A Comparative Study of Three Sudoku Input Methods for Touch Displays

Aslak Burheim Sommervold[1], Benjamin Nils Øvergaard[1], Eskil Nysether[1], Mohamed Yusuf Nur[1] and Frode Eika Sandnes[1,2][0000-0001-7781-748X]

[1] Oslo Metropolitan University, Oslo, 0130 Oslo, Norway
[2] Kristiania University College, Oslo, Norway
aslak.sommervold@gmail.com,
ben.overgaard@gmail.com,s330457@oslomet.no,s333725@oslomet.no,frodes@oslomet.no

Abstract. Sudoku is a popular recreational game which is claimed to have positive health effects. It can be played using paper or electronically using computers. Only very few studies have explored Sudoku interaction methods. We therefore designed a controlled within-groups experiment involving $N = 18$ participants to empirically compare three Sudoku interaction methods implemented in a popular Sudoku smartphone app. Our results show that the participants entered digits faster when they selected the location first, followed by selecting the input digit, compared to selecting the digit first followed by selecting the cell location. Participants also preferred selecting cell first over selecting input digit first. No effects of error rates were found.

Keywords: Sudoku, smartphone game design, touch interaction, cell first, digit first, game interaction, preference.

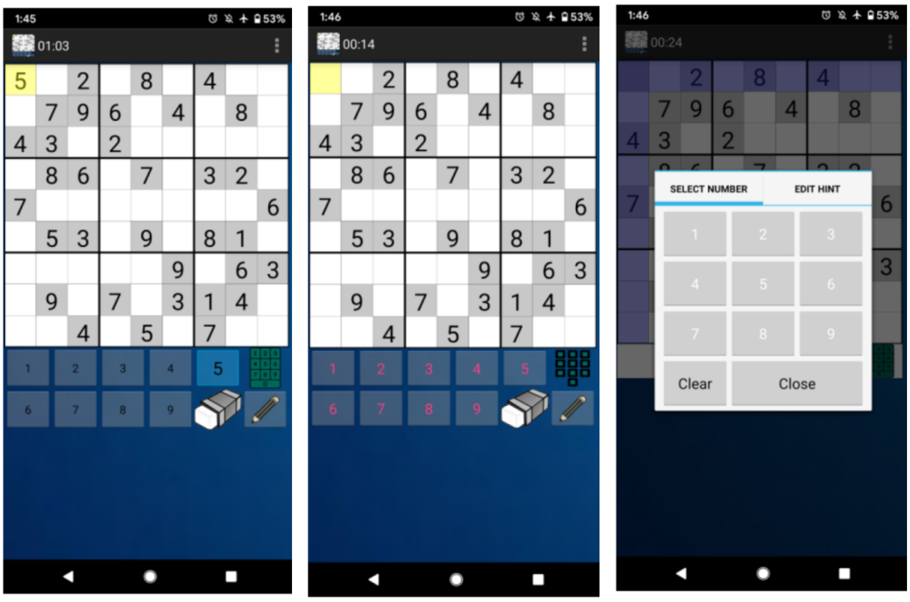
1 Introduction

Sudoku is a popular game that is believed to reduce cognitive aging and help train working memory [1]. It is a game of numbers that involves placing digits on a grid. The traditional Sudoku grid comprises 9 × 9 cells divided into nine squares of 3 × 3 cells. The game involves inserting digits from 1 to 9 such that all vertical rows and horizontal columns contain each of the 9 digits exactly once. At the same time the player must adhere to a set of digits already inserted into the board. The difficulty of the game is controlled by the way the fixed digits are inserted into the Sudoku grid. Sudoku can also be played using letters, words or symbols instead of digits.

Traditionally, Sudoku puzzles were printed in newspapers or in special Sudoku puzzle books. Several electronic computer-based Sudoku games have also been developed. To the best of our knowledge, there are only a handful of studies addressing the interaction aspects of Sudoku and little is therefore known about Sudoku game interactions. This study therefore set out collect empirical data about three interaction methods implemented in a popular Sudoku app.

