A comparative evaluation of mouse, pen- and touch-input in computerized version of the Torrance tests of creative thinking

Stanislaw Zabramski, Dimitrios Gkouskos, Mats Lind
Department of Informatics and Media
Uppsala University
751 20 Uppsala, Sweden
firstname.lastname@im.uu.se

ABSTRACT
This paper presents the results of an experiment comparing mouse, pen- and touch-inputs in computerized figural creativity test. The results show no statistically significant differences in creativity scores obtained by participants using mentioned devices as computer input methods. The results indicate that creativity testing can be introduced to a variety of modern computer devices with insignificant influence of the three most popular input devices tested. That also shows that the differences between the computerized and paper-based TTCT noticed in a previous research might be the result of User Interface artifacts used.

Keywords
Mouse, stylus, pen, touch, evaluation, comparison, creativity, test, TTCT, figural, drawing

INTRODUCTION
Issues related to the use of computer input devices are one of the core topics in Human-Computer Interaction (HCI) research. Recent market trends show that touch- and pen-sensitive displays become more widespread and have been changing the way people interact with computers. This is also the case for creative or artistic endeavors. For example idea-sketching has an important role in the creative process in an early phase of design, therefore sketching behavior has been experimentally studied and idea-sketching is supported by many computerized design tools [9]. On the other hand research on computer input methods has traditionally focused on the accuracy in navigational pointing or selecting tasks. Tracing tasks that require a creative approach demand much more freedom and therefore are more complex and harder to predict by traditional spatio-temporarily restrictive HCI models [1; 5].

In previous studies on unconstrained shape tracing [10] it has been found that university students using mouse, pen- and touch-input were comparably accurate in shape-pattern replication task while being significantly different in the time needed to perform that task. The results show that students using touch-input outperformed pen-input users by a factor of 1.54 and mouse users by a factor of 2.3.

DESIRE'11, 19-21 October, 2011, Eindhoven, The Netherlands

Therefore we decided to check if that time-accuracy tradeoff would affect a person using mentioned input methods for computer-supported creative drawing.

TTCT-Figural
We have found that solving a figural part of the Torrance Tests of Creative Thinking (TTCT) is a task which provides the participant with the drawing situation and additionally offers a standardized measure of the participant's creativity. The TTCT has been developed in 1960s and its validity has been assessed by longitudinal studies carried out internationally. The test includes two subtests. The TTCT-Verbal consists of forms A and B that include the following subtests: Asking Questions, Guessing Causes, Guessing Consequences, Product Improvement, Unusual Uses, and Just Suppose. The TTCT-Figural also has two forms with three subtests: Compose a drawing (one closed shape as a stimulus), Finish a drawing (10 open shapes as stimuli), and Compose a different drawings (based on sets of parallel lines or circles as stimuli) [3].

The TTCT's scoring procedures can be taught easily [8] and produce valid and reliable assessments of four principal cognitive processes of creativity: Fluency - number of relevant responses, Originality - novelty of unusual but relevant responses, Abstractness of Titles - a verbal measure, Elaboration - the number of details used to extend a response, and Resistance to Premature Closure - a person's ability to stay open and tolerate gestalt ambiguity.

In one of the previous attempts of creating a computer-based version of the figural TTCT, it has been found not to be the equivalent to the paper-pencil version [4]. That version of the TTCT was operated with a mouse and had elaborated User Interface (UI) artifacts that were constantly displayed on the screen. Knowing that UI design has a profound impact on the user's experience and performance, we have decided to take a different approach and minimize that effect by creating a program with UI that would resemble the functionality of a pen and a piece of paper. It has been later confirmed that the original, pen-based, paper version of the figural part of the TTCT and its digital, screen-based version, deliver similar results when the User

Interface (UI) of the computer version offers similar user experience as the paper version [11].

We have decided to follow that approach towards UI and check if there is any influence of other popular input methods on the computer user's creative process in a drawing task. We have designed and performed an experiment using a computerized version of the TTCT and comparing mouse and touch-input with a pen as a reference.

METHODS AND PROCEDURES

It would be not possible to obtain unbiased results from a series of full consecutive creativity tests taken by each user with each input method. Therefore, in order to compare the results of mouse, pen- and touch-input in a creative drawing task we performed the test of a factorial design. Every participant used one of the three input methods (within subject, assigned in randomized order for counterbalancing) to perform one of three activities of the figural part of the TTCT (Form B) in original order.

