

Dwell-Based Pointing in Applications of Human Computer Interaction

Christian Müller-Tomfelde

CSIRO ICT Centre, Cnr Pembroke & Vimiera Roads,
Marsfield NSW 2122 Australia
Christian.Mueller-Tomfelde@csiro.au

Abstract. This paper describes exploratory studies and a formal experiment that investigate a particular temporal aspect of human pointing actions. Humans can express their intentions and refer to an external entity by pointing at distant objects with their fingers or a tool. The focus of this research is on the dwell time, the time span that people remain nearly motionless during pointing at objects. We address two questions: Is there a common or natural dwell time in human pointing actions? What implications does this have for Human Computer Interaction? Especially in virtual environments, feedback about the referred object is usually provided to the user to confirm actions such as object selection. A literature review and two studies led to a formal experiment in a hand-immersive virtual environment in search for an appropriate feedback delay time for dwell-based pointing actions. The results and implications for applications for Human Computer Interaction are discussed.

Keywords: Pointing gesture, dwell time, interactive systems.

1 Introduction

Some of humans' intentions can be gained in real life situations from their actions and the context in which they are performed. An important and frequent gesture is the manual pointing action that allows humans to refer naturally and intuitively to distant objects in the environment. In a more general definition, pointing by a human at an object is understood as "how one organism manipulates the visual attention of another to some distant entity" [13]. Manual pointing can be considered as a "referential act" and therefore, represents a basic means to communicate with others beyond speech. This embodiment of communication makes use of implicit references, whereby movements of the body bind objects in the world to cognitive programs [3]. A single human can easily coordinate and execute actions and it requires little effort for another human to predict the actions and recognise the intention merely by watching the movements. Therefore, pointing represents a basic and ubiquitous device to communicate with others [14] and is often used in a communicative situation to establish a common ground over the course of the conversation [2]. Consequently, this also has implications for ways that humans interact with computers. It is important to understand what "behaviour" has to be exhibited by the computer in

order to meet expectations of users, especially when a computer should operate like a dialogue partner or accept more intuitive input signals from humans. The concrete question about how long users dwell on a target to express their selections and what an appropriate delay time for a corresponding feedback is, has not been discussed in the related literature in sufficient details. It is the central subject of this paper.

The paper is structured as follows. First, we describe the motivation and review the scientific context for this research. Then two exploratory studies are presented that further provide data related to typical manual dwell actions. The information of the review and the findings of the exploratory studies lead to a qualitative experiment to test hypotheses about the subjective experience of participants performing pointing actions. The paper concludes with a discussion about the results and a recommendation for a feedback delay time for dwell-based interaction.

2 Pointing in the context of HCI

In the typical situation of human computer interaction the user moves a physical device so that a screen cursor is placed over an element of the two-dimensional graphical user interface. With respect to the prior definition of pointing, the user's referential act raises the "attention" of the computer (as a dialogue partner) for the object located under the screen cursor. After placing the cursor on the target, the user usually clicks on a mouse button and therefore explicitly signals the intention to execute the function or the command that is associated with that interface element. This overall interaction is known as "point and click". Furthermore, an interaction technique referred to as hovering is used in situations in which a mouse-click is not applicable or available for the user, e.g., during "drag and drop" or in pen-based interfaces [10]. Nevertheless, to select a function, the user must hold the pointing device motionless for a certain period of time, to trigger a "hover event" by temporal discrimination. This provides information about the object in the focus of the attention. The "act" of hovering replaces the explicit click and allows for the selection of the target object.

In non-WIMP (Windows Icons Menus Pointing devices) based computer systems, the explicit act of clicking becomes substituted either by an event of another modality or by an event created by discriminating temporal segments. In multimodal research, information of multiple modalities becomes fused to model the content of the information at a high level of abstraction [19]. Furthermore, in virtual environment applications, the pointing gesture provides an intuitive access to distant objects using a laser pointer metaphor [5, 6]. This paper concentrates on three-dimensional pointing interaction using exclusively movements of a hand or a tool.