This is a post-peer-review, pre-copyedit version of a conference proceeding that has been published in HCII 2020: Universal Access in Human-Computer Interaction. Design Approaches and Supporting Technologies. 14th International Conference, 22nd HCI International Conference, Proceedings, Part I, which is part of the Lecture Notes in Computer Science book series (volume 12188). The final authenticated version is available online at: http://doi.org/10.1007/978-3-030-49282-3_15

We have coined the three input methods *digit first*, *cell first* and *popup*. The *digit first* input method involves first selecting the digit at the bottom of the screen organized in two rows (see Fig. 1 (a)) and then select the cell in which the digit is to be inserted. Once a digit is selected, this digit can be inserted in multiple cells without having to re-select the digit. Hence, the digit first method holds potential for reducing the users' workload. With the *cell first* method the user first selects the cell on the Sudoku board where the digit is to be placed, followed by selecting the input digit (see Fig. 1 (b). To insert another digit the cell-digit selection process needs to be repeated. *Popup* is similar to the *cell first* method in that the user first selects the cell where the digit is to be input. Next, a digit-selection popup is presented to the user (see Fig. 1 (c)) in the form of a numeric keypad. Again, the entire process is repeated to insert additional digits.



(a) Digit first (b) Cell first (c) Popup

Fig. 1. Sudoku input methods.

This study thus set out to explore if there are any differences in the mean time to use these input methods, the errors involved and users' preferences for the three methods. This paper is organized as follows. The next section reviews related work, followed by a description of the method used and the results obtained. Next, the results are discussed. The conclusions section closes the paper.

2 Related work

Sudoku has received much attention in the research literature, see for instance [2, 3, 4]. Within computer science most of the attention has evolved around algorithms for the automatic solution of Sudoku puzzles [5, 6, 7], for example using metaheuristics [8]. A Sudoku inspired encoding has also been used for information hiding [9]. The potential of Sudoku as a pedagogical resource for learning computer science has also been discussed [10].

Comparatively, little has been written about Sudoku interaction. One exception is Norte and Lobo's [11] Sudoku game aimed at users with reduced motor function where the game is controlled via switches and scanning or via voice input. Sudoku has also been studied for visually impaired players [12], and very challenging problem as it is difficult to effectively communicate two-dimensional structures such as tables with non-visual means [13]. Echtler et al. [14] discussed a Sudoku interface for tabletop displays allowing users to move items around or access the game using a mobile device. Kondraju [15] proposed a possible extension of Sudoku into three dimensions. Bernsen and Dybkjær [16] addressed usability testing with Sudoku as a case. A doctoral dissertation [17] (written in Slovenian) was devoted to the implementation of an Android Sudoku app.

Traditional Sudoku uses digits and the interaction can in some ways be considered a specialized form of digit input. The problem of digit input has been studied extensively as it is a commonly performed task in many domains. Issues that have been studied includes speech-based digit input [18], written digit recognition [19] and keyboard layouts for digit input [20, 21]. It has been demonstrated that numeric keypads are more efficient than the number keys found on small laptop computers [21]. In Norway digit copying and input is heavily used when performing bank transactions, both for copying temporary passcodes [22] and long customer identification numbers [23]. The copying and input of numbers also appear in other domains such as prescription parameters in medical equipment [24, 25, 26, 27, 28].

One of the input methods studied herein includes popups. Popup messages versus in text messages have been measured systematically [29]. Specialized methods for setting quantities has also been studied such as setting the time [30] and color values [31, 32].

3 Method

3.1 Experimental design

A controlled within-groups experiment was conducted with input method as an independent variable with three levels, namely *cell first*, *popup* and *digit first*. Previous experience playing Sudoku was also used as a secondary between-groups independent variable. Three dependent variables were measured, namely the time to input digits, error rate the participants' subjective input method preferences.

3.2 Participants

A total of 18 participants was recruited for the experiment. There nearly a balance of female and male participants (with slightly more males) and mostly in their twenties, and a few in their thirties. The participants were all computer science students at the authors' university. Of the participants, 13 had played Sudoku before, while 5 have never played Sudoku. Of the 13 participants who had played Sudoku 4 had also used the app before.

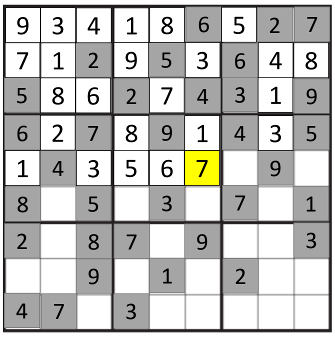


Fig. 2. The cell with its content highlighted in yellow as used in the Sudoku copying task. The empty cells on the board were traversed left-to-right, top-to-bottom.