Participants

24 participants aged between 20 and 38 (median of 28 years) have been recruited among the students and staff of Uppsala University by convenience sampling and voluntarily participated in the study. They were active computer users (6.9 hours/day) using mostly the mouse as the input device (5.5 hours/day) with the touch-input as secondary one (1.6 hours/day), followed by the pen-input (0.42 hour/day). None of them had any prior knowledge of the TTCT method.

Hardware

An HP Touchsmart TM2-1090eo Tablet PC with a 12.1 inch diagonal LED display and a resolution of 1280*800 pixels, equipped with stylus and finger input, as well as a Logitech basic optical mouse were used. The HP TM2 was used in "tablet mode" with the stylus and finger input, lying flat on the desk or in "laptop mode" while used with the mouse.

Software

The previous attempt of creating the computerized TTCT [4] was operated with a mouse and had elaborated User Interface (UI) artifacts that were constantly displayed on the screen. We have decided to minimize that effect by creating a program with UI that would resemble the functionality of a pen and a piece of paper. Taking into consideration the unique characteristics of the three different input methods we wanted to compare, any advanced functionality could not be translated to the simple and uniform user experience of pen and paper. Therefore, we have created a program that allowed the user to leave a 5 pixel thick solid black trace of drawing over the black contour shapes generated with 5 pixel thick solid black lines and displayed on a white background.

The test moderator was switching manually between three activities of the TTCT figural B whenever instructed by a participant.

We kept the standard MS Windows 7 settings unmodified for all three input methods and corresponding cursors. The pen and touch cursors were visible while interacting and the mouse cursor was always visible.

All the user's interactions, system events, computer's screen view and a picture from external video camera were recorded with TechSmith Morae v.3.2.

Test scenario

As pre-test arrangements participants had to sign consent forms and take part in a short introductory session for the mouse, stylus and touch input in MS Paint. Then, they were presented with a greyed-out random shape and instructed to trace over the shape in one stroke using every input method.

Then participants filled in pre-test questionnaires and performed all three activities of the TTCT figural form B according to the TTCT Directions Manual [7] where:

I. Picture Construction Activity – required the participant to draw a picture in which a closed shape provided should be an integral part.

II. Incomplete Figures Activity - where the participant was provided with 10 different open shapes (one screen with first 6 shapes and then 4 shapes on another screen) and was required to draw as many pictures as possible with each shape as an integral part.

III. Repeated Figures Activity - provides 30 sets of paired parallel lines or 36 circles distributed over 3 screens for a participant to make and draw multiple associations to a single stimulus.

In each activity participants had to write the titles of their drawings on a separate sheet of paper using a regular pencil. These titles were later digitized and combined with the drawings for grading purposes.

After finishing the last activity participants were asked to fill in a post-test questionnaire regarding their preferences and opinions about the input devices they were using.

We have collected screen snapshots with the participants' drawings after 10 minutes - that is on the end of every activity, and also intermediary snapshots whenever instructed by the participant to switch to another page/screen of given activity if more pages were available.

Additionally, we have collected snapshots of participants' drawings at the beginning of the 5^{th} minute in tasks performed using touch and 8^{th} minute in case of tasks performed using stylus. These time-points approximately corresponded to the time-wise performance differences between input methods in shape replication tasks.

The snapshots have been analyzed and an individual Creativity Index (CI) score of every participant has been assessed. CI incorporates standardized evaluations of four principal cognitive processes of creativity manifested in the participant's responses to the various stimuli of the three activities of the TTCT. However, these standardized scores

are calculated between activities and are based on multiple raw scores describing the user's performance in each activity. Even though it is not directly supported by the TTCT method, we decided to use these intermediary raw scores and group them to describe user's performance in every activity of the TTCT separately. That allowed us to have a look at the participants' scores from the perspective of each input method. To achieve that, the participants were randomly assigned to "triads" based on the order of input device used (Latin-square design). The raw scores of the three participants constituting each triad have been combined what created 8 "virtual" users performing every activity of the TTCT with every input method.

RESULTS

The scoring of the participant's responses was performed by two untrained raters using the TTCT scoring manual [8] as a reference. These results have been later compared to estimate inter-rater reliability and Pearson's correlation coefficient of 0,785 has been achieved indicating a strong positive correlation between raters [6].

