Situated Breakdown Analysis for the Evaluation of a Virtual Environment

Spagnolli Anna♣, Gamberini Luciano, Gasparini Daniele

Department of General Psychology

University of Padova, Italy

ABSTRACT

It is a widely-supported tenet in human-computer interaction that the meaningful unit of analysis is not the technical device alone, but the technical device together with the person interacting with it; the reason is that what is a relevant property of a technology is only understandable with respect to the specific goals and resources activated during its usage. This basic reflection should also inspire the procedure followed to evaluate the usability of a technology, namely its efficiency and satisfaction for a specific class of users. The topic of this paper is precisely to describe a method developed in compliance with this observation and aimed at evaluating the usability of virtual environments. Two main requirements were set forth: first, the method should take the strong connection between humans and technology as its building block, by linking a property of the virtual environment to a particular use that makes that property relevant. To this goal, action has been placed at the center of the analysis; the functional properties of the VE are then observed in the general economy of users' interaction with the technology and the whole ensemble is the appropriate object of evaluation. Such 'action-based' approach (Gamberini, Spagnolli, 2002) is reminiscent of the Situated Action theory (Suchman, 1987) and Activity Theory (Nardi, 1996); the former proposes a detailed analysis of the sequential interaction with the technology and provides a rich examination of the structure given to it by the users. The latter focuses more on specific phenomena, such as contradictions and breakdowns, identified by the evaluators; it allows to profit from data poor in comments and verbalizations, and to analyze the interaction with the technology from a structural and organizational level. As a second requisite for the method, we wanted it to benefit from the advantages of both approaches; thus we decided to concentrate on the breakdowns occurring during users' interaction with the VE but to study these episodes from a situated point of view. In our definition, breakdowns reveal an inappropriate interpretation of the possibilities for action offered by the virtual environment and are to be analyzed in their sequential, contextual unfolding. This version of breakdown analysis highlights the spontaneous, subjective problems in the use of a technology and connects them to specific aspects of users' action. It renews the ergonomic tradition of error studies (Reason, 1990; Rasmussen, 1980) with an ethnographic contamination, that pays attention to users' contextualized practices. It also suits the kind of data the interaction with a virtual environment is mostly made of, namely bodily action in a three-dimensional space. Few methods with these characteristics have been employed so far to analyze the interaction with the VE. After a brief introduction, the paper will describe the basics of this approach and illustrate them with instances from the evaluation of a virtual library.

♣ Corresponding Author:
Anna Spagnolli, Dipartimento di Psicologia Generale,
Università degli Studi di Padova, via Venezia 8, 35131 Padova, Italy.
Tel: +39-049-8276644
Fax: +39-049-8276600
E-mail: anna.spagnolli@ unipd.it

1. Evaluating a virtual environment in use: situated, action -based approach.

The structure and features of a virtual environment (VE), like the structure and features of any technical artifact, are much more plastic than we may think. Let alone inexperience or exceptional misunderstandings, there is an inescapable process that shapes an artifact according to the practices of use, so that the artifact in the context of use can differ substantially from how it appears in its engineering description. This description offers but one perspective on the artifact, from the viewpoint of the designers and for the benefit of their practical concerns. In fact, each class of people that will get in touch with the technology comes up with own pictures of the technical artifact, based on the actions performed with it (Law. 1992; Carroll et al. 1994, Kling, 1980; 1992; Button, 1993; Mantovani, 1996; Mantovani, Spagnolli, 2001; Lea, 1992; Zucchermaglio et al. 1995; Greenbaum, Kyng, 1991; Ciborra, Lanzara, 1990): the properties of a car will differ substantially depending on whether one wants to mend it, advertise it, buy it, park it. Those interpretations are unpredictable, since nobody can figure in advance the vast variety of settings in which a technological product will be eventually placed and what they will look like. For this reason, before releasing a technical artifact in the market and sometimes even periodically throughout its life, it is highly recommended to test the users' interpretations (Gamberini, Valentini, 2001).