3.3 Equipment

The Classic Sudoku Pro (No Ads) Android app was used for the experiments. The experiments were run on a Google pixel 2 smartphone. The numbers to be input were presented using a Microsoft PowerPoint presentation on a laptop computer placed on a desk in front of the participants. This PowerPoint presentation contained 146 pages shown in sequence. For each page the cell with the digit to be input was highlighted (see Fig. 2). Hence, the participants simply had to copy what was shown on the PowerPoint. The cells were filled in from left-to-right, top-to-bottom. Google forms was used to collect the users' preferences for the various input methods.

3.4 Procedure

Each participant was tested individually in a quiet room. They were seated in front of a laptop computer and was asked to hold the smartphone with the Sudoku app. The participants were asked to input 47 digits displayed in the PowerPoint presentation into the app. Hence, the participants did not have to solve the Sudoku puzzle, but rather focus on copying the digits into their respective positions. Each participant completed the Sudoku board with each of the three input methods. The order of the input methods was varied to minimize learning effects and biases. There are six possible combinations with the three input methods and the 18 participants were recruited so that each combination was executed by three participants. The task completion times for the entire Sudoku boards were measured using a stopwatch. This total board completion time was used to calculate the time per digit for each participant. Errors were observed manually and counted. After completing the session with the three interfaces the participants were asked to indicate their subjective preference for the three interfaces using a 5-point Likert scale using Google forms, i.e., how easy they found each interface to use, respectively. The questionnaire also asked if the participants had played Sudoku before, and if they had used this app or similar apps before.

Participation was voluntary and anonymous. As the experiment was conducted in one session no identifying mechanisms had to be employed. The General Data Protection Regulations (GDPR) therefore did not apply for this study.

3.5 Analysis

The results were analyzed using JASP version 0.10.0.0. The task completion times were analyzed using a repeated measures anova with Bonferonni post-hoc tests, while the error rates and preference data were analyzed using a Friedman test with Connovers post hoc testing since the error rates and Likert data did not satisfy the assumptions of the parametric testing procedures. The non-parametric Spearman correlations and Mann Whitney tests were also used for the same reasons.

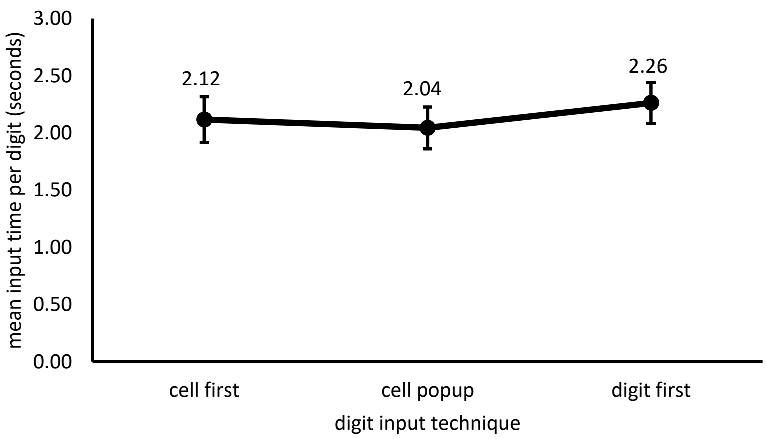


Fig. 3. Mean time to input a digit with the three interfaces. Error bars show 95% confidence intervals.

4 Results

4.1 Digit input time

Fig. 3 shows the results of the digit input time measurements. A repeated measures anova shows that the digit input times with the three interfaces were significantly different ($F(1, 34) = 11.17$, $p < .001$, $\eta^2 = 0.397$). Bonferroni post-hoc tests reveal that the mean digit input time with the *digit first* interface (M = 2.26 seconds, 95% CI [2.08, 2.44]) was significantly different to both the *cell first* interface (M = 2.12 seconds, 95% CI [1.92, 2.32], p =.027, Cohen's d = -0.694) and the *popup* interface (M = 2.04 seconds, 95% CI [1.86, 2.23], p =.002, Cohen's d = -0.961), while the *cell first* and *popup* interface were not significantly different (p =. 188). In other words, the *digit first* input method was the slowest, while the *popup* digit input method was the fastest, although not significantly faster than the *cell first* input method.

