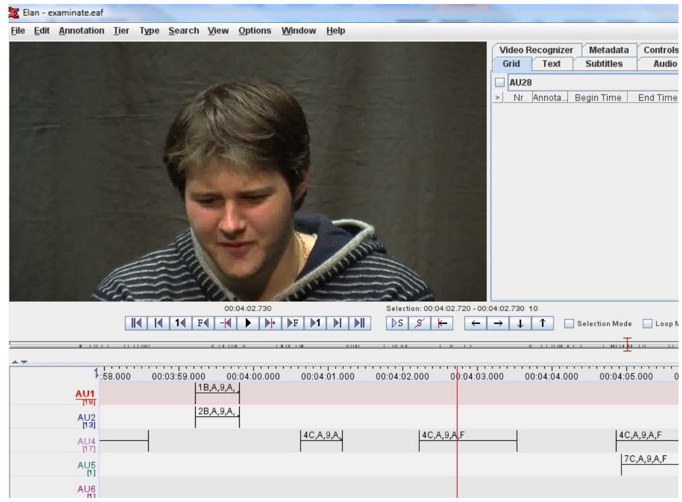


Figure 2. Illustration of the Language Archiving Technology, ELAN, used by our FACS coders in annotating the Facial Action Units.

TABLE III. The list of facial AUs occurred in our FACS-coded database

AU	Meaning	AU	Meaning
AU1	Inner Brow Raise	AU50	Speech
AU2	Outer Brow Raise	AU51	Head Turn Left
AU4	Brow Lowerer	AU52	Head Turn Right
AU5	Upper Lid Raiser	AU53	Head up
AU6	Cheek Raise	AU54	Head Down
AU7	Lids Tight	AU55	Head Tilt left
AU9	Nose wrinkle	AU56	Head Tilt Right
AU10	Upper lip raiser	AU57	Head Forward
AU11	Nasolabial Furrow Deepener	AU59	Head Nod
AU12	Lip Corner Puller	AU60	Head Shakes
AU13	Sharp Lip Puller	AU61	Eyes turn left
AU14	Dimpler	AU62	Eyes turn right
AU15	Corner Depressor	AU63	Eyes up
AU16	Lower Lip Depress	AU64	Eyes down
AU17	Chin Raiser	AU33	Blow
AU18	Lip Pucker	AU36	Tongue Bulge
AU19	Tongue Show	AU37	Lip wipe
AU20	Lip Stretch	AU38	Nostril Dilate
AU21	Neck Tightener	AU40	Sniff
AU23	Lip tightener	AU43	Eye Closure
AU24	Lip presser	AU45	Blink
AU25	Lips Part	AU68	Eye Rolling
AU26	Jaw Drop	AU72	Lower Face not visible
AU28	Lips Suck	AU80	Swallow
AU29	Jaw Thrust	AU82	Shoulder shrug
AU30	Jaw sideways	AU84	Head shake back and forth
AU31	Jaw Clencher	AU85	Head nod up and down
AU32	Bite	AU92	Partial Flash

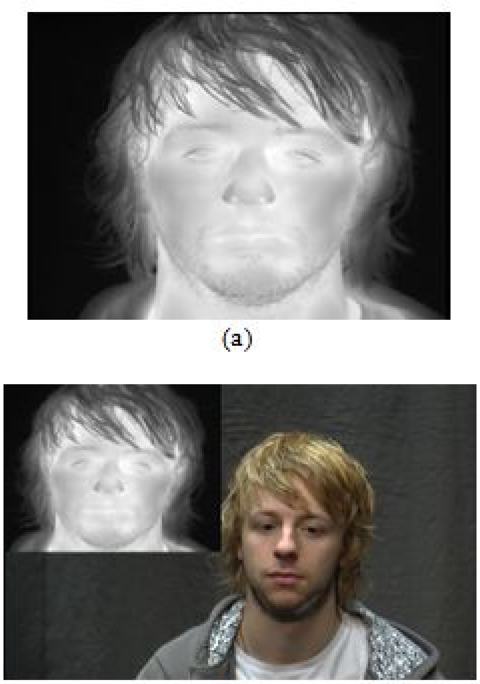
We found 56 facial AUs in our dataset. Table III lists the AUs with the respective meaning. The 56 AUs are the standard AUs in Ekman & Friesen's guidelines [18].