The variance of CI in the triads has reduced and the standard deviation of 7.2 from initial 12.3 for the individual CIs. That suggests that these triads present a more even level of creativity and became comparable more easily.

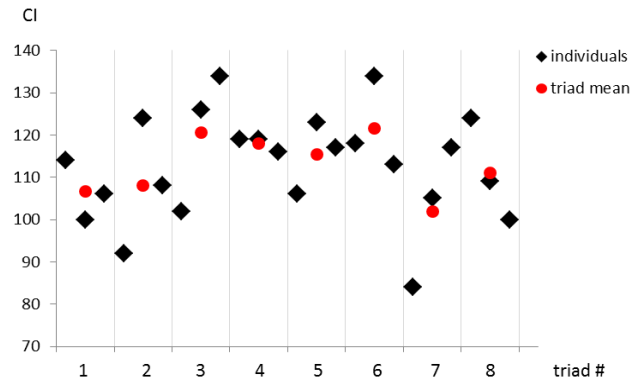

Figure 1. Scatter plot of participants' individual CI *(the black diamonds)* and mean CI values of each triad *(the red circles)*.

The raw scores within each triad have been grouped by input device and then analyzed. The large differences in variance have been observed between the different groups therefore, non-parametric tests were used.

The first tested hypothesis was that no difference exists between the input methods when the same 10 minutes time frame is used. A Friedman non-parametric ANOVA for dependent samples were used to do this. The result showed that the hypothesis cannot be rejected (Chi Sqr. (N=8, df=2) = 1.75; p=0.4169). The same procedure was then used to test the hypothesis that there were no differences between the input methods when a modified time-frame was used (8 minutes for pen- input and 5 minutes for touch-input). This hypothesis was rejected (Chi Sqr. (N=8, df=2) = 8.58; p=0.0137). A post-hoc test was then performed by pair wise testing these conditions using the Wilcoxon matched pair test. It revealed that in 10 minutes mouse input produced significantly higher creativity scores both compared to 8 minutes of pen usage (N=7, T=1, Z=2.1974; p= 0.028) as well as compared to 5 minutes of touch input usage (N=8, T=1.5, Z=2.3105; p=0.021). However, no significant difference was found between 8 minutes of pen usage and 5 minutes of touch input (N=8, T=5.0, Z=1.8204; p= 0.069).

Our participants on average were skilled mouse users (5.5 hours per day) while having rather brief contact with touch inputs (1.5 hour per day) and accidental contact with stylus as an input device (0.5 hour per day).

Post-test questionnaires were based on The Creativity Support Index survey [2] and showed participants' positive or negative responses to evaluation statements in the 10-point Likert scale, where 1 = "agree" and 10 = "disagree":

Q2. "I was very absorbed/engaged in this activity – I enjoyed it and would do it again" got median response = 2.

Q3. "What I was able to produce was worth the effort required to produce it" with median responses: 7 for "Mouse", 4 for "Pen" and 5 for "Touch" sub option.

Q4. "While I was doing the activity, the input tool "disappeared" and I was able to concentrate on the activity" with median responses: 7 for "Mouse", 2 for "Pen" and 5 for the "Touch" sub option.

DISCUSSION

In the previous attempts of creating a computerized version of the TTCT a mouse was offered to the participants to respond to the stimuli [4]. That was reported to be more difficult and time-consuming compared with the paper-and-pencil version of the TTCT. It also involved elaborated UIs that also could be blamed for introducing an additional cognitive load significantly influencing participant's scores. We avoided this problem by not introducing any artificial elements of UI that would differentiate the digital version from paper even more. On the other hand, our implementation of the TTCT into computer software was a result of many compromises therefore it was not a perfect solution. E.g. we have decided not to implement an eraser's functionality since it would be hard to implement it e.g. for touch-input, without creating its visual metaphor as an UI artifact which we wanted to avoid. Nevertheless, only 6 of 24 participants asked about some sort of "erase" or "undo" function. The design of this experiment forced us to use TTCT's raw scores but we were interested in potential input method induced differences not in the absolute values of users' Creativity Index - which was used only for reference. Activity-based results could not be obtained otherwise because the TTCT raw scores are summarized between activities later during the scoring process, and blend into values characterizing four principal cognitive processes measured by the TTCT. The results showing no significant influence of the digital input methods used in these creative tasks might be potentially surprising in the light of previous research on input methods. Obviously, the level of