Pearson correlations showed that participants who were faster with one type of input method were also faster with the other input methods. That is, digit input time with *digit first* correlated strongly and positively with digit input using *cell first* ($r(18) = 0.876$, $p < .001$), *digit first* correlated strongly and positively with *popup* ($r(18) = 0.833$, $p < .001$) and *cell first* correlated strongly and positively with *popup* ($r(18) = 0.936$, $p < .001$).

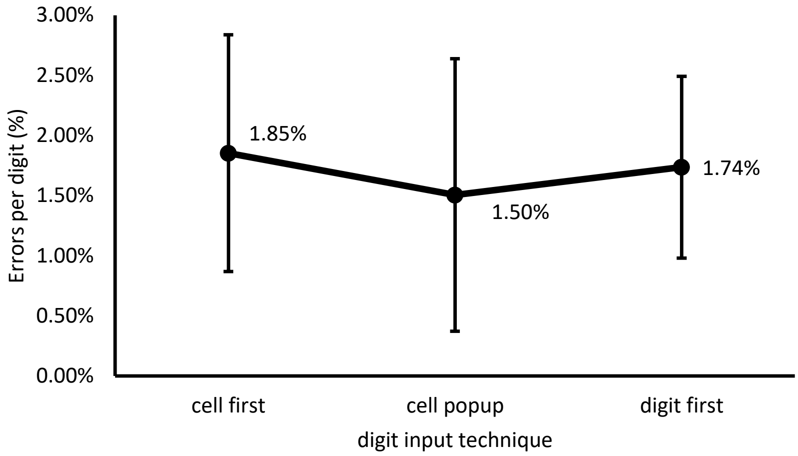


Fig. 4. Mean percentage of errors per digit. Error bars show 95% confidence intervals.

4.2 Errors

Fig. 4 shows the error rates in percentage for the three digit input methods. Overall, the error rates with the digit input methods were below 2%. The *popup* interface appears to have been associated with the fewest mean number of errors, although this input method also had the largest spread. The confidence intervals for the three methods overlap greatly and a Friedman test confirms that there was no significant difference in error rate between the three digit input methods ($\chi^2(2) = 0.905$, $p = .636$).

There was also a significant positive medium to strong correlation between the digit input times and error rates for the three methods, namely *digit first* ($r(18) = 0.729$, $p < .001$), *cell first* ($r(18) = 0.645$, $p = .004$) and *popup* ($r(18) = 0.578$, $p = .010$). Clearly, participants who perform the task faster also makes more mistakes.

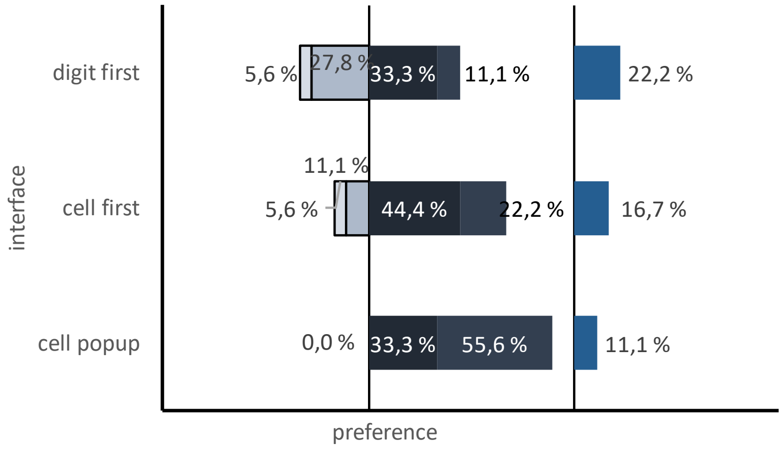


Fig. 5. Diverging stacked bar graph showing the distribution of preferences in percentages (left: negative, middle: positive and right: neutral.

4.3 Preferences

Fig. 5 shows the participants' subjective preferences for the digit input methods. Clearly, the *popup* method was associated with the most positive responses as there were no negative responses and it exhibited the lowest quantity of neutral responses. *Popup* is followed by the *cell first* digit input method. *Digit first* was associated with the lowest preference with a balanced set of negative and positive responses as well as being the input method with the most neutral responses. A Friedman test reveals that there is a significant difference between the preferences for the three methods ($\chi^2(2)$ = 8.716, p = .013, Kendall's W = 0.137). Connover's post hoc tests show that the low preference for *digit first* was significantly different to *popup* (p =.007) which was the most preferred. However, there were no significant differences between *popup* and *cell first* (p = .124) and *cell first* and *digit first* (p = .186).