In the case of virtual environments, this recommendation is still overlooked. Human factors are usually considered very early in the design process, except for the measurement of the sense of presence conveyed by the simulation, which is assessed at the end, but usually covers only perceptual and sensory-motor processes (Wann, Mon-Williams, 1996; Stanney et al., 1998; Steuer, 1992). What is largely missing is a systematic study of the process of interaction with the VE, to have a comprehensive appreciation of how users interpret the functioning of the system. We can look for inspiration in the parent field of Human Computer Interaction, where we can find two particularly interesting frameworks conceptualizing the interaction with a technology, namely the 'situated' and the action-oriented frameworks. Taken together, those perspectives see users' interpretation as an embodied, practical phenomenon, instead of a mental, abstract one (activity theory: Engeström et al., 1999; Nardi, 1996; phenomenology: Ihde, 2002), which takes shape in the contingent, sequential unfolding of the interaction (ethnography: Button, 1993; Suchman, 1987; Hutchins, 1995 and discourse/interaction analysis: Luff et al., 1990; Jordan, Henderson, 1995; Engeström, Middleton, 1996). They look especially useful in case of virtual environments, where interaction is basically action in a three-dimensional space performed with material and

virtual resources. We then adopted this perspective and evaluate how the properties of the VE figure in users' situated action with the VE.

2. What to look for: breakdowns

A good rationale to evaluate the usability of a technology is to start from those events that reveal problematic, in other words, where user's interpretation of the artifact results inadequate.

To evaluators, problems and errors have always proved an insightful locus of analysis (Reason, 1990; Engeström, 1996; Carroll, 1993; Flanagan, 1954). The selection of problematic episodes can be accomplished in two ways, 'normative' or 'open'. In the former, the evaluator refers to a pre-established list of expected results; interactions are then inspected to single out the circumstances under which the actual interaction and the expectations are mismatched. The 'open' approach, instead, is more explorative. The procedure consists again of collecting and analyzing fragments of interaction in which some problems occur; only, this time the selection criterion is not the designer's, but the users', for the identification of a problem depends on some signals coming from the interaction itself. This latter approach applies when one prefers to pay more attention to the structure of the interaction in order to decide whether a passage is problematic or not, or when no specification is provided of the expected results, either because evaluators are interested in unexpected events, or because designers are not available to provide the list of expectations altogether. This approach is even more valuable if it can work in absence of verbal cues, for this would help in all cases in which users react to problems by quitting talk and concentrating on the difficulty, instead of asking questions and making comments.

The criterion we applied to collect problematic episodes without relying on verbal cues *only* and on designers' expectations was to look for spontaneous breakdowns. They are crisis in the interpretation of the situation, that force actors to suspend the current activity and mend the interpretative flaw (Winograd, Flores, 1986). From a situated, action-based perspective, in addition, breakdowns are not mental events, located in the cognitive processes of the user, but episodes involving the action of the user in the environment. T*he actor is forced to abandon the environment-action-person configuration adopted up to that moment and mobilize resources to obtain a more effective one* [1].

Procedurally, that means that:

- the observational focus is on the user's projected course of action (a certain actor-environment-action configuration), and its expected evolution;
- when a suspension or interruption of the course of action occurs, this is taken as an index of a breakdown episode, along with other concurrent evidences such as unexpected outcomes, verbal cues, gestures, pauses.

For example:

1. PROJECTED COURSE OF ACTION. The user is approaching a door in the virtual environment (Figure 1); the fact that he is moving towards the door, that he has been suggested to explore the virtual library and its features and he says 'let's go out, let's see if we can exit' suggests that the projected action is an attempt at opening the door.

Figure 1

2. BREAKDOWN. The course of action does not go through; the usual strategy to interact with an object (namely clicking on a dedicated button of the joystick) does not produce any results. The analyst registers the frame at which this interruption occurs and includes the following attempts at opening the door as part of the breakdown episode, which stops when the course of action leaves room to a new one. In this specific example, the episode stops after a series of attempts, when the user states that the door would not open and goes on with the navigation.

