# Perception and Manipulation of Game Control

Danny Plass-Oude Bos, Bram van de Laar, Boris Reuderink,
Mannes Poel, and Anton Nijholt

University of Twente, Human Media Interaction Group,
P.O. Box 217, 7500 AE Enschede, The Netherlands
`d.plass@utwente.nl`

**Abstract.** Brain-computer interfaces do not provide perfect recognition of user input, for similar reasons as natural input modalities. How well can users assess the amount of control they have, and how much control do they need? We describe an experiment where we manipulated the control users had in a keyboard-controlled browser game. The data of 211 runs from 87 individuals indicates a significant linear correlation between users' sense of control and the amount of control they really had in terms of mutual information (not accuracy!). If users know what they put in, they can assess quite well how much control they have over the system. In our case, from an amount of control of above 0.68 bits in mutual information (a 5-class accuracy of 65%), this aspect of control no longer seems to be the critical factor for finishing the game. Deliberate manipulation of perception may offer a way to make imperfect, uncertain input modalities more acceptable, especially in combination with games.

**Keywords:** Human-computer interaction, brain-computer interfaces, manipulation of control, perception of control.

## 1   Introduction

Brain-computer interfaces (BCIs) provide systems with input based on the user's mind, so devices and applications can respond to specific mental states. Like other input modalities based on observations of the body, BCIs do not provide perfect recognition of what a user tries to convey [1,2,3,4]. These inputs suffer from problems related to noise, non-stationarities, and ambiguity [5]. This can be problematic, as input is the basis for usable systems in general, and recognition accuracy is most important to users of BCIs [6].

We focus on games specifically. A large part of the population plays games, and it is known that gamers are among the first to adopt new technology [7]. Learning a new skill, like providing brain-based input, could be part of the challenge of the game [8]. It comes as no surprise then that many of the current commercial BCI applications are game-oriented. For scientific purposes, games can help experiment participants to stay motivated and focused for longer periods [9].

We have done many demonstrations and experiments in which people could try our brain-computer interface (BCI) games. Sometimes people seemed to overestimate their level of control, and sometimes to underestimate it. This made us

wonder: how well can people assess how much control they really have? Additionally, what would be the minimum amount of control necessary to operate a given system?

Previous analysis of data from this experiment has been published in [5], which posed that perfect control may not always be optimal. People actually experienced more fun in the experiment game when the control was not perfect. Our focus in *this* paper is the *perception* of control. Additionally, we investigate how much control might be necessary so users don't give up.

## 2   Background and Related Work

*Perception of Control.* There are many reasons to suspect that there is no simple linear relationship between the perception of control and how much control people actually have. People overestimate their influence on things with a positive outcome, and underestimate their effect on negative outcomes [10,11]. Even when people have no control at all, they may experience 'the illusion of control' [12]. Additionally, people assess more beautiful systems as being more usable, even if they are not [13]. Norman goes even further, arguing that pleasing things (not necessarily through beauty alone) actually work better [14]. One way pleasing aspect might be using a novel input modality. People appear to be more lenient towards mistakes made by a brain-computer interface than towards errors made with a non-BCI input (in this case, a variation on mouse selection) [15]. For more on how the user experience may affect BCI performance, see [16].

*Controlled Simulation of Uncertain Control.* Brain-computer interfaces do not allow us to control the amount of control of a user over the full range from having no control at all to complete control. So we needed a substitute of which we could also be certain that the user has the level of control that we wanted to provide. We looked at various alternatives: simulating imaginary-movement-based BCI input [17], manipulating mouse input, [18], or issuing incorrect commands at selection level [19]. Carlson, et al. evaluated the effect of shared control (with an AI) on driving a BCI-controlled wheelchair [20] using another alternative. To make the evaluation less time-consuming, they decided not to use actual BCI control, but to simulate it with input transformation matrices. Such a matrix provides a probability for each input to transform into something else, which can then be used to actually transform input actions into other actions with different system responses. Then an 'expert driver' would *pretend* to control the wheelchair with a brain-computer interface, by simply pressing buttons on a keyboard. In the end, we opted for these *input transformation matrices*. Its simplicity allows us to assess only this aspect of the input. Additionally, it is easier to implement and adjust.