Spearman correlations also revealed a significant positive medium correlation between digit input times using the digit first input method and preference for the digit first input method ($r_s(18)$ = 0.566, p = .014).

4.4 Prior Sudoku experience

Of the participants, 13 had played Sudoku before while 5 had never played the game. In terms of digit input time a mixed repeated measures anova did not reveal any between-group effects ($F(1, 16)$ = 1.437, p = .248), however, an interaction between digit input times and past experience with Sudoku was found ($F(1, 32)$ = 8.601, p < .001).

To explore other between-groups effect a series of non-parametric Mann Whitney tests were performed. Three significant differences were observed. First, the number of errors were significantly lower with the *popup* method for participants who had played

Sudoku before compared to those who had not played the game before (W = 12.00, p = .024). Next, participants who had prior experience with Sudoku reported a significantly higher preference (W = 52.5, p = .031) for the *popup* method (M = 4.69, SD = 0.48) than the group without prior experience (M = 3.8, SD = 0.84). Finally, the participants with prior Sudoku experience exhibited a significantly lower preference (W = 8.5, p = .016) for the *digit first* input method (M = 2.77, SD = 1.09) compared to those without prior experience (M = 4.20, SD = 0.45) who showed a high preference for the *digit first* method.

No effects were found connected to participants prior experience with the Sudoku app that was used in the experiments.

5 Discussion

The results show that it takes around two seconds to input a digit with the three Sudoku input methods. *Popup* resulted in the shortest digit input times and the *digit first* method resulted in the longest digit input times, although the difference between the *popup* and *cell first* methods was practical and not statistical. Moreover, the results showed that participants who are faster with one method is also faster with the other methods, suggesting that there was an underlying effect of individual skill.

The error rate results did not show any differences between the methods. We are therefore unable to claim any effect of input method on error. Instead, the results showed that participants who entered digits faster also made more mistakes. This observation is consistent with what one would expect.

The input time observations were consistent with the preference observations, namely that the *popup* method is the fastest, and also the most preferred, while the *digit first* method was the slowest and least preferred method. One may speculate that a faster interface is more likely to be preferred over a slower interface, although this may not always be true such as has been demonstrated in the text entry literature where time has demonstrated that faster optimized keyboard layouts cannot compete with the ubiquitous but slower Qwerty keyboard [20].

One explanation for why *digit first* yielded the slowest digit entry times and lowest preference scores is that a participant primarily focus on the Sudoku board and its location first and the digit to be inserted second. Therefore, asking a participant to first select the digit before selecting the location on the board may violate the mental models of users. Hence, the intended time saving made possible by inserting the same digit into multiple cells does not justify the added cognitive load of violating the cognitive model of the users.

Clearly, the *popup* method was only practically faster and more preferred than the *cell first* method and not significantly faster. A possible explanation for this small, but insignificant difference could be the following. First, the user may have been more familiar with the numeric popup dialog as it had the visual appearance of a traditional numeric keypad, while the two-row numeric layout used with the two other methods was not standardized and therefore not as recognizable. Another explanation may be that the popup immediately drew the users' attention towards the digit input step after selecting a cell. While with the *cell first* method the user must explicitly initiate a selection of a digit from the bottom of the screen.

It is indeed interesting that participants who performed worse with the *digit first* input method also preferred this input method. Perhaps one explanation is that for participants who were generally faster would prefer one of the other two input methods while the participants who were generally slower preferred the *digit first* method.

These observations are also consistent with the between-groups results, namely that the participants with prior experience with Sudoku exhibited a higher preference for the *popup* method compared to those that had not played Sudoku before, while those who had not played Sudoku before preferred the *digit first* method to a much higher degree than those who had Sudoku experience. A noticeable between groups effect was also observed for the *popup* method further confirming experienced participants' confidence with this method over the inexperienced participants. One reason for these differences could be that the participants with Sudoku experience approached the task in a Sudoku manner, although they were not asked to solve the Sudoku puzzles, while the participants without Sudoku experience approached the task as a simple digit entry task where digits displayed had to be mechanically placed on the board without placing any particular significance with regards to the board locations.