2 Tricky cases.

2.1 Complex breakdowns.

As we explained in the previous paragraph, each breakdown refers to one course of action, namely to a certain relationship person-environment-action, which tends recognizably to some consequence [1]. There are particularly tricky cases in which multiple connected problems occur. Here, a precise reference to the course of action is useful to establish when a breakdown episode is over and decide if a problem belongs to the same episode or ushers a new one. For example, when a new course of action intervenes before the previous one is through (either resolved or abandoned), like when the evaluator tries to help and a misunderstanding occurs, inserting a new breakdown into the previous one. For example, see the following episode (see the appendix for transcription symbols):

```
1    P:  ((he stops in front of the windows of an
         office, clicks a button of the joystick
         several times; nothing happens; so he
         goes on to his left))

2    R:  that one (.) is the window,

3        the other one is the door.

4    P:  pardon:?
```

Here the breakdown episode is re-opened by the researcher who refers to the just abandoned course of action (the attempt at opening a door) and suggests a solution: it is possible to enter the room, just the participant was mistaking the window-wall for the door. While addressing the breakdown with this suggestion, another breakdown occurs, this time communicative, since the participant cannot hear what the researcher is saying and initiates a 'repair' (in conversation analysis terms) by asking 'pardon?'

In this case we have multiple breakdowns because we have different connected courses of action. Otherwise, we are assisting to a series of problems within the same breakdown episode, like in the remaining of the sequence reported above:

```
5        ((stopping and touching
         the headphones;
         a door is in his view))
```

```
6   R:   the other one is the door if you want to enter there.

7        (go) more towards your right,

8   P: ⌈((he goes towards another door))

9   R: ⌊ (.4) no the o-

10  P: ((he stops; he's in front of the second door))

11  R: not that one.

12  P: ((he turns to his right; the first door is in his view))

13       [thi:s one.
```

Here there are multiple attempts at conveying a helping instruction. Each attempt and the corresponding failure is not a breakdown on its own, but part of a series of attempts in the same episode, since they all try to deliver the same course of action, namely entering the room.

2.2 Verbalizations.

Videotapes do not speak for themselves, but are interpreted by the evaluator (Suchman, 1995; Shotter, 1983; Biggs, 1983), whose work will be facilitated by familiarity with the context (by interacting several times with the VE, being present during the videorecording and being cognizant of the goals of the virtual environment and the interaction) and by eliciting some verbalization from the participant. How should those verbalizations be considered? Discourse analysis reminds us that they shouldn't be taken literally, as neutral descriptions: words do not label actions, they are actions themselves, either concurrent or divergent with their non verbal actions. When the participant talks to the evaluator about the ongoing breakdown, then, she is not describing it but articulating it, making sense of what is happening at the interlocutors' benefit (Smagorinski, 1998). In the following fragment, the participant turns to the right, where a wooden board appears in front of her and she retracts in a sudden, effective movement that reverses her previous turn. Her exclamation is not simply a spontaneous expression of surprise; it is prolonged from an 'o' into an 'o:god' which extends until her retraction is over and conveys a strained attitude.

3 Procedure

The breakdown analysis consists in two basic steps: identifying and collecting breakdown episodes and than analyzing their structure and development. The first steps has been dealt with in the previous paragraph. Once the episodes have been collected, the evaluator wants to *analyze the structure of the course of action (the actor-environmental affordances-action configuration) and its development, to gain some indication on the users' interpretation, the circumstances under which it turned out as inappropriate and the resources deployed during the breakdown episode.*

For example, in one evaluation we carried out, we built a series of grids to guide the analysis of each breakdown episode. We built four grids each of them was analyzing the same episode from a different analytic focus (possible actions afforded by the environment, strategies to exit the breakdown, handling the interactive device).

Frame	Description of the breakdown episode	Circumstances of breakdown	**Possible action**	Comments

Once all episodes have been analyzed, they have all been compared for similarities in order to draw up some general categories; the list of outlined categories has then been tested on another set of episodes and refined.

If the analysts want to reach a finer degree of analysis, for example because some episodes are intricate or some specific phenomenon are to be unearthed, they can carry out an interaction analysis (Jordan, Henderson, 1995; Ochs, 1996). As in the previous method, the detailed sequence of verbal and nonverbal actions is analyzed by looking at the resources that make this action recognizable as such and by tying them to the context in which they are performed. The difference is that the analysis proceeds utterance by utterance, move by move, trying to see how discursive practices already identified in the literature are used. Since this method time-consuming, evaluators may want to combine it strategically with faster solutions. For example:

- a deep exploration of a selected collection of cases and a faster examination of the remaining ones to check the interpretation and integrate the recurring results with new ones;
- a brief observation of all cases and then a deeper analysis of significant episodes;

- the adoption of the first stage of discourse analysis (transcription), as a means to empower the observational capacity of the evaluator: transcribing allows the evaluator to sharpen her view, so to speak, and have a remarkably greater closeness to the structure of the data.