*Minimum Amount of Control.* It has been said that for BCIs a selection accuracy of 70% is acceptable [21]. Another BCI research group determined the minimum level of accuracy to be 77% for four input classes [19].

*Uncertainty in Applications.* There can be uncertainty about whether the input will be interpreted correcty by the input device, but uncertainty can also be purposefully introduced in an application [22]. For example, we observed that uncertainty can increase the sense of fun [5].

## 3   Methods

### 3.1   Experiment Protocol

To reach a large number of participants and gather enough data for each level of control, the experiment was run from a web browser, so people could participate from anywhere. Social media was actively used to get the word out.

The input was provided by, the widely available, keyboard. To manipulate the amount of control, each run, an input transformation matrix would be randomly selected from the database. This randomization had two benefits: (1) It allowed for a distribution of samples over the different levels of control, and (2) if the previous experienced level of control affects the experience of the current level, this avoids order effects in the results.

Each run started with an explanation on how to play the game. During a run, the player tried to guide a laboratory test hamster to freedom through four levels (Figure 1)[1]. After one minute, the player could decide to skip the rest of the run. At the end of each run, a questionnaire would pop up, after which the player would be encouraged to play another round.

### 3.2   Manipulation of the Input

The amount of control is varied by manipulating the keyboard input, which consists of the four directional arrows. If the user takes no action, a 'no action' input is generated, also known as 'idle state', 'no control', or 'no operation'.

This input is transformed by a matrix which dictates probabilities for each system response (what the hamster you control ends up doing) given a particular provided input. The probabilities for a non-matching output are set equal for all non-matching outputs. The amount of control is thus defined by the probability for correct classification (the hamster obeys), which is equal to the accuracy level.

We generated 15 input transformation matrices, evenly spread out over the whole possible range of *mutual information*. Mutual information is defined as the amount of information one sequence provides over another, in bits. In this case, the amount of information that is being shared between the actual input, and the transformed input with a lower accuracy. This measure may be a little less intuitive to interpret compared to accuracy, but it is more suitable, as it is comparable for different numbers of classes and different prior probabilities [23]. The relationship between mutual information and accuracy is logarithmic.

---

[1] The software and additional notes can be found on the following website:
`www.dannyplass.nl/control`
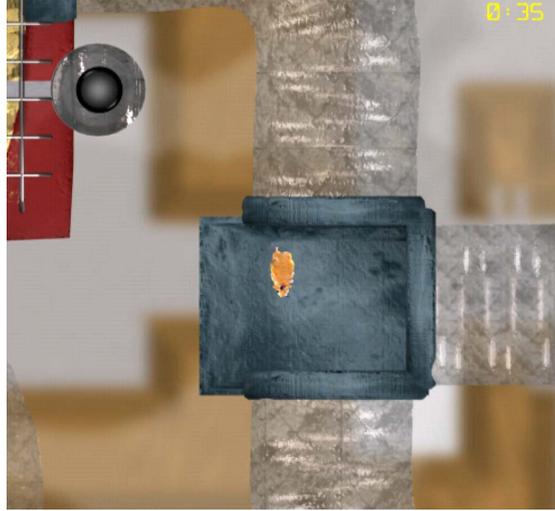
**Fig. 1.** A screenshot of the game used in the experiment

To determine the mutual information $I(X;Y)$ for a given input transformation matrix concerning inputs $X$ and responses $Y$, we assume a uniform probability function over the input space (verification in the Results section). This probability function is also known as the marginal probability function $p(x)$ for $x$. The mutual information is then computed as follows:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log_2 \frac{p(x,y)}{p(x)p(y)}$$

where $p(x) = \frac{1}{n}$, with $n$ the number of input types

### 3.3   Data Collection

The questionnaire consisted of 6 visual analogue scale (VAS) questions [24], and 3 open questions: age, gender, and a field for remarks. The VAS scales went from 0 to 100, and were initialized at 50. For this analysis, the two questions of importance are those related to control: *"I had the feeling that the hamster did what I wanted it to do"* and *"I had the feeling the computer was following my commands"*. Additionally, detailed action logs were maintained for each participant, containing all keyboard inputs and the resulting transformed actions, as well as starting, skipping, finishing, pausing, and resuming levels.