2.1 Components of the pointing gesture

In order to clarify what part of the pointing gesture is addressed by this paper, we provide a detailed description of a pointing gesture. The pointing gesture is a process that consists of three major elements or sections. Firstly, persons have to position and orient their pointing tools towards the object at which they intend to point. Secondly,

persons remain motionless for a certain period of time while holding the pointing device towards the target. This period of time is called the *dwell time*. It is assumed that a pointing gesture can only be recognised when the clear separation between these two periods is possible for an observer. The observer combines the information about the object the person is pointing at with the “act” of dwelling and recognises the selection of the target object. Thirdly, the person starts to move the pointing device away from the focused target. This idealistic temporal structure of a pointing gesture is less overt in real life situations. For instance, the transition between the first and second phase can happen multiple times when the person adjusts the orientation of the pointing tool. In that case, the tool does not remain motionless during the second phase. Another ambiguous situation can occur when the person aimlessly dwells in a resting position of the hand. Moreover, the definition of motionless is relative and depends on the perspective of the observer. The physical velocity of the pointing tool is a good candidate for a parameter for discrimination. The particular discrimination level of velocity has to be determined in specific applications.

The first phase of the gesture, i.e., moving toward a target can usually be described using Fitts’ law [7]. In contrast, the focus of this work lies on the second phase of the pointing gesture, the period when the person holds the tool nearly motionless on the target. In some applications, feedback is provided to the user to indicate that the target has been recognised. For these situations we call the time from the beginning of the dwell period until the feedback stimulus (S) the feedback delay time (FT) (see Figure 1). It must be smaller than the dwell time (DT). Finally, we call the time the user remains on the target after she has received a confirmation feedback stimulus about the recognition of the gesture the exit time (ET), similar to the definition in [27]. The exit time can also be considered as a simple reaction time of the user to the feedback stimulus, as described, for example, in [1].

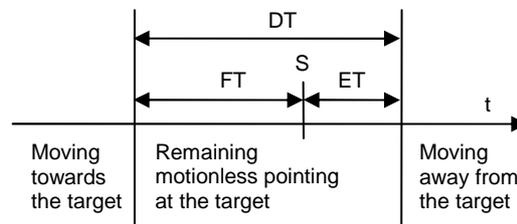


Figure 1. The temporal structure of a human pointing gesture with confirmation feedback. The dwell time (DT) is the sum of the feedback delay time (FT) and the exit time (ET) after the feedback stimulus (S).

Touching or clicking on an object with the hand or tool can be understood as an explicit interaction, whereas pointing at objects must be considered as a more implicit interaction. It requires a method to decide whether a selection has been made or not. A particular problem that can occur during these sorts of interactions is the *Midas Touch Problem*, coined by Jacob in the context of interaction systems based on eye movement tracking [11]. This means that, for manual pointing, unwanted selections can happen without a users’ intention, and, for instance, the context of the interaction has to be taken into account to further disambiguate the situation.

2.2 Pointing in HCI applications

With the emergence of the notion of disappearing computer and intelligent environments, pointing interaction is now a popular choice. Most applications that are using pointing interactions address issues of novel forms of interaction in information enriched environments or with large computer displays. An early example of an interactive system that involves pointing in combination with other modalities, such as speech, is Richard Bolt's "Put-that-there" system from 1980 [5]. In the literature several applications can be found in which the time the user remains motionless pointing at an object is used to control an application. The research problem of detecting a pointing gesture is usually tackled using the spatial as well as the temporal information of the users' actions.

In literature we reviewed, the time the user must remain motionless to allow the system to detect an act of selection varies. The range of the dwell time begins with 300 ms [20, 4], continues with 500 ms [16], 1 second [21, 28, 14] and stops at values of about 2 second [18]. The adjustment of the dwell time is usually based on ad hoc experiences of the developer. Formal evaluations have been done only in applications that use eye movements to control a cursor [24, 27]. Manual pointing interactions are used, for instance, to control home entertainment components, such as CD players, televisions, amplifiers etc. [28]. In this, the core interaction technique is based on pointing with a dedicated tool or wand towards the component the user wants to control. The user points at the target and holds the device motionless for a moment or presses a button to select the target. In an experimental setup, participants received an audio feedback when the pointing device was held motionless for "a brief period (less than one second)". Aspects of wall-based laser pointer interaction are examined and reported in [21]. In that study participants were asked to aim at a target with the laser pointing device. A dwell time of about one second was proposed to acquire the target. However, the selection of this dwell time was basically motivated by the accuracy constraint of the underlying technology and was not based on user preferences.