This last method has been employed in another evaluation carried out by part of the authors and resulting in a narrative description of the most recurrent breakdowns with an emphasis on the relevant environmental elements involved and temporal details of interest; each description was accompanied by a correspondent suggestion to the designer.

  The list of aspects the evaluator may want to pay attention to is endless. For example, the breakdown episode may be seen as a case of practical problem-solving, namely a spontaneous problem faced by the person engaged in a particular course of action, which causes that person to employ the available resources to solve it. It is a practical process because it does not start by elaborating mental solutions to be subsequently implemented into action, but by performing concrete actions in accordance to the affordances of the situation, in order to turn it into a more desirable one (Lave, 1988; Rogoff, Lave, 1984; Suchman, 1987). Those resources are various, ranging from a logic examination of the situation to ready, immediate moves. Distinguishing among this different kind of resources may be a good source of information. The availability of ready resources for example may be associated to the users' expertise or their growing familiarity with the VE. The kind of resources deployed to solve the breakdown can also sketch a picture of how generalization works, by indicating which circumstances are seen as similar and reacted to with similar strategies. The extent and criteria of generalization, in fact, should not be presupposed a priori, since more often than not what looks like a familiar situation to the evaluator strikes the user with puzzlement. This is illustrated in figure 2 below, which refers to two actions, namely turning to the left and moving laterally to the left (Figure 2). Some participants were not able to adopt for the latter the operation already employed for the former, treating the two actions as different and then associating them with different resources and possibilities.

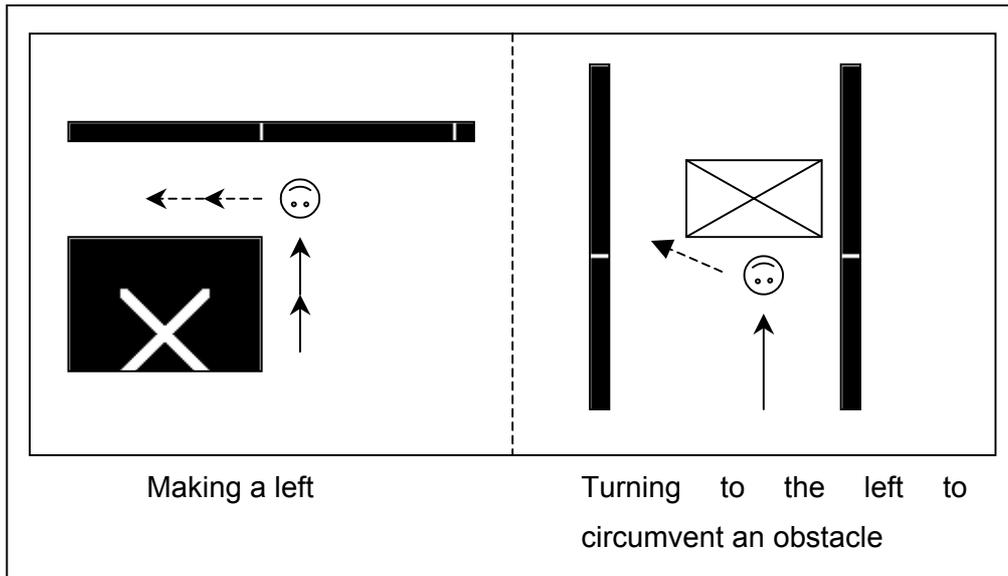| Making a left | Turning to the left to circumvent an obstacle |

Figure 2

Finally, some strategies that can improve the quality of the breakdown analysis and are recommended for any qualitative method in general include the following:

- o   to *anchor* the interpretation to a set of synergic evidences, such as the local resources the actor is orienting to or the sequence of moves she performs.
- o   to consider *alternative* interpretations;
- o   to grow *familiar* with the context in which the interaction takes place
- o   to *confront* with other evaluators;
- o   to *broaden* the corpus of data with occurrences that the previous collection of episodes lack
- o   to adopt an *integrated* method of analysis that includes multiple techniques to address different aspects of the phenomenon
- o   to keep track of the choices made during the analysis and discuss them constantly (*reflexivity*)

In addition to the qualitative analysis we have dwelled on so far, the various aspects of the breakdown episodes may undergo quantitative analysis of various kinds, according to the questions that are relevant for the evaluator or the designer and can help reducing the costs in temporal terms associated with the deployment of a qualitative analysis.