### 3.4   Participants

We could identify 87 individuals based on filled out gender-age answer pairs in combination with the IP addresses. Of these individuals 39% was male, 29% female, and 28% unknown, with an average (provided) age of 24.9 years (in the range of 10-58, with a standard deviation of 7.5).

For the main analysis concerning the perception of control, we excluded runs for which not all the VAS questions were filled out. This filtering resulted in 211 runs, with at least 9 runs and at most 22 runs per level of control.

To determine when people gave up, we looked at the action logs for all started runs, and analysed the final entries for each run, which gives an indication of how it was ended. This resulted in the analysis of 465 runs.

## 4   Results

The two main questions are: (1) How well can people assess their level of control, and (2) How much control is sufficient so users do not give up?

### 4.1   From Theory to Reality

The input transformation matrices were computed based on an equal occurrence for each input. In practice, there was indeed a fairly equal distribution among the classes (medians around 20%), but with a preference for 'right' (about 30%, due to level design), and a lower occurrence for 'no action' (around 10%).

How does this affect the amount of control people had? Based on the confusion matrix of observed inputs and into what system responses they were transformed, we computed the *observed mutual information*. The theoretical and observed mutual information are tightly correlated, see Figure 2.

We decided to use the medians of *observed* mutual information, instead of its purely theoretical counterpart, to group the data points for each input transformation matrix. These observed medians per matrix are the dots in Figure 2. This grouping of data allows us to provide box plots with more statistical information about the data. The exact details per run are lost in this approximation, but because of the close relationship between the theoretical and observed values, this effect should be minimal.

### 4.2   Sense of Control

The questionnaire contained two questions related to the user's sense of control: *"I had the feeling that the hamster did what I wanted it to do"* and *"I had the feeling the computer was following my commands"*. These items averaged together form the combined *sense of control* scale.

Figure 3 shows the sense of control results grouped per input transformation matrix represented by the median observed mutual information related to it.

Theoretical and observed mutual information
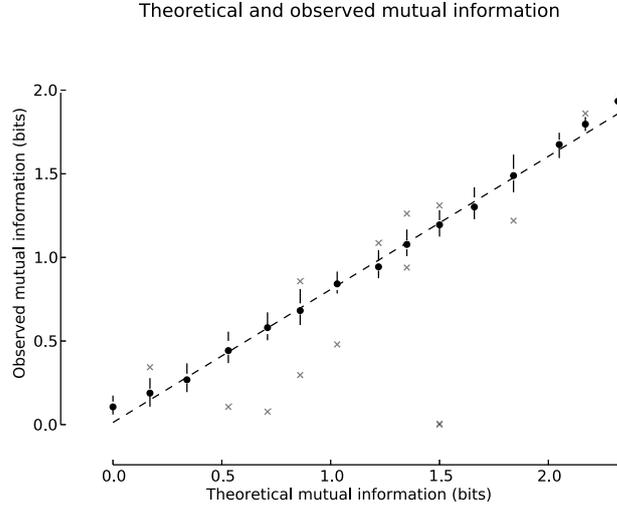


**Fig. 2.** The theoretical mutual information shows a tight linear relationship with the observed mutual information. *How to interpret Tufte box plots:* Each vertical 'bar' represents the minimum, lower quartile, median, upper quartile, and maximum. The dot indicates the median, and the inter-quartile range is visualized as the vertical white space around this dot. Crosses are outliers, which are beyond 1.5 times the inter-quartile range from the lower and upper quartiles.

The strong and significant fit of the linear regression analysis between mutual information and sense of control ($\beta = 36.51, p < 0.001$) indicates that people are quite good at estimating their level of control.