In other approaches for multimodal interaction in virtual environments, the accuracy of the pointing gesture detection increases by fusing hand and head orientation detection [18]. The authors presented work on real time tracking for 3D pointing interaction using video cameras. They stated that the average length of the pointing gesture was 1.8 s. The motionless phase revealed the highest duration variance (from 100 ms to 2.5 s). In [16] dwell-based interaction is used in combination with a force feedback device in the context of the training of motor skills. A dwell time of 500 ms is used here to discriminate gestures using a hierarchical network of sensors of activity. [14] describes another interactive application with large displays that rely only on the manual pointing gesture of the user. The author suggests a dwell time on the target of approximately one second to transform the "act" of dwelling into a "button-selected" event. However, the author did not give any further background information about this choice for the time period. In a study about interaction on tabletop computer displays with objects on the far side of a table, a dwell of the input cursor was also used to indicate a selection [20]. The dwell time there was set to 300 ms but wasn't studied further, since the authors primarily addressed issues of Fitts' law. A comparison of the performance of the selection by mouse-click and by mouse-dwelling was a subject of the study presented

in [9]. The authors stated that the interaction techniques based on a mouse-click was on average about 150 ms faster than the selection by dwelling. The dwell time activation introduced more errors and was therefore less efficient. A large number of subjects also reported that the fixed dwell time of 500 ms was too short. Therefore, the authors recommended a dwell time of 750 ms for novices and proposed an adaptation of the dwell time for individual users but a formal evaluation was not reported.

2.3 Temporal aspects of actions

The prevailing approach in the context of temporal aspects of human actions is the hierarchical organisation of human behaviour in time scales [17, 3]. The scales start at the neuron activity level where it takes about 1 millisecond to generate an electrical signal to communicate with another neuron. The scale ends at the cognitive level with period of 2-3 s [3]. The interesting interval for the work and experiment presented in this paper is assumed to be in the range of a couple of hundreds of milliseconds to about one second, since most of the previously described approaches can be found here. We now present results of related work which concentrates on the perception of the delay between an intentional action and its effect in the environment.

In [8] the authors report on an experiment that was addressing the issue of intentional binding between voluntary actions and their effects. The Participants' task was to press a key and to receive a delayed auditory feedback. They were asked to judge the onset times of their actions and the perceived feedback. The onset time of pressing the key was judged to happen later while the onset time of the feedback was judged to be earlier. At delay times between the key press and the auditory feedback of about 250 ms, the participant perceived the time gap between both events to be relatively shorter than for a delay time of about 650 ms. In other words, with a decreasing delay time between the key press action and the auditory feedback, the "perceptual attraction" or binding effect became increased. The authors argued that this phenomenon helps humans experience their own actions and provides a sense of controlling events in the outside world.

In another recent study about the action-effect relationship [23], participants were asked to which extent they felt that an effect was a consequence of their own action or of the action of the experimenter. Different temporal delays of up to 600 ms were introduced between the action and an auditory feedback. The experiment revealed that the feeling of being the author of the effect decreased with increased temporal delay. In the context of pointing and selecting of objects, it is assumed that a short or zero delay, for instance, causes a strong feeling of being author of the selection, while a longer delay suggests that the selection might be done by someone or something else.

Based on the idea of the hierarchical organisation of human behaviour suggested by Newell [17], Ballard *et al.* [3] argue for a time scale that defines a special level of abstraction. The level is called "Embodiment Level" because the constraints of basic physical acts influence the nature of cognitive operations at a time period of about a third of a second. It is at this level that primitives, such as movements of the eyes or taps on a key of the keyboard happen [1]. These primitives are distinguished from those of the next higher cognitive level, such as typing a word or dialing a number

[3]. We assume that a pointing action is more than basic movement primitive since it includes also cognitive aspects. An appropriate temporal scale of these referential primitives is expected to be greater than the Embodiment Level (300 ms), while it should be less than the temporal scale of action of a higher cognitive level (2-3 s).

In the research area of hypertext presentation and exploration, Meyer *et al.* presented a study [15] about the relationship of system response delay and the time the user is watching a hypertext page before moving onto a new page. The experiment revealed interesting effects under the condition of different delay times between the click on a hyperlink and the appearance of the corresponding new page. The time the participants remained on the page was correlated with the time they waited for the page until a response delay of 3 s. Furthermore, the emotional workload increased above this temporal threshold. The results underpin the existence of a universal constant of segmentation for the temporal integration of successive events into perceptual units of about 2 to 3 s duration [22]. In [25], Schleidt *et al.* presented a study about the temporal integration of successive events and short term movements. Videotapes of behaviour scenes of different cultures were analysed. The result of the study supported the assumption that both the human perception and the “overt behaviour is characterised by a similar time constant”. This refers also to a temporal window of 2 to 3 s, in which the “short term movement episodes” with obvious segmentations happen. The question that arises from the latter two studies is whether or not the motionless time period while pointing has a similar constant character.