	A	B	C	D	E	F	G	H	I
1	Subject 02								
2									
3	Action Units	begin time	end time	duration	AU Intensity	Topic	Question	A/Q	GroundTruth
4	AU1	32.89	34.28	1.39	1C	A	1	A	1
5	AU1	185.33	186.72	1.39	1C	A	8	A	1
6	AU4	70.08	71.34	1.26	4C	A	3	A	1
7	AU4	179.39	180.59	1.2	4C	A	8	A	1
8	AU12	60.61	64.36	3.75	12C	A	2	A	1
9	AU12	77.85	80.2	2.35	12B	A	3	A	1
10	AU12	93.97	96.2	2.23	12C	A	4	A	1
11	AU12	107.88	110.35	2.47	12B	A	4	A	1
12	AU12	135.3	144.4	9.1	12B	A	6	A	1
13	AU12	160.32	164.62	4.3	12C	A	7	A	1
14	AU12	210.71	212.45	1.74	12B	A	9	A	1
15	AU12	219.96	224.79	4.83	12B	A	10	A	1
16	AU12	226.15	228.95	2.8	12A	A	10	A	1
17	AU14	125.9	127.54	1.64	R14D	A	5	A	1

Figure 3. The layout of the partially exported AUs annotation into an excel spreadsheet.

C. Thermal Images

Thermal images were also captured and are being analysed separately. Figure 4(a) illustrates the high quality thermal images in our experiment, while Figure 4(b) shows the corresponding images for thermal and visual. Future studies may benefit from multivariate analysis incorporating both visual and thermal cues.



(a)

(b)

Figure 4: Illustrations of thermal data: (a) thermal image, (b) thermal image with the correspondence visual image.

V. Discussion and Conclusion

The literature review identified those psychological and physiological behaviours that might plausibly be used in evaluating truthfulness and credibility assessment. In particular, it addressed the behaviours that are detectable in the visual domains of facial behaviours. In addition, our research established a rich FACS coded database alongside with high quality thermal images is important in future research development.

Problem with laboratory study in evaluating truthfulness is that it contextualized the human actions and choices [19]. It is necessary to analyze on real life data. But there is a need for cautious in putting the experimental studies into real-life application. The challenge is how to evaluating truthfulness within the context of complex social interactions and how to develop paradigms in which subjects have a real choice as to whether and when to lie. The real intention of a subject to deceive the examiner is crucial. The problem of giving instruction to lie eliminates the voluntary intention to deceive. There are not consequences for the subjects' action (negatively), no harm can come to anyone and we do not achieve a valid representation of the process of deceptive acts. In the future, we have to consider the pragmatics of human communication [20] in our experimental design.

In future work, we will investigate into five communication channels [21], which combine facial behavioural analysis, gait analysis (gesture behavioural analysis), speech analysis, voice analysis, and physiological methods (thermal analysis). Strong case supports from psychology research are important in spotting lies. Recently, researchers are also looking into self-deception [22]. Human is fallible in detecting deception therefore automated detection tools to augment human judgment can greatly increase detection accuracy. More research under a variety of contexts will determine which indicators and systems are the most reliable.

Acknowledgment

This work was supported by EPSRC grant on "Facial Analysis for Real-Time Profiling" (EP/G004137/1). The authors would like to thank Doherty Victoria, Andy, and Ashley from QinetiQ for providing help and advices in experimental design and statistical analysis. We also express our appreciation to Bashar Rajoub, Norhayati Ahmed, Hamad Alawar, Christopher Watkins, and Xia Han in conducting the experiment.

