Interface Design for Autism in an Ever-Updating World

Lewis D Sawyer
BioSensing for Health Laboratory
School of Computing, University of Kent
Canterbury, United Kingdom
L.Sawyer@ieee.org

Ramaswamy Palaniappan
BioSensing for Health Laboratory
School of Computing, University of Kent
Canterbury, United Kingdom
R.Palani@kent.ac.uk

***Abstract*—It is well-known that people diagnosed with autism spectrum disorder experience difficulty coping with change. However, the extent to which this applies to small changes in user interface design is not known. Through exposing autistic ($n = 11$) and control ($n = 10$) users to seven controlled, cumulative design changes within an e-calendar appointment creation interface, and systematically recording usability scores, a bespoke comfortability score and qualitative responses, we confirm that the autistic discomfort around change extends to small changes in user interface design, negatively impacting system usability and comfortability scores consistently in the autistic group for some tests. To widen the applicability of our findings for industry use and promote further research in this under-explored yet critical area, we also transform our results into preliminary non-exhaustive design heuristics for reducing negative impacts of interface design changes in autistic users.**

***Index Terms*—accessibility, autism, human factors, software updates, user interfaces, user modelling**

I. Introduction

One of the greatest challenges of the modern Human-Computer Interface practitioner is ensuring inclusivity consistently throughout their work. While accessibility in HCI is a growing field, a specific subset of the discipline, software interface design change, and its impacts on accessibility, is a much lesser-studied area. Generally, software design change is viewed as positive by HCI practitioners: a chance to improve a design for the benefit of the user. However, as users with autism spectrum disorder may experience difficulty dealing with small changes [1], do autistic users share this view?

To answer this, we set out to investigate the possibility that the autistic difficulty dealing with small changes extends to small, singular-component HCI design changes. Through reviewing the small amount of existing literature, which we use to create informed research goals, we devise, describe and execute an experiment to capture this reaction, the results of which provide valuable insights into the autistic reaction to change in software design.

Our primary motivation for this study is to raise awareness of the difficulties faced by autistic users when exposed to change in software design, so that they can be considered during the design change process as a part of inclusive design practices.

II. Background

A. Autism Spectrum Disorder

Autism spectrum disorder (‘*ASD*’ or ‘*autism*’), a neurodevelopmental disorder, is defined thoroughly in the ‘*DSM-5*’ [1] – a highly-cited psychology industry “*gold standard*” definition source for “*psychiatric illnesses*” [2]. As a spectrum disorder, the manifestation and extent of symptoms may differ between individuals. There are five required diagnostic criteria, one of which is “*restricted, repetitive patterns of behaviour*”, which can manifest through “*insistence on sameness” and “extreme distress at apparently small changes*” [1].

Autistic people are more likely to form reliance on routines, allowing for a sense of predictability and safety [3]. They may also have a particular affinity for using computers [4], where the repetitive, routine-based nature of computer systems allows for this predictability and safety. While this may bring comfort to autistic users, design-changing software updates pose a threat to these sought attributes. This presents a problem which must be accounted for in inclusive design practices.

B. HCI Change Studies

Some studies explore the impacts of software updates, however there have been very little focus on negative reactions to design changes. One study [5] identifies the stages of installing software updates and lists factors which can negatively impact perceptions of the update at each stage. Despite focusing on the process of updating itself, authors note that “*changes updates bring can also be unwelcome to users who liked the way their systems used to function*”, however impacts of design change itself in the updated software were not explored further.

Authors in [6] explore opinions of updates in hedonic software (video games in this instance), through determine impacts on ‘*continuance intentions*’ (‘*CI*’) based on past experience. Participants mostly exhibited greater CI opinions post-update, indicating that change in video games is seen as positive. However, hedonic software is designed to provoke enjoyment and excitement; change in information systems (e.g. productivity software) may be seen differently due to the functional and routine-based nature of the software. One existing study [7] focuses on the impacts of update frequency and magnitude on CI in information systems. Authors found that small, frequent, incremental updates are preferred over large packaged updates bundling multiple changes. But again, impacts of individual changes themselves were not explored.

Both [6], [7] attribute the increasing relevance of software update reaction studies to the rise of high-speed internet, as it facilitates the rapid distribution of software updates. The choice around updating software is reducing, presenting an increased challenge to autistic users, as there is no option to resist the change. By contrast, in the past, distributing updates through physical media or optional download provided greater scope for update avoidance, allowing autistic users an easy escape from potentially distressing scenarios.

While [7] indicates that smaller, frequent, incremental updates are preferred by users, these may still present as great a threat to autistic users as infrequent, larger updates, due to the greater sensitivity to change in ASD [1].

C. Autism in HCI Studies

Some existing works look at usability of systems for autistic users outside of a software update context, many of which however focus on autistic children, or are based on specialised technologies [8]. No empirical works focusing specifically on autistic adults and reaction to change in software interface design could be found in our literature search attempts.

The lack of autism-HCI studies is explored in [8], where authors attribute the lack of research in the area to autistic users being a "*sensitive user base*", requiring specialist considerations in HCI experiments (such as those listed in [9]). This gap has led to the adoption of low-literacy accessibility guidelines for ASD, to account for the reading comprehension difficulties that impact some autistic users, without considering other challenges faced. This is the case in the current W3C Web Content Accessibility Guidelines [10], and in a study which attempts to compile autism-friendly design rules [11].

Whilst some limited, specialised autism usability studies do exist, as far as we have searched, the autistic reaction to change and its application to HCI continues to be overlooked, despite clearly being a relevant, critical topic.