In brief, the literature we studied strongly indicates a time period of about 500-1000 ms for an appropriate feedback delay time for dwell-based pointing actions. At this delay the intentional binding between action and effect still supports the users’ experience of controlling the environment. The suggested time span is confined by particular time scales of human information processing and the majority of existing HCI approaches use delayed feedback below 1 s. Assuming a standard reaction time to a visual stimulus of about 300 ms (exit time, see Figure 1) a resulting natural dwell time is expected to be about 800-1300 ms.

3 Exploratory studies

To gain further knowledge about temporal aspects of dwelling on targets during pointing actions we conducted two exploratory studies in a similar interactive environment. We separated the process of passing on a reference from an actor to an observer into two parts. Study A was about performing a pointing action, while study B focused on the observation of such an action. The studies were done independently from each other, no direct or mediated verbal communication happened between the actor and the observer, except pointing gestures. It was assumed that the average dwell times of actors are greater than the average response times of observers or receivers. Instead of imposing a particular dwell time on the participant by providing a feedback for a selection, the actors performed simply based on of their inherent temporal organization of pointing actions. The idea of the studies was that the results would provide further indications to identify an appropriately time period for feedback for the interaction with computers.

3.1 Study A: Performing a pointing action

We chose a hand-immersive virtual environment for this study. The participants interacted with a stylus in a 3D scene using shutter glasses providing stereo vision. The input device reliably tracked the participants' interactions. The study was conducted with 15 unpaid participants recruited from the research organisation of the authors. Ten of the participants were male and five female. The task of the participants was to answer questions about the two most common colours of objects by pointing with a tool at the correspondingly coloured objects in the environment. For example: What are the two most common colours of flowers? As a possible answer, the participant pointed at the green object and then at the red object. Furthermore, participants answered 50 of these questions, each with a different topic and were asked to perform their tasks in a comfortable and relaxed manner. The participants were advised to perform the task so that another person would be able to understand their selections. The participants were given a verbal introduction about the environment. Written instructions about the task were handed out, and questions about the task were discussed. We observed the first few interactions then allowed the participants to finish the study alone.

3.2 Result study A

The result of study A can be illustrated in a histogram of the detected dwell times with an interval width of 333 ms (see Figure 2).

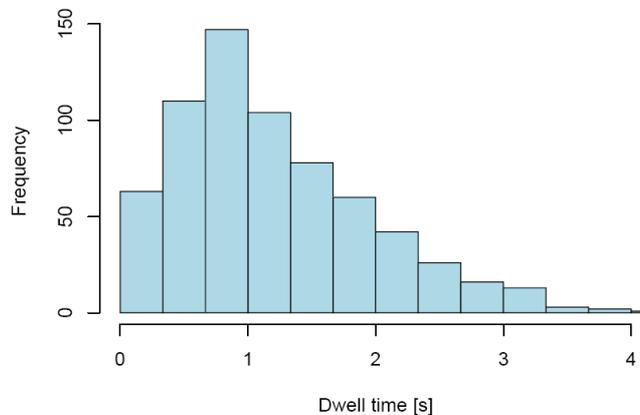


Figure 2. Histogram of the detected dwell times of all participants in study A. The histogram is skewed to the right.

The histogram is right-skewed, because only positive dwell times can occur. The skewness of the distribution is 0.9602 and standard deviation is 0.7575. Data analysis revealed an average dwell time of 1.175 s and a median value of 0.999 s. 50 % of the data lies between 0.599 and 1.598 s. Because of the asymmetric shape of the histogram the average value is greater than the median value and the latter is more

appropriate for description. The expected numbers of dwell events were 1500. Five participants touched target objects rather than pointed at them and were omitted for the analysis. The decreased number of dwell events (1000) is further reduced due to insufficiently detected dwells. Therefore, the total number of dwells contributing to this histogram is 702.