Conclusions.

In this paper, we described the basic assumptions of a situated breakdown analysis and the kind of aspects to extract from the videorecorded data. The main advantages of this analysis are its closeness to data, the attention to contextual elements, the ability to handle both verbal and bodily actions. Breakdowns can be studied in order to redesign the system's affordances for a certain class of users and hence prevent misunderstanding on the functioning of the system; on the other hand, breakdowns represent a chance for the users to expand their knowledge of the technology (Winograd, Flores, 1986; Koschmann, 1990), so they can be administered deliberately in a customized training path.

Notes.

[1] The exhaustion of a course of action is not predictable a priori, since it can be extended no matter how completed the action seems at the moment and can be considered finished only when a new one starts. This criterion is borrowed from conversation analysis and its description of a sequence of talk-in-interaction.

References.

Biggs S. J. (1983). Choosing to change in video feedback: On common-sense and the empiricist error. In P. W. Dowrick, S. J. Biggs (eds), *Using video. Psychological and social applications*. Chichester: John Wiley and Sons, 211-226.

Button G. (ed) (1993). *Technology in working order*. London: Routledge.

Carroll J. M., Mack R. L., Robertson S. P., Rosson M.B. (1994). Binding objects to scenarios of use. *International Journal of Human-Computer Studies*. 41: 243-276.

Carroll J.M., Neale D.C., Isenhour P.L. (1993). Critical incidents and critical themes in empirical usability evaluation. In *Proceedings of the BCSHCI93 People and computers VIII*, 279-292, Cambridge: Cambridge University Press.

Ciborra C., Lanzara F. (1990). Designing dynamic artifacts: computer systems as formative contexts. In P. Gagliardi (ed) Symbols and artifacts: Views of the corporate landscape.

Engeström, Y. and Middleton, D., Eds, (1996). *Cognition and communication at work*.. Cambridge: Cambridge University Press.

Engeström Y., Escalante V. (1996). Mundane tool or object of affection? The rise and fall of the postal buddy. In Nardi B. (1996). *Context and consciousness. Activity theory and human-computer interaction*. Cambridge, MA: The MIT Press.

Engeström Y., Miettinen R., Punamäki R. (1999) (eds). *Perspectives on activity theory*. Cambridge, MA: Cambridge University Press.

Flanagan J.C. (1954) The critical incident technique. *Psychological Bulletin*, 51 (4), 327-358.

Gamberini L., Valentini E. (2001) Web usability Today: Theories, Approach and Methods. In G. Riva, C. Galimberti *Towards Cyberpsychology: Mind, Cognition and Society in the Internet Age*. Amsterdam: IOS Press.

Gamberini L., Spagnolli A. (2002) On the relationship between presence and usability in virtual environments: A situated, action based approach. In G. Riva, F. Davide (a cura di) *'Being there'* Amsterdam: IOS Press.

Greenbaum, J., Kyng, M., (Eds) (1991). *Design at work:Ccooperative design of computer systems.* Hillsdale, NJ: Lawrence Erlbaum.

Jordan, B, Henderson, A. (1995). Interaction Analysis: Foundations and practice. *The Journal of the Learning Sciences* 4(1), 39-103.

Hutchins E. (1995). Cognition in the wild. Cambridge, MA: The MIT Press.

Ihde D. (2002) *Bodies in technology*. Minneapolis, MN: University of Minnesota Press.

Jordan, B, Henderson, A. (1995). Interaction Analysis: Foundations and practice. *The Journal of the Learning Sciences* 4(1), 39-103.

Kling R. (1980). Social analysis of computing: Theoretical perspectives in recent empirical research. *Computing surveys*, 12: 61-110.

Kling R. (1992) Behind the terminal: The critical role of computing infrastructure in effective information systems' development and use. In W. Cotterman, J. J. Senn (eds) *Challenges and strategies for research in system development*. London: Wiley.