How well does *accuracy* do as a predictor of sense of control? Again, we use the actual accuracy as observed from the interaction logs. Accuracy is a less accurate linear predictor of sense of control than mutual information, explaining 67% of the variance as opposed to 72% (with $p < 0.001$, the same). The medians indicate an exponential relationship, which is to be expected based on the logarithmic relation between mutual information and accuracy.

### 4.3   Sufficient Control

The amount of frustration decreases when the amount of control (in mutual information) increases ($\beta = -23.35, p < 0.001$). However, this does not tell us the minimum amount of control users need. One could put an imaginary boundary at some level of the VAS item, but what level of *frustration* is unacceptable?

Another source of information on how much control is needed is the way runs were ended. Participants could simply leave the website, or they could wait a minute and then skip to the questionnaire by pressing a button, or they could finish the level by bringing the hamster to safety. At the necessary amount of
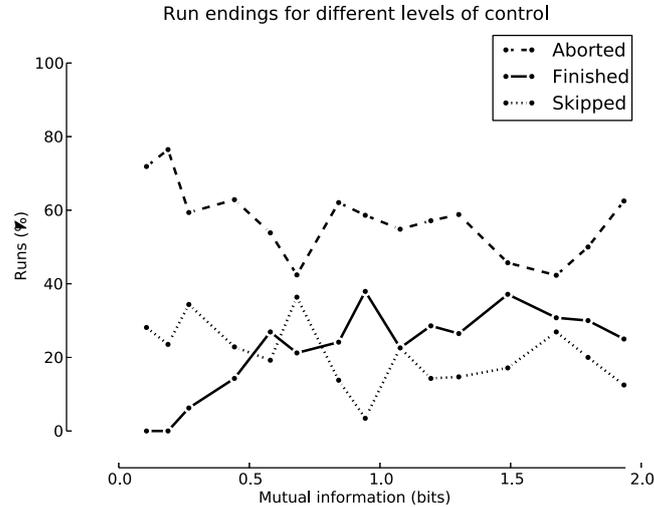
Sense of control scale
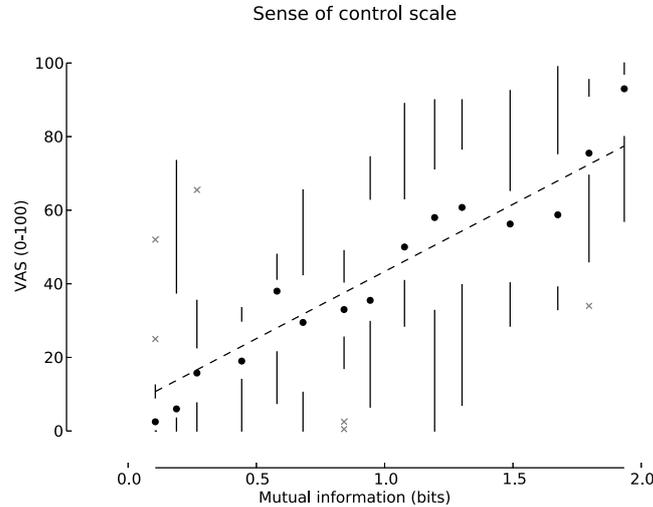


**Fig. 3.** The relationship between median mutual information (representing the input transformation matrices) and the combined control scale. A linear regression analysis (the dashed line) showed that mutual information is a highly significant predictor of the sense of control scale ($\beta = 36.51, p < 0.001$), accounting for 72% of the variance. The indicated outliers were included in the regression analysis.

control, more runs should be finished, and less runs should be aborted. People can have various reasons not to finish a run, which are not related to control. Perhaps they did not like the game enough, or something more important came up. As long as they are not connected to the amount of control, we can assume that it will have an equal chance to occur at each of the levels of control.