3.3 Study B: Observing a pointing action

Study B was done to understand how long observers need to receive the reference of a pointing action by watching the action. A standard desktop computer with a keyboard and an LCD panel was used to collect data about the reactions of the observers. Nine unpaid participants contributed to this study, 2 female and 7 male. The participants were asked to watch on the screen an animated typical trial of one participant of study A. The action was represented only by the replay of the recorded movement of the pointing tool. The task was to press the space bar of the keyboard to confirm the 100 selections that had been made by the person moving the stylus. The participants were asked to perform their tasks in a timely but comfortable and relaxed manner. The participants were given instructions about the task and questions were discussed. We observed the first few reactions then restarted the experiment and the participants finished the study alone.

3.2 Result study B

The result of study B can be illustrated in a histogram of the detected response times with an interval width of 100 ms (see Figure 3).

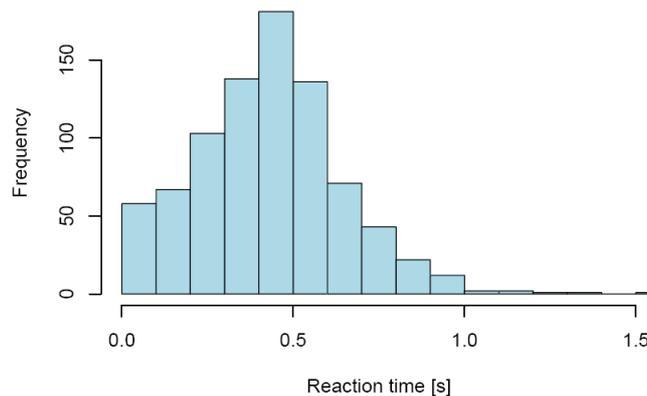


Figure 3. Histogram of the response times of all participants observing a recorded and animated trial of one participant of study A.

In the study 855 confirmations of the 9 participants had been recorded. The missing 45 confirmations are due to the fact that some selections had been overseen

by participants probably because of drops in concentration or because of misinterpretations. This inaccuracy of 5 % will be ignored for the following considerations. All confirmations were set into relation to the start point of the dwelling activity, that is, when stylus started remaining nearly motionless while it was held towards the target object. 17 confirmations (2 %) revealed negative reaction times and were ignored also for further analysis. The histogram of all 838 confirmation time periods can be seen in Figure 3, where 50 percent of the measured reaction times occurred between 0.292 s and 0.552 s. Data analysis revealed an average reaction time of 0.4325 s and the same value as median. The standard deviation is 0.215 and the skewness of the distribution is 0.5484.

The independent studies A and B showed that the median dwell time (0.999 s) is greater than the median response time (0.4325 s), although there was no feedback given to the pointing participants and the studies were independently conducted. The variances are relatively high compared to those other response time experiments because of the self-paced character of the studies without an explicit feedback stimulus to react to. Nevertheless, the studies indicated a possible existence of a natural dwell time and suggest a corresponding feedback delay time of about 0.4325 s. In the next section the latter period of time is further investigated in a formal experiment that uses an explicit and temporally controlled visual feedback.

4 Experiment with visual feedback

Based on the results of the literature review and the exploratory studies we formulate hypotheses. We want to test whether a variation of the delay of an explicit visual feedback for a pointing action has an effect of the perception of the interaction process. First, feedback delay time above approximately 430 ms is experienced by users to happen late. Second, for a feedback delay time above approximately 430 ms users experience waiting for feedback to happen and third, feedback delay below 430 ms is considered by users to be natural as in real life conversations.

In order to test the hypotheses the interaction environment of study A was modified to provide a discrete visual feedback. A simple mechanism was implemented to detect dwell events and the feedback delay time was adjustable. The design of the experiment had one within-participant factor, the feedback delay time (FT, see Figure 1). The dependant variables were the ratings of a questionnaire. The 15 participants of the experiment were the same as in study A.

The experiment was conducted in 8 blocks each with 10 questions about the two most common colours of objects. The task of the participant was similar to that of study A except that they were told that their answers should be confirmed by the computer. As a feedback the colour of the selected object would change to white. The feedback delay time was set in block 1 to 100 ms and was incremented for each block by 250 ms. The last delay time of block 8 was 1850 ms. The participants were not informed about these changes but rather encouraged to act as in study A. After each of the 8 blocks the participants were asked to fill out a questionnaire with possible answers on a 7 point Likert scale. The questions were:

- Question 1: Do you have the impression that the system feedback happened in a reasonable time according to your action?
Answer: confirmation occurred too fast (1), too late (7).
- Question 2: Did you have the feeling to wait for the feedback to happen?
Answer: no I didn't have to wait (1), yes, I waited (7).
- Question 3: Did you have the impression that the time delay for the feedback was natural? (i.e., as in a real life communication situation)
Answer: time delay is not natural (1), quite natural (7).