Koschmann T. '*Dewey's contribution to a standard of problem-based learning practice*' available at http://www.mmi.unimaas.nl/euro-cscl/Papers/90.pdf)

Lave J. (1988). *Cognition in practice. Mind, mathematics and culture in everyday life*. Cambridge: Cambridge University Press.

Law J. (1992) *Notes on the Theory of the Actor Network: Ordering, Strategy and Heterogeneity*. Available at: http://www.comp.lancs.ac.uk/sociology/soc054jl.html.

Lea M. (1992). *Contexts of computer-mediated communication*. New York: Harvester Wheatsheaf.

Luff, P., Gilbert, N. and Frohlich, D., Eds, (1990). *Computers and conversation*. London: Academic Press.

Mantovani G. (1996). Social context in human-computer interaction: a new framwork for mental models, cooperation and communication. *Cognitive Science* 20: 237-269.

Mantovani G., Spagnolli A. (2001). Legitimating technologies. Ambiguity as a premise for negotiation in a networked institution. *Information, technology and people*, 14 (3): 304-320.

Nardi B. (1996). *Context and consciousness. Activity theory and human-computer interaction*. Cambridge, MA: The MIT Press.

Ochs, E., Schegloff, E.A., Thompson S. A., Eds, (1996). *Interaction and grammar.* Cambridge: Cambridge University Press.

Reason J.T. (1990) *Human error*. Cambridge: Cambridge University Press.

Rasmussen J. (1980) What can be learned from human error reports?, In K. Duncan, M. Gruneberg, D. Wallis (eds) *Changes in working life*. London: Wiley.

Rogoff B., Lave J. (1984) (eds) *Everyday cognition: Its development in social context*. .Cambridge, MA: Harvard University Press.

Shotter J. (1983). On viewing videotape records of oneself and others: A hermeneutical analysis. In P. W. Dowrick, S. J. Biggs (eds), *Using video. Psychological and social applications*. Chihester: John Wiley and Sons, 199-210.

Smagorinsky P. (1998). Thinking and speech and protocol analysis. *Mind, culture and activity,* 5 (3), 157-177.

Stanney, K.M., Mourant, R.R. & Kennedy, R.S. (1998). Human factors in virtual environments: A review of the literature. *Presence.* 7 (4), 327-351.

Steuer, J. (1992). Defining Virtual Reality: Dimensions Defining Telepresence. *Journal of Communication,* 42(4), 23-72.

Suchman, L. (1987). *Plans and situated actions. The problem of human-machine communication*. New York Cambridge University Press.

Suchman L. (1995) *Making work visible*. Communications of the ACM 38 (9) : 56-64.

Wann, J. & Mon-Williams, M. (1996). What does virtual reality NEED?: human factors issues in the design of three-dimensional computer environments. *International Journal of Human-Computer Studies,* 44, 829-847.

Winograd T., Flores S. (1986) Understanding computers and cognition. Norwood, NJ: Ablex.

Zucchermaglio C., Bagnara S., Stucky S. U. (eds) (1995) *Organizational learning and technological change*. Berlin: Springer Verlag.

TRANSCRIPTION CONVENTIONS

(base on the code elaborated by Gail Jefferson; for a broader version, refer to Ochs, Schegloff and Thompson, 1996, pp. 461-465).

[[point of overlap onset at the start of an utterance

[point of overlap onset

=   latched utterances

(0.5)   pause, represented in tenth of a second

(.)   micropause

:   stretching of the preceding sound

_: falling intonation contour

:   rising intonation contour

.   falling or final intonation contour

-    cut-off or self-interruption

↑↓   sharp rise/fall in pitch or resetting of the pitch register

word   emphasis; represented by the length of the underlining

WORD   especially laud sound

°°   softer sound

hh   marked expiration, whose length is represented by the number of letters

(h)   expiration within a word (e.g. while laughing)

.h   inspiration

(( ))   transcriber's descriptions of events (e.g. cough, telephone rings,) or non-verbal actions

><   compressed talk (rushed pace)

<>   stretched talk (slowed pace)

(word)   uncertain identification of the word

(word A)/

(word B) alternative hearings of the same strip of talk

(   )   inaudible talk; the distance among the brackets should represent the length of the missing talk

,   'continuing' intonation

?   rising intonation

¿   mild rising intonation