

Video Coding Manual for Sandtray Experiments

James Kennedy, Paul Baxter
WP1, ALIZ-E
Plymouth University, U.K.

October 2013

Abstract

This document describes the coding process for the video data obtained from child-robot interaction with the Sandtray setup. The initial purpose of this coding scheme is to characterise the robot- versus touchscreen-directed overt behaviours of the child to assess the social interaction.

Contents

1	Purpose of this Behaviour Coding	2
2	Tools and Files	2
2.1	ANVIL	2
2.2	Video Files	3
3	Behaviour Coding Categories	3
3.1	Group: Interaction	3
3.2	Group: Child Behaviour	4
3.3	Group: Robot Behaviour	5
3.4	Group: Other	5
4	Coding Procedure	6
4.1	ANVIL Issues	7
5	Validation	7
A	Robot Scripted Phrases	8

1 Purpose of this Behaviour Coding

The purpose of video coding is to characterise the events that occur in a video stream, according to some defined set of measures/characteristics. This document outlines one such coding scheme, and provides guidelines for its application to a video dataset.

This particular coding scheme is concerned with characterising the video data obtained from the Sandtray study run at Salisbury Road Primary School, July 2013. The coding scheme is roughly based on the information provided in [1, 2], and the coding categories inspired by a range of HRI studies, including [3].

The initial purpose of this preliminary coding scheme is to characterise the robot- versus touchscreen-directed overt behaviours of the child to assess the social behaviours taking place.

2 Tools and Files

2.1 ANVIL

ANVIL is a freeware/opensource video annotation tool: it was chosen for its cross-platform support, its various import/export options, its prior application and validation on a number of datasets, and the prior experience that certain people within the ALIZ-E project have with it. Information and installation instructions can be found here: <http://www.anvil-software.org/>, and a brief description here: <http://scholarspace.manoa.hawaii.edu/bitstream/handle/10125/4492/tanmartin.pdf?sequence=1>.

Regarding the use of ANVIL for video annotation, it is useful to distinguish between three levels of organisation, which are used in the coding description below:

- *Element*: is an attribute/value/characterisation of a particular type, has ‘start’, ‘end’ and ‘value’ attributes.
- *Track*: is a group of elements for one type of information. Within a track, only one element can be attributed at any one time, elements within a track are non-overlapping.
- *Group*: is a collection of tracks related to a type of information.

The definition of the coding scheme is contained within a “Specification file” (an XML document, **.xml*). This is provided for the coding scheme described below. For compatibility, and to enable use of the various statistical methods built into ANVIL, *the provided specification file must remain unchanged* over the course of coding and analysis for this study.

Each coder should create a separate “Annotation file” (**.anvil*) for each of the videos to be coded: use the same naming convention as for the videos (described in next section), but append the initials of the coder (e.g. *d1i01-JK.anvil*).

2.2 Video Files

For compatibility with ANVIL, the videos have the following characteristics: *.mov container, using the *Apple Cinepak codec*, with *uncompressed audio*, and a *70% quality* of the original video file (to reduce the file size - ANVIL recommends to only deal with video sizes of less than 1Gb).

Split video file names have the following format:

dDiXXcine.MOV

Where:

- “D” is the day; either 1 or 2.
- “XX” is the interaction number on that day.

There are a total of 28 video files to code, but for second coding, only a proportion of these will be done. The files to code will be:

- dli08mov.cine
- dli11mov.cine
- dli13mov.cine
- d2i06mov.cine
- d2i13mov.cine

3 Behaviour Coding Categories

The coding relates only to overt behaviours of the robot and child, and not to perceived/inferred internal states or intentions. This is to keep the study an objective one, and which should also lead to a higher inter-rater reliability.

3.1 Group: Interaction

- *Track: Interaction stage.* Composed of the following elements: *Please try to assign one of these elements to the entire timeline.*
 - Introduction - from the moment the video starts until immediately before the robot starts to wave.
 - Robot Instructions - from the start of the robot wave, until it finishes saying ‘now you can start’.
 - Pre-Test - from the end of the robot instructions speech until the robot moves to change the library when there are no images on the touchscreen.

- Main Activity - from the proceeding session until the robot introduces the post-test ‘right, we’ll do just one more set of aliens...’
- Post-Test - from the start of the ‘right, we’ll do just one more set of aliens...’ speech until the child has categorised all touchscreen images.
- Robot Goodbye - from the proceeding category until the robot has finished waving.
- Debriefing - from the proceeding category until the child leaves the room (normally when the video ends).