5.1 Limitations and future work

This experiment essentially comprised a digit copying task. The nature of the task may have affected the results as the performance of the participants may have been different if they were exposed to an actual game context rather than digit copying context.

The cohort was limited to relatively young computer science students. It is therefore unlikely that the results generalize as computer science students are likely to be more enthusiastic about using computers and more computer literate than a typical computer user. Future work should therefore also consider broader sampling.

It would also have been relevant to explore effects over time through a longitudinal study. One may expect that the practical differences between the popup numeric keypad and the two row digit layout in the cell first method would disappear with practice as users would learn to locate the desired digits with equal speed, while the mental mismatch of asking for the digit before the location may be less likely to disappear.

It would have been interesting to explore Sudoku game interaction using digit handwriting as it would be more directly related to the traditional paper-based Sudoku games. Touch displays also affords handwriting. With direct handwriting on a Sudoku board the selection of location and insertion of the digit would be done in one integrated step. Although more physical effort may be needed to articulate the digit gestures compared to the two display presses needed to select location and content, such digit gestures are an established part of most people's fundamental skill set utilizing their motor memory.

6 Conclusion

Sudoku game interaction was studied. Three input methods for placing digits on Sudoku boards were explored. The results show that user performed the task faster and also prefered to first select the location of the digits followed by selecting the digit. On the contrary, users without Sudoku experience preferred to enter the digits first and location second. Small non-significant, but practical differences were found in favor of numeric keypad popups over the two-line row of digits at the bottom of the screen. One possible implication of the results is that this type of games should be designed with the principle of going from broad/global to narrow/local, i.e., from location to content.

References

1. Grabbe, J. W.: Sudoku and working memory performance for older adults. Activities, Adaptation & Aging **35**(3), 241-254 (2011). doi: 10.1080/01924788.2011.596748
2. Delahaye, J. P.: The science behind Sudoku. Scientific American, 294(6), 80-87 (2006).
3. Felgenhauer, B., Jarvis, F.: Mathematics of sudoku I. Mathematical Spectrum **39**(1), 15-22 (2006).
4. Herzberg, A. M., Murty, M. R.: Sudoku squares and chromatic polynomials. Notices of the AMS **54**(6), 708-717 (2007).
5. Geem, Z. W.: Harmony search algorithm for solving sudoku. In: International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp. 371-378. Springer, Berlin, Heidelberg (2007). doi: 10.1007/978-3-540-74819-9_46
6. Simonis, H.: Sudoku as a constraint problem. In CP Workshop on modeling and reformulating Constraint Satisfaction Problems, Vol. 12, pp. 13-27 (2005).
7. Felgenhauer, B., Jarvis, F.: Enumerating possible Sudoku grids. Preprint available at http://www.afjarvis.staff.Shef.ac.uk/sudoku/sudoku.pdf. (2005).
8. Lewis, R.: Metaheuristics can solve sudoku puzzles. Journal of heuristics **13**(4), 387-401 (2007). doi: 10.1007/s10732-007-9012-8
9. Chang, C. C., Chou, Y. C., Kieu, T. D.: An information hiding scheme using Sudoku. In: 2008 3rd international conference on innovative computing information and control, pp. 17-17. IEEE (2008). doi: 10.1109/ICICIC.2008.149
10. Kurmas, Z., Dulimarta, H., Tao, Y.: Sudoku: nifty tools. Journal of Computing Sciences in Colleges **23**(1), 109-110 (2007).
11. Norte, S., Lobo, F. G.: Sudoku access: a sudoku game for people with motor disabilities. In: Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility, pp. 161-168. ACM (2008). doi: 10.1145/1414471.1414502
12. Libuša, A.: Sudoku for Visually Impaired, Doctoral dissertation, Masarykova univerzita, Fakulta informatiky (2008).
13. Sandnes, F. E.: Universell utforming av IKT-systemer, Oslo: Universitetsforlaget, 2nd edition (2018).
14. Echtler, F., Nestler, S., Dippon, A., Klinker, G.: Supporting casual interactions between board games on public tabletop displays and mobile devices. Personal and Ubiquitous Computing **13**(8), 609-617 (2009). doi: 10.1007/s00779-009-0246-3
15. Kondraju, V. K. M. K.: An Exploration of HCI Design Features and Usability Techniques in Gaming. IOSR Journal of Computer Engineering **15**(3), 53-57 (2003).
16. Bernsen, N. O., Dybkjær, L.: Intermezzo 5: Sudoku Usability Evaluation. In Multimodal Usability, pp. 387-412. Springer, London (2010). doi: 10.1007/978-1-84882-553-6_17