References

[1] R. Bull, A. Vrij, and I. McKenzie, "Passenger Profiling: Assessment of verbal and non-verbal cues. ," QinetiQ report. QinetiQ/KI/CHS/CR0221812002.

[2] P. Ekman, *Telling Lies*. New York. London.: W.W.Norton & Company, 2001.

[3] S. Helman, M. Christie, P. Collett, V. Pelly, and G. Huddy, "Behavioural analysis to predict terror attacks: scoping study," QinetiQ report: QINETIQ/07/027852007.
[4] P. Collett, *The Books of Tells: Reading People's Minds from their Actions*. London: Transworld, 2003.
[5] P. Ekman, "An argument for basic emotions," *Cognition and Emotion,* vol. 6, pp. 169-200, 1992.
[6] A. Vrij, "Detecting lies and deceit: The pscychology of lying and its implications for professional practice," *Chichester: John Wiley and Sons,* 2000.
[7] A. Vrij, "Why professionals fail to catch liars and how they can improve," *Legal and Criminological Psychology,* vol. 9, pp. 159-181, 2004.
[8] J. K. Burgoon, D. B. Buller, and L. K. Guerrero, "Interpersonal deception IX: Effects of social skill and nonverbal communication on deception success and detection accuracy," *Journal of Language and Social Psychology,* vol. 14, pp. 289-311, 1995.
[9] B. M. DePaulo, J. J. Lindsay, B. E. Malone, L. Muhlenbruck, K. Charlton, and H. Cooper, "Cues to deception," *Psychological Bulletin,* vol. 129, pp. 74-118, 2003.
[10] T. Edwards and L. Hardy, "The interactive effects of intensity and direction of cognitive and somantic anxiety and self-confidence upon performance," *Journal of Sport and Exercise Psychology,* vol. 19, pp. 296-312, 1996.
[11] P. Ekman, *Darwin and facial expression - a century of research in review*: Academic Press New York and London, 1973.
[12] J. A. Russell. (1997) The Psychology of Facial Expression. *Cambridge University Press*.
[13] J. M. Fernandez-Dols and M. A. Ruiz-Belda, "Are Smiles a sign of happiness? Gold medal winners at the Olympic Games," *Journal of Personality and Social Psychology,* vol. 69, pp. 1113-1119, 1995.
[14] R. E. Kraut and R. E. Johnston, "Social and Emotional Messages of Smiling: An Ethological Approach," *Journal of Personality and Social Psychology,* vol. 37, pp. 1539-1553, 1979.
[15] M. G. Frank, M. A. Menasco, and M. O'Sullivan, *Handbook of Science and Technology for Homeland Security* vol. 5: John Wiley & Sons, Inc, 2008.
[16] P. Ekman and W. V. Friesen. (1978) Facial Action Coding System: A Technique for the Measurement of Facial Movement. *Consulting Psychologies Press*
[17] H. Lausberg and H. Sloetjes, "Coding gestural behaviour with NEUROGES-ELAN system," *Behavior Research Methods, Instruments & Computers,* vol. 41, pp. 841-849, 2009.
[18] H. Sloetjes and P. Wittenburg, "Annotation by category - ELAN and ISO DCR," in *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, 2008.
[19] K. E. Sip, A. Roepstorff, W. McGregor, and C. D. Frith, "Detecting deception: the scope and limits," *Trends in Cognitive Sciences, Elsevier,* vol. 12, pp. 48-53, 2007.
[20] J. Verschueren, "Understanding Pragmatics," *Arnold,* 1999.
[21] J. Pearse and C. Lansley. (2010, 27 May 2012). Reading Others. *Feature,* 62-65. Available: www.trainingjournal.com
[22] W. V. Hippel and R. Trivers, "The evolution and psychology of self-deception," *Behavioural and Brain Sciences,* vol. 34, pp. 1-56, 2011.