3.2 Group: Child Behaviour

- *Track: Child gaze.* Based on eye direction, where the child’s attention is directed, composed of the following elements. *Please try to assign one of these elements to the entire timeline.*
 - Towards robot
 - Towards touchscreen
 - Towards another person
 - Other: including occluded gaze
- *Track: Child gestures.* *No need to provide complete coverage of the timeline.*
 - Preparation/Moving: moving towards/away from screen with hand or moving between images on screen
 - Touching screen: whether moving or not
 - Pointing/Gesturing towards robot
 - Pointing/Gesturing towards another person
 - Other: primarily for annotating interesting looking behaviours that are not part of the defined coding scheme - fist pumps, pulling at clothes, repetitive scratching e.t.c.
- *Track: Child vocalisations.* Where distinguishable from the video. Just code for the length of the speech; there is no need to annotate the words which are spoken.
 - Towards robot
 - Towards another person
 - Towards self
 - Other
- *Track: Other.* *No need to provide complete coverage of the timeline.*
 - Undefined: no aspect of child behaviour is distinguishable, cannot be characterised using above track definitions, or is another behaviour of interest not covered above.

3.3 Group: Robot Behaviour

- *Track: Robot gaze.* Based on head direction (height of gaze must be inferred due to camera angle), with the following elements (bear in mind that in the current study, the robot gaze direction is randomly defined, except where a touchscreen directed movement is taking place). *Please try to assign one of these elements to the entire timeline.*
 - Towards Child
 - Towards Touchscreen
 - Other: including occluded gaze and gaze which is at neither the child or touchscreen, e.g. around the room. Please include a short comment if other.
- *Track: Robot gestures.* Bear in mind that the default behaviour of the robot is to shift its weight randomly, this behaviour need not be coded. *No need to provide complete coverage of the timeline.*
 - Highlight screen item - from the moment the robot starts to raise its arm to move a screen item, until the moment it has finished lowering.
 - Change screen library - again from the start until the end of movement, performed when screen items have run out.
 - Speech gesture - from the start of the movement, until the end of the gesture. Do not include the slow return to the ‘default’ position.
 - Other: including motions errors, etc.
- *Track: Robot Vocalisations.* These should cover the length of each phrase while the robot is talking. There is no need to annotate the words the robot is saying, just the length.
 - Unscripted/Feedback to child - these are used after the child does a move, or the robot is highlighting a move for the child
 - Scripted - at certain points in the interaction, the robot will follow a script. The scripted items are listed in the appendix of this document and each phrase should be coded as one element.
 - Other: not present in this case
- *Track: Other.* This track is primarily for annotations of interesting behaviour that is not covered in the defined coding scheme.

3.4 Group: Other

- *Track: Undefined.* This is a string field that may be used to either annotate situations that the camera is occluded, other people walk into the room, or other features not captured by the defined coding scheme. *No need to provide complete coverage of the timeline.*

4 Coding Procedure

To begin a new annotation, locate and open the desired video, select the “sals-sandtray-spec.xml” specification file, and “UTF-16” encoding (no particular reason for this encoding choice - but stick to it for consistency).