The questionnaire was discussed with each participant. Then the first couple of interactions were observed by the experiment supervisor while the participant finished the task alone.

4.1 Results of the experiment

The average ratings for the three questions are shown in Figure 4. There is a trend in the average values of the ratings under the changing condition of the feedback delay time.

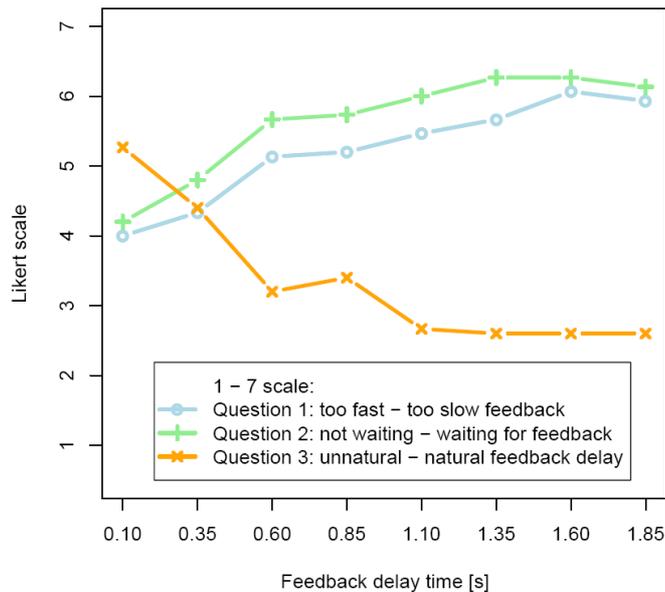


Figure 4. Results of the questionnaire for the experiment. The ratings show a general trend and the largest changes can be reported between 350 and 600 ms.

The effects of feedback delay time on the three ratings were tested by one-way repeated measures analysis of variance (ANOVA). There was a significant main effect in the ratings of question 1, $F(7,98) = 11.289$, $p < 0.01$, question 2, $F(7,98) = 6.1828$, $p < 0.01$ and question 3, $F(7,98) = 12.667$, $p < 0.01$. The ratings for questions 1 and 2 increase, while ratings for question 3 decrease with the increasing

independent variable. The largest changes of the ratings happen between 350 and 600 ms of feedback delay time (see Figure 4).

We carried out further detailed analysis of the ratings. Several paired t-tests were calculated for the ratings of the questions. The tests revealed that the ratings of question 1 at a feedback delay time of 100 ms are not significantly different to those at 350 ms. In contrast, all other paired t-test between rating of question 1 at 100 ms and those equal and above 600 ms are significantly different. This step of significance between 350 and 600 ms is also valid for the ratings of question 3. The step in ratings of question 2 also exists but the p-values are not significant.

The results of the experiment can be summarized as followed. At a delay time of 100 ms the feedback is rated as neutral (4), i.e., the feedback is neither too fast nor too slow (Question 1). Question 2, waiting for feedback, is also judged neutral. There is a tendency to judge the system feedback as natural at the feedback delay time of 100 ms (Question 3). Towards the delay of 1850 ms the feedback was rated less natural and feedback time as too slow. On average the participant also rated that they waited for the feedback. At the step in significance between 350 and 600 ms we assume that the judgments of the participants are about to change.

As an additional outcome of the experiment the exit time is calculated from the interaction data. The dwell time of each recorded pointing action was reduced by the associated feedback delay time (see Figure 1). The average exit time is about 0.4562 ms with a standard deviation of 0.3814. The median value is 0.3380 s while 50 % of all exit times are between 0.2890 and 0.4470 s. The median exit time as a response to the feedback stimulus can be considered as a standard movement response times of about 330 ms.