17. Bartol, M.: Aplikacija uganke Sudoku za Android, Doctoral dissertation, Univerza v Ljubljani (2012).
18. Ainsworth, W. A.: Optimization of string length for spoken digit input with error correction. International Journal of Man-Machine Studies **28**(6), 573-581 (1988). doi: 10.1016/S0020-7373(88)80061-0
19. LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., Jackel, L. D.: Handwritten digit recognition with a back-propagation network. In: Advances in neural information processing systems, pp. 396-404 (1990).
20. Sandnes, F. E.: Effects of common keyboard layouts on physical effort: Implications for kiosks and Internet banking. In: The proceedings of Unitech2010: International Conference on Universal Technologies, pp. 91-100. (2010).
21. Tsang, S. N., Chan, A. H. S., Chen, K.: A study on touch screen numeric keypads: Effects of key size and key layout. In: International MultiConference of Engineers and Computer Scientists, Vol. 324 (2013).
22. Vinbæk, E. O., Pettersen, F. M. B., Carlsen, J. E., Fremstad, K., Edvinsen, N., Sandnes, F. E.: On Online Banking Authentication for All: A Comparison of BankID Login Efficiency Using Smartphones Versus Code Generators. In: International Conference on Human-Computer Interaction, pp. 365-374. Springer, Cham. (2019). doi: 10.1007/978-3-030-23560-4_27
23. Sandnes, F. E.: A memory aid for reduced cognitive load in manually entered online bank transactions. In: Norsk informatikkonferanse. Tapir Academic Pubhlishers (2012).
24. Sandnes, F. E.: An error tolerant memory aid for reduced cognitive load in number copying tasks. In: International Conference on Universal Access in Human-Computer Interaction, pp. 614-623. Springer, Berlin, Heidelberg. (2013). doi: 10.1007/978-3-642-39191-0_66
25. Oladimeji, P.: Towards safer number entry in interactive medical systems. In: Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems, pp. 329-332. ACM (2012). doi: 10.1145/2305484.2305543
26. Sandnes, F. E., Huang, Y. P.: A computer supported memory aid for copying prescription parameters into medical equipment based on linguistic phrases. In: 2013 IEEE International Conference on Systems, Man, and Cybernetics, pp. 3447-3451. IEEE (2013). doi: 10.1109/SMC.2013.588
27. Taxis, K., Barber, N.: Incidence and severity of intravenous drug errors in a German hospital. European journal of clinical pharmacology **59**(11), 815-817 (2004). doi: 10.1007/s00228-003-0689-9
28. Wiseman, S., Cairns, P., Cox, A.: A taxonomy of number entry error. In: Proceedings of the 25th BCS Conference on Human-Computer Interaction, pp. 187-196. British Computer Society (2011).
29. Hofseth, K. Å., Haga, L. K., Sørlie, V., Sandnes, F. E.: Form Feedback on the Web: A Comparison of Popup Alerts and In-Form Error Messages. In: Innovation in Medicine and Healthcare Systems, and Multimedia, pp. 369-379. Springer, Singapore (2019). doi: 10.1007/978-981-13-8566-7_35
30. Skogstrøm, N. A. B., Igeltjørn, A., Knudsen, K. M., Diallo, A. D., Krivonos, D., Sandnes, F. E.: A comparison of two smartphone time-picking interfaces: convention versus efficiency. In: Proceedings of the 10th Nordic Conference on Human-Computer Interaction, pp. 874-879. ACM (2018). doi: 10.1145/3240167.3240233
31. Brathovde, K., Farner, M. B., Brun, F. K., Sandnes, F. E.: Effectiveness of Color-Picking Interfaces among Non-Designers. In: International Conference on Cooperative Design, Visualization and Engineering, pp. 181-189. Springer, Cham. (2019). doi: 10.1007/978-3-030-30949-7_21

32. Hansen, F., Krivan, J. J., Sandnes, F. E.: Still Not Readable? An Interactive Tool for Recommending Color Pairs with Sufficient Contrast based on Existing Visual Designs. In: The 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 636-638. ACM (2019). doi: 10.1145/3308561.3354585