Figure 4 shows the different ways runs were ended for each of the different levels of control. The numbers of aborted and skipped runs slightly decrease with increasing control, and the amount of finished runs increases accordingly. Surprisingly, the effect of the amount of control does not seem to be very strong, especially for higher levels of mutual information. However, on the low control side, up to a mutual information of 0.68 bits, there does seem to be a clear effect, with a steep decrease in aborted runs, and a similar increase in finished runs. This seems to indicate that up until this point, the amount of control was the critical reason to stop playing. Beyond this level of control, other unknown, but control-independent, reasons seem to become dominant as the percentages become more steady. The increase in aborted runs after 1.5 bits could be related to the decrease in fun participants experienced when the level of control gets (close to) perfect – see our previous data analysis in [5].

This critical point where the amount of control is no longer a critical factor for finishing the game is at about 0.68 bits. This corresponds to 65% observed accuracy for 5 classes. At the 5 inputs per second this game allowed, this is an information transfer rate of 204 bits per minute.

**Fig. 4.** The percentage of different ways of ending a run for different levels of control

## 5   Discussion and Conclusions

### 5.1   Sense of Control

People are quite good at estimating their level of control over keyboard input in this game. There is a strong and significant linear relationship between people's sense of control and how much control (in mutual information) they actually had. This observation may be generalizable to other inputs, other applications, and less immediate effects of the input. To confirm this, more research is required.

We suspect that the key aspect for users to be able to assess their level of control is that they are certain about what input they provide. With brain-computer interfaces, this is not yet the case. Even with relatively simple mental tasks such as focusing on a flickering target in the case of SSVEP, participants can be uncertain whether they are focusing in the right way or with the right intensity. In such situations, the psychological effects on the perception of control might be stronger, which would correspond to what we have informally observed in practice. However, with more practice this uncertainty, and any positive effects from the novelty of this type of interface, will diminish. As a result, the actual amount of control will become more dominant in the perception of control.

Some level of uncertainty will remain even with practice. This opens up a way to make uncertain input modalities more accepted, for example through the psychological phenomena described in the introduction. Such deliberate manipulation of perception is particularly applicable for games, as the goals and results of user actions are designed by the game designers instead of following from user goals. Besides, ambiguity can be used as a way to enhance user engagement [22].

## 5.2   Sufficient Control

We observed a critical level of control at 0.68 bits of mutual information, below which the amount of control affects the number of finished runs. In this application, this corresponds to an accuracy of 65% for 5 classes. This is slightly lower than the 70% indicated by BCI research groups (see Introduction).

Again, this result is based on this one application, with keyboard input. Besides, there are different ways of determining what amount of control is sufficient. This concerns just one specific aspect, which is whether people give up playing the game. Besides, the potential other factors that may affect the sense of control, may likely affect the necessary amount of control as well.

All this begs for more research in this area. Not only to increase the amount of information that can be provided through interfaces like BCIs, but also to investigate how this critical amount of control might be reduced. Again, games provide the perfect vehicle for this kind of research, as the goals and results are designed for a specific experience.

## References

1. Deng, L., Huang, X.: Challenges in adopting speech recognition. Communications of the ACM 47(1), 69–75 (2004)
2. Jacob, R., Karn, K.: Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. Mind 2(3), 4 (2003)
3. Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B., et al.: A review of classification algorithms for EEG-based brain–computer interfaces. Journal of Neural Engineering 4 (2007)
4. Plass-Oude Bos, D., Gürkök, H., Reuderink, B., Poel, M.: Improving BCI performance after classification. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction, pp. 587–594. ACM (2012)
5. van de Laar, B., Plass-Oude Bos, D., Reuderink, B., Poel, M., Nijholt, A.: How much control is enough? Influence of unreliable input on user experience. IEEE Transactions on Cybernetics 43(6), 1584–1592 (2013)
6. Plass-Oude Bos, D., Poel, M., Nijholt, A.: A study in user-centered design and evaluation of mental tasks for BCI. In: Lee, K.-T., Tsai, W.-H., Liao, H.-Y.M., Chen, T., Hsieh, J.-W., Tseng, C.-C. (eds.) MMM 2011 Part II. LNCS, vol. 6524, pp. 122–134. Springer, Heidelberg (2011)
7. Nijholt, A., Tan, D.: Playing with your brain: Brain-computer interfaces and games. In: Proceedings of the International Conference on Advances in Computer Entertainment Technology, pp. 305–306. ACM (2007)
8. Nijholt, A., Plass-Oude Bos, D., Reuderink, B.: Turning shortcomings into challenges: Brain–computer interfaces for games. Entertainment Computing 1(2), 85–94 (2009)