5 Discussion

The findings of the two studies and the results of the experiment provide empirical evidence that there is a possible common or *natural dwell time* in human pointing actions that implies a corresponding delay time for feedback. The results of the experiment show that for a feedback delay below 600 ms it is likely that the feedback is experienced as natural and that the users rate the feedback neither too fast nor too slow. The significant differences of the ratings at 350 and 600 ms mark an interval in which judgments of the participants are about to change. Furthermore, the interval encloses the average feedback delay time of 430 ms gained from study B and matches some delay times proposed in the literature. However, these had been constrained by technology instead of being based, like ours, on empirical evidence. Although our study was designed with high generalisability, it is clear that the findings strictly hold for the particular environment and task that was used. However, the applications reviewed in the literature resemble in many aspects the setup used in this study. We assume that our results can be even compared to those of experiments with manual 2D interaction environments with mouse and pen input devices.

In order to further link the results of the studies and the experiment we introduce the notion of an *internal stimulus*. The internal stimulus is a hypothetical event that marks the point in time when the participant's intention to act emerges. Then the

motor programming stage prepares the physical action that finally becomes executed. We assume now that the process that follows the internal stimulus is equivalent to that which follows the external stimulus in the experiment with feedback. In other words, the motor programming that leads to the movement of the hand is assumed to be the same in study A and the experiment. Therefore, the time that elapses from the start of the dwell period until the participant experiences an internal stimulus to act can be deduced. We subtract the exit time of approximately 340 ms gained from the experiment from the natural dwell time of study A, 1 s. The resulting period for the internal stimulus is approx. 660 ms. The difference of about 230 ms between this derived internal stimulus time and the external stimulus time of study B is probably due to the high skewness of the data of study A. It is more likely that the natural dwell time is less than 1 s and is located in the centre of the mode at about 830 ms (see Figure 2). Future studies have to provide more evidence.

The detection of a motionless period always requires time and will introduce a temporal delay. A feedback delay time of 350 ms to 600 ms gives detection mechanisms in dwell-based pointing applications enough time to derive the referred object from the real time interaction data and to avoid the ambiguity of short dwells. At the same time the acceptance of the feedback based on the rating is still high.

The idea of a natural dwell time is not entirely new and has been motivated by the research presented in [25, 22], which favours a universal time constant in short term human actions and by the notion of the Embodiment Level in human behaviour as proposed in [3]. Although time periods in human behaviour cannot be literally constant, we argue that the natural dwell time in pointing could indicate the existence of a particular foundational time scale for human communication. Therefore, it is important to consider the natural dwell when designing the interaction of humans with computers.

6 Conclusion

We conducted two exploratory studies and a formal experiment focusing on the dwell time of the human pointing gestures. We were focusing mainly on applications of Human Computer Interaction where no explicit and discrete actions, such as button clicks are practical or available to capture user input. Our aim was to provide empirical evidence for a possible natural dwell time during pointing to express the selection of a target object and to convey this reference to a communication partner. The findings of our studies were combined with the results of a pointing experiment with delayed visual feedback for selection. This allows us to recommend a feedback delay time for manual pointing actions of approximately 350 to 600 ms as a starting point for the development of interactive applications. We have shown that this feedback delay is experienced by users as natural and convenient and that the majority of observers of pointing actions gave feedback within a similar time span. Any further delay is counterproductive and impedes the progress of a task [26]. Therefore, it is not recommended to let technical constraints bias the choice for a feedback delay time. Instead, the user's behaviour and expectations as studied in this paper should inform the system design to achieve a high interaction quality.

Acknowledgements. The author would like to thank Carl Mistler for his valuable contribution in building the tool to play back the interaction data. Thanks also to Cara Stitzlein and Claudia Schremmer for their insight and critical review of this manuscript.