participant's dexterity with use of any given input method is increasing through practice and for a skilled user the input device is no longer an obstacle and becomes a "transparent" extension of the body. Therefore, we may assume that mouse users' high skills have compensated the fact that the pen- and touch-input offer more direct (in a sense of spatial separation between human input and computer output) and less cognitively demanding interaction. However, the finding that there was no difference between 8 minutes of pen usage and 5 minutes of touch input usage suggests that the participants were using strategies that resulted in valuable elements of their work like titles or elaborated details being added on the end of each activity. Also, the fact that the input methods used produced comparable results at these intermediary stages and at the end of the activity indicates a similar and uniform influence they have on the results of the creative idea generating process – esp. if the esthetical appeal, time consumption or accuracy of drawing are not graded.

Figure 2. Exemplary responses to the Repeated Figures activity generated with the use of the mouse *(the left column)*, stylus *(the middle column)*, and finger *(the right column)*.

Qualitative post-test data shows that the TTCT test was perceived as an enjoyable experience. The pen seemed to be a tool that helped to reach a state of flow where its presence was no longer noticed, enabling participants to produce the outcome worth the effort – even though the participants do not use the pen as the computer input device. That reflects the perception of the pen's intuitiveness and ease of use in drawing situations. The touch input got medium scores and the mouse was perceived the worst option of all reflecting low level of user's satisfaction.

CONCLUSION

Creativity is the basis for inventive problem solving. We thought it was important to identify the computer input devices that affect user's creativity most. It seems however, that the user's creativity in the context of use of the computer based creativity testing application or brainstorming facilitating tool may be more affected by the design of the UI than by the input method used. Access to digitized results of the user's creative work is an important feature of that kind of software therefore, we think that as the next step it would be necessary to create different mouse-based versions of UIs of e.g. the digitized TTCT and to compare them to the previous attempts [4] to see how and to what extent the problem of UI's influence could be minimized.

ACKNOWLEDGMENTS

This study was partially financed by Anna Maria Lundin Foundation and The Swedish IT-User Centre (NITA).

REFERENCES

1. Accot, J. and Zhai, S. Performance evaluation of input devices in trajectory-based tasks: an application of the steering law, in *Proceedings of the the SIGCHI conference on Human factors in computing systems: the CHI is the limit* (Pittsburgh, Pennsylvania, United States, 1999), ACM,
2. Carroll, E.A. and Latulipe, C. The creativity support index, in *Proceedings of the the 27th international conference extended abstracts on Human factors in computing systems* (Boston MA USA, 2009), ACM, 4009-4014.
3. Cramond, B., Matthews-Morgan, J., Bandalos, D. and Zuo, L. A Report on the 40-Year Follow-Up of the Torrance Tests of Creative Thinking: Alive and Well in the New Millennium. *Gifted Child Quarterly* (49), 283-291.
4. Kwon, M.C. An exploratory study of a computerized creativity test: comparing paper-pencil and computer-based versions of the torrance tests of creative thinking, in Texas A&M University 1996, 156.
5. Mackenzie, I.S., Sellen, A. and Buxton, W.a.S. A comparison of input devices in element pointing and dragging tasks, in *Proceedings of the the SIGCHI conference on Human factors in computing systems: Reaching through technology* (New Orleans LA USA, 1991), ACM,
6. Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. *Numerical Recipes in C*. Cambridge University Press, 1996.
7. Torrance, E.P. *Torrance Tests of Creative Thinking: Directions Manual: Figural A and B*. Scholastic Testing Service, 1990.
8. Torrance, E.P., Ball, O.E. and Safter, H.T. *Torrance Tests of Creative Thinking: Streamlined Scoring Guide: Figural A and B*. Scholastic Testing Service, 1992.
9. Verstijnen, I.M., Van Leeuwen, C., Goldschmidt, G., Hamel, R. and Hennessey, J.M. Sketching and creative discovery. *Design Studies* (19), 519-546.
10. Zabramski, S., Gkouskos, D. and Lind, M. A comparative evaluation of mouse, stylus and finger input in shape tracing, in *Proceedings of the 1st European Workshop on HCI Design and Evaluation* (Limassol, Cyprus, 2011). IRIT Press, Toulouse, France, 102.
11. Zabramski, S. and Neelakannan, S. Paper equals screen; A comparison of a pen-based figural creativity test in computerized and paper form, in *Proceedings of the DESIRE'11* (Eindhoven, the Netherlands, 2011).