9. Graimann, B., Allison, B., Gräser, A.: New applications for non-invasive brain-computer interfaces and the need for engaging training environments. In: BRAIN-PLAY 2007 Brain-Computer Interfaces and Games Workshop at ACE (Advances in Computer Entertainment), pp. 25–28 (2007)

10. Allan, L.G., Jenkins, H.M.: The judgment of contingency and the nature of the response alternatives. Canadian Journal of Psychology 34(1), 1 (1980)

11. Thompson, S., Armstrong, W., Thomas, C.: Illusions of control, underestimations, and accuracy: A control heuristic explanation. Psychological Bulletin 123(2), 143 (1998)

12. Langer, E.: The illusion of control. Journal of Personality and Social Psychology 32(2), 311 (1975)

13. Tractinsky, N., Katz, A., Ikar, D.: What is beautiful is usable. Interacting With Computers 13(2), 127–145 (2000)

14. Norman, D.: Emotion & design: Attractive things work better. Interactions 9(4), 36–42 (2002)

15. Hakvoort, G., Gürkök, H., Plass-Oude Bos, D., Obbink, M., Poel, M.: Measuring immersion and affect in a brain-computer interface game. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part I. LNCS, vol. 6946, pp. 115–128. Springer, Heidelberg (2011)

16. van de Laar, B., Gürkök, H., Plass-Oude Bos, D., Nijboer, F., Nijholt, A.: Brain-computer interfaces and user experience evaluation. In: Allison, B.Z., Dunne, S., Leeb, R., Del, R., Millán, J., Nijholt, A. (eds.) Towards Practical Brain-Computer Interfaces, pp. 223–237. Springer (2012)

17. Quek, M., Boland, D., Williamson, J., Murray-Smith, R., Tavella, M., Perdikis, S., Schreuder, M., Tangermann, M.: Simulating the feel of brain-computer interfaces for design, development and social interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011, pp. 25–28. ACM (2011)

18. Cincotti, F., Kauhanen, L., Aloise, F., Palomäki, T., Caporusso, N., Jylänki, P., Mattia, D., Babiloni, F., Vanacker, G., Nuttin, M., et al.: Vibrotactile feedback for brain-computer interface operation. Computational Intelligence and Neuroscience 2007 (2007)

19. Ware, M., McCullagh, P., McRoberts, A., Lightbody, G., Nugent, C., McAllister, G., Mulvenna, M., Thomson, E., Martin, S.: Contrasting levels of accuracy in command interaction sequences for a domestic brain-computer interface using SSVEP. In: Biomedical Engineering Conference, pp. 150–153. IEEE (2010)

20. Carlson, T., Monnard, G., Millán, J.: Vision-based shared control for a BCI wheelchair. International Journal of Bioelectromagnetism 13(1), 20–21 (2011)

21. Quek, M., Höhne, J., Murray-Smith, R., Tangermann, M.: Designing future bcis: Beyond the bit rate. In: Allison, B.Z., Dunne, S., Leeb, R., Del R. Millán, J., Nijholt, A. (eds.) Towards Practical Brain-Computer Interfaces, pp. 173–196. Springer (2012)

22. Gaver, W.W., Beaver, J., Benford, S.: Ambiguity as a resource for design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 233–240. ACM (2003)

23. MacKay, D.J.: Information theory, inference and learning algorithms. Cambridge university press (2003)

24. Wewers, M., Lowe, N.: A critical review of visual analogue scales in the measurement of clinical phenomena. Research in Nursing & Health 13(4), 227–236 (2007)