References

1. Card, S.K., Moran, T.P. and Newell, A. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum, Hillsdale, N.J., (1983).
2. Clark, H. and Wilkes-Gibbs, D. Referring as a collaborative process. *Cognition* 22, (1986), 1-39.
3. Ballard, D.H., Hayhoe, M.M., Pook, P.K. and Rao, R.P.N. Deictic codes for the embodiment of cognition, *Behavioral and Brain Sciences*, 20 (1997), 723-767.
4. Bohan, M., Chaparro, A. and Scarlett, D. The Effects of selection technique on target acquisition movements made with a mouse. In *Proceedings of the 42nd Annual Meeting of the Human Factors Society*, Human Factors and Ergonomics Society (1998), 473-475.
5. Bolt, R.A. Put-that-there: Voice and Gesture at the Graphic Interface. In *Proceedings of Computer Graphics (SIGGRAPH'80)*, 14, (1980), 262-270.
6. Bowman, D.A. and Hodges, L.F. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the Symposium on Interactive 3D Graphics* (1997), 35-38.
7. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Experimental Psychology*, 47 (1997), 381-391.
8. Haggard P., Clark S. and Kalogeras J. Voluntary Action and Conscious Awareness. *Nature Neurosciences*, 5 (2002), 382-385.
9. Hansen, J.P., Johansen, A.S., Hansen, D.W., Itoh, K. and Mashino, S. Command without a click: Dwell time typing by mouse and gaze selections. In *Proceeding of INTERACT 2003*, IOS Press: Amsterdam (2003), 121-128.
10. Hinckley, K., Jacob, R.J. and Ware, C., Input/Output Devices and Interaction Techniques. In *The Computer Science and Engineering Handbook*, A. Tucker, ed., CRC Press, 2004.
11. Jacob, R.J. What You Look at is What You Get: Eye Movement-Based Interaction Techniques, In *Proceedings CHI 1990*, ACM Press (1990), 11-18.
12. Leavens, D.A. Manual deixis in apes and humans. *Interaction Studies*, 5 (2004), 387-408.
13. Kita, S. Pointing. A foundational building block of human communication. In *Pointing: Where Language, Culture, and Cognition Meet* (Kita, S., ed.), Mahwah: Lawrence Erlbaum, (2003).
14. Malerczyk, C. Interactive Museum Exhibit Using Pointing Gesture Recognition. In *Proceedings of the 12-th International Conference in Central Europe on*

Computer Graphics, Visualization and Computer Vision'2004, Short Communications Volume II. (2004), 165-171.

15. Meyer, H.A., Hänze, M. and Hildebrandt, M. Das Zusammenwirken von Systemresponsezeiten und Verweilzeiten beim Explorieren von Hypertextstrukturen: Empirische Evidenz für einen zeitlichen Integrationsmechanismus? In I. Wachsmuth and B. Jung (Eds.), *KogWis99: Proceedings der 4. Fachtagung der Gesellschaft für Kognitionswissenschaft* (1999), 86-91.
16. Müller-Tomfelde, C. and Paris, C. Explicit task representation based on gesture interaction. In *Proceedings of the NICTA-HCSNet Multimodal User Interaction Workshop MMUI '05*, ACM Press (2006), 39-45.
17. Newell, A. *Unified theories of cognition*. Harvard University Press, (1990).
18. Nickel, K. and Stiefelhagen, R. Pointing gesture recognition based on 3D-tracking of face, hands and head orientation. In *Proceedings of the 5th international Conference on Multimodal Interfaces ICMI '03*, ACM Press (2003), 140-146.
19. Nigay, L. and Coutaz, J. A Design Space for Multimodal Systems Concurrent Processing and Data Fusion. In *Proceedings of the INTERCHI'93 Conference*, ACM Press (1993), 172- 178.
20. Parker, J.K., Mandryk, R. and Inkpen, K. TractorBeam: Seamless Integration of Local and Remote Pointing for Tabletop Displays. In *Proceedings of Graphics Interface*, (2005), 33-40.
21. Peck, C.H. Useful parameters for the design of laser pointer interaction techniques. *Extended Abstracts CHI 2001*, ACM Press (2001), 461-462.
22. Pöppel, E. A hierarchical model of temporal perception. *Trends in Cognitive Science*, 1(1997), 56-61.
23. Sato, A. and Yasuda, A. Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership, *Cognition*, 94, 3 (2005), 241-255.
24. Sibert, L.E. and Jacob, R.J.K. "Evaluation of Eye Gaze Interaction," In *Proceedings CHI 2000*, ACM Press (2000), 281-288.
25. Schleidt M., Eibl-Eibesfeldt, I. and Pöppel, E. A Universal Constant in Temporal Segmentation of Human Short-Term Behaviour, *Naturwissenschaften*, 74 (1987), 289-290.
26. Shneiderman, B. Response time and display rate in human performance with computers. *ACM Computing Surveys*, 16, 3 (1984), 265-285.
27. Špakov, O. and Miniotos, D. On-line adjustment of dwell time for target selection by gaze. In *Proceedings of the Third Nordic Conference on Human-Computer interaction 2004 NordiCHI '04*, ACM Press (2004), 203-206.
28. Wilson, A. and Shafer, S. XWand: UI for intelligent spaces. In *Proceedings CHI 2003*, ACM Press (2003), 545-552.