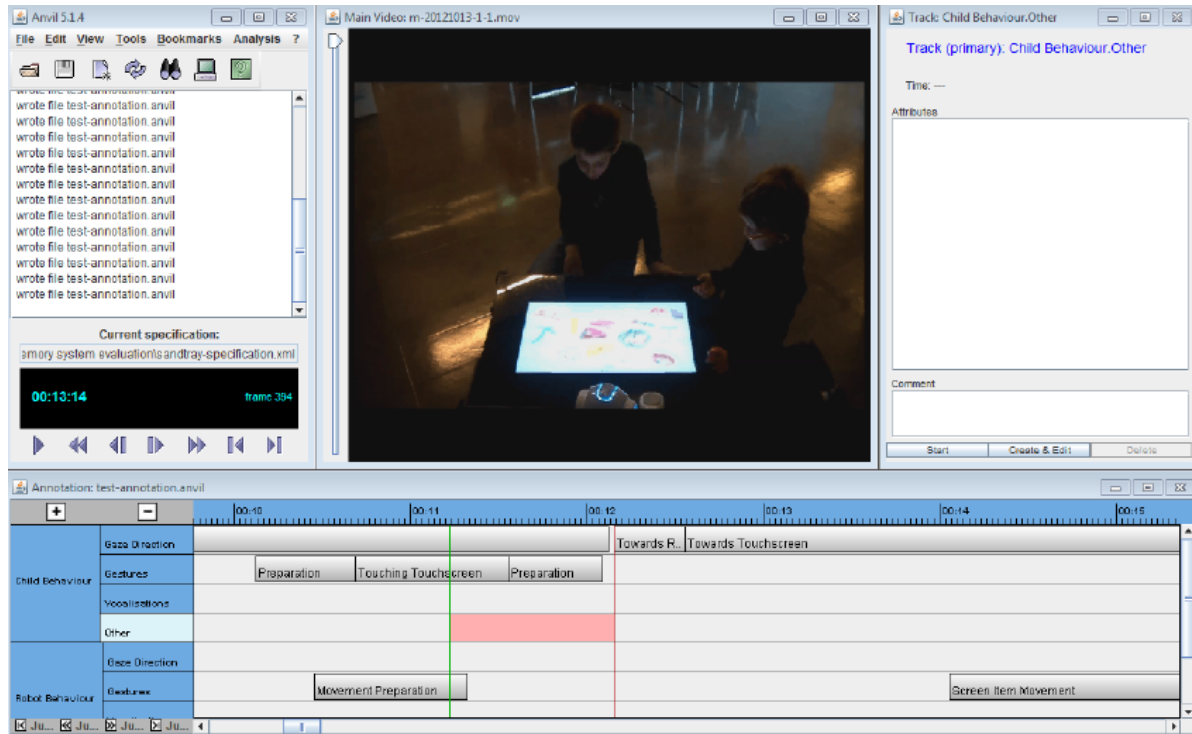


Figure 1: Screenshot of the ANVIL interface (In Windows), with a sample video and associated annotations.

In Figure 1, four panels of ANVIL are shown: (top left) main control panel with video playback control, open/save annotation files here; (top centre) the video playback panel (the SPACE key can be used to start/pause video play); (top-right) the track property panel; (bottom) the annotation timeline panel.

The various tracks and associated elements can be seen on the left hand side of this lower panel. Two coloured vertical lines can be seen here: the green one can be moved by double-clicking on the timeline (or press F1 on Windows; 1 on Mac) and indicates the start of an annotation object (element), the red one indicates the current play location in the video, and the end of an annotation object if desired. To add an annotation, right-click on the highlighted area (shown in pink in the lower panel of figure 1) and select ‘*Create and Edit*’ (or press F3 on Windows; 3 on Mac). This brings up the set of allowed annotation options for that element. Choose the appropriate one, and add a clarification note if necessary. Click OK to, or Ctrl + Enter to save the element.

Recommended coding procedure: Making multiple passes through the video and coding only

one element per pass works best (rather than stepping through the video). When coding tracks which can change at a very quick rate, such as child gaze, it is recommended to use the arrow keys (left and right) to move forwards and backwards through the video on a frame-by-frame basis.

4.1 ANVIL Issues

Occasionally, when deleting or editing an item, ANVIL will appear to freeze. What has happened is the confirmation box has opened off-screen! To fix this (at least in Windows):

1. Open the Task Manager (ctrl + shift + esc)
2. Select the 'Warning' entry
3. Click 'End Task'

This should allow ANVIL to be used again as before. Your edit/delete will not have been completed, so you will have to try it again.

5 Validation

Validation of the video coding: two coders per video, use of built-in ANVIL tool to calculate inter-rater reliability (Cohen's corrected kappa, κ). Based on the characterisation in (Landis & Koch, 1977), we may consider that $0.41 < \kappa < 0.6$ to be 'moderate' agreement, $0.61 < \kappa < 0.8$ to be 'substantial' agreement, and $0.81 < \kappa < 1.0$ to be 'almost perfect' agreement.

References

- [1] Peter H. Kahn, Batya Friedman, Nathan G. Freier, and Rachel Severson. *Coding Manual for Childrens Interactions with AIBO, the Robotic Dog: the Preschool Study*. 2003.
- [2] Jean-marc Colletta, Olga Capirci, Carla Cristilli, Susan Goldin-meadow, Michle Guidetti, and Susan Levine. *Coding Manual*. 2009.
- [3] Marcel Heerink, Marta Daz, Jordi Albo-Canals, Cecilio Angulo, Alex Barco, Judit Casacuberta, and Carles Garriga. A field study with primary school children on perception of social presence and interactive behavior with a pet robot. In *RoMan 2012*. IEEE Press, 2012.

A Robot Scripted Phrases

Below is a list of the robot scripted phrases and where they occur in the interaction. These should be coded as 'scripted' under robot vocalisations. Each phrase should be coded as one element.

- Robot Instructions
 - Hello! I'm Pop/Crackle.
 - Right, what we are going to be doing today is sorting out some aliens.
 - We have two species of aliens that are lost in space and we have to return them to their home planet. Okay?.
 - So here we have our different types of aliens and our two planets, the purple and the orange.
 - We need to sort them into their two different groups.
 - I'd like you to see if you can guess which planets the aliens are from.
 - You can touch an alien and you drag it to the planet you think it's from, and it'll tell you whether you are right or not.
 - I won't help you on your first go. Let's see how well you can do on your own!
 - Now you can start.
- Main Activity
 - Lovely, well done.
 - Now I'll give you a clue, the aliens from the purple planet all have something in common.
- Post-Test
 - Right, we'll do just one more set of aliens.
 - Using the practice we've just done, let's see how well you can do.
 - I won't help you this time.
 - Have a go.
- Robot Goodbye
 - Well done. thank you very much.
 - Thank you for helping me out today.
 - You can go back to your class.
 - Goodbye!