III. Research Goals

Based on the literature gaps and nuances of autism highlighted in our review, we aim to:

1) Show that autistic users react to change in user interface components negatively, to a greater degree than neurotypical users,
2) Determine the impacts of this reaction on both perceived system usability and the user themselves, and
3) Identify which changes cause statistically significant difficulty in autistic users, to produce non-exhaustive design heuristics for further exploration.

IV. The Method

To meet these goals, we designed an online questionnaire to measure reactions to interface design changes. An initial interface was presented, followed by seven cumulative changes. User responses to each change were recorded systematically between exposures, upon completion of a consistent task. This systematic testing design was inspired by a recommendation in [7].

A. Measures Used

Perceived system usability was selected as a measure based on our assumption that an adverse reaction to a change may result in lower perceived usability scores due to the difficulty encountered by the user. The need for repetitive, simple testing resulted in the selection of the System Usability Scale ('*SUS*') [12]. Other usability measures were deemed unsuitable for repetitive testing due to their length.

To measure affective impacts of change on the user, through taking a symptomatic approach, we introduce a new measure: the '*comfortability score*' ('*CMF*') – a Likert scale from 0 to 10, with custom descriptors at 0 ('*extremely distressed*'), 5 ('*neutral*') and 10 ('*fully comfortable*'), based on descriptions in clinical literature of the reaction to change in ASD [1]. Intermediate points were left blank to allow for a level of subjectivity, as some autistic people may experience '*alexithymia*' (the inability to put feelings into words) [13].

As a previous software update study identified that the scalar nature of quantitative measures can limit depth of response [6], and as autistic subjects may find expressing their thoughts through a rigid scalar format difficult [8], we allowed for optional qualitative comments for each SUS point on each interface, and for each interface iteration overall, with the prompt: "*Are there any other comments you would like to share about this change, or how it made you feel?*".

B. The E-Calendar System

The change tests were presented as iterations of an e-calendar system '*create calendar appointment*' interface, with a ribbon-style button system, as is commonplace in similar productivity software. Our justifications for this choice are threefold: for simplicity, removing elements of complexity that could cloud responses, for modularity, to allow for the easy automation of interface design change delivery, and to leverage a pre-existing schema of use within the user's memory, allowing for a more realistic reaction to change. This design was inspired by [7], where a basic word processing application is used for similar reasons.

The seven tests were chosen based on identifiable design components within the initial design and the technical feasibility of implementation, and were justified by assessing the symptoms of ASD that may be triggered. Changes were applied cumulatively to mitigate reaction to change removal. For each design change, there is an implementation and justification description in Table I.

See supplementary material **I** for the given task to complete and **II** for the interface iterations.

C. Test and Control Groups

The test group for this study were human participants aged 18 or above with a clinical diagnosis of autism spectrum disorder. The control group were human participants aged

TABLE I: Change tests, implementations and justifications

Change	Implementation	Justification
1: Colour scheme (Fig. **S4**, Fig. **S5** and Fig. **S6**)	Colour scheme changes from blue to green.	Green stimulates different eyecolour receptors to blue [14], which may trigger discomfort around change due to the contrasting visual input.
2: Font face (Fig. **S7**)	Interface font changes from Arial to Helvetica.	There are minor noticeable differences between the two fonts, but due to distress at "small changes" [1], autistic users may notice this.
3: Text wording (Fig. **S8**)	The wording of text in the interface changes to synonyms. The task wording does not change.	Changing wording (despite location remaining constant) may introduce a considerable cognitive workload and potentially impact autistic users.
4: Button height (consistent sizing change) (Fig. **S9**)	The height of the ribbon (and the buttons within) reduces from 100px to 60px.	The sizing change is consistent and non-functional, however as a visual change in itself, it may cause distress in autistic users.
5: Button width (inconsistent sizing change) (Fig. **S10**)	The width of the 'send' button increases from 75px to 125px. The text within is made bold.	Introducing inconsistency tests whether the autistic "insistence on sameness" [1] spans to sameness in formerly uniform design elements.
6: Button order (Fig. **S11**)	The order of ribbon buttons within their group is changed, apart from 'submit' and 'clear'.	A memory mismatch situation where the interface no longer reflects the learned routine could impact autistic users to a greater degree.
7: Ribbon split (Fig. **S12**)	Functional buttons are moved above the main text box, and terminal buttons are moved below.	This tests whether changing an interface to better represent the order of the task (routine) will be welcomed, or rejected, in autistic users.

18 or above that do not have a clinical diagnosis of autism spectrum disorder.

Exclusion criteria included diagnosis of any other condition which may impact cognitive or visual function, including dyslexia, colour blindness, conditions impacting memory retention, and other neurodevelopmental disorders (e.g. ADHD), or taking medication which can impair such functions.

The questionnaire required a PC to complete; mobile device browsers could not access the questionnaire. Participants self-indicated eligibility and group membership, and ineligible participants could not complete the questionnaire.

D. Recruitment

To directly recruit autistic participants for this introductory study, the University of Kent's '*Student Support and Wellbeing*' department distributed a call for participants to registered undergraduate and postgraduate students aged 18 and above, with a diagnosis with autism spectrum disorder and no other disability. Control and further autistic participants were recruited via posters displayed in strategic areas across the University of Kent's Canterbury campus.

E. Institutional Review Board Approval and Positionality

This study received an institutional review board approval from: *Central Research Ethics Advisory Group, University of Kent, Canterbury, United Kingdom* (ref: '*CREAG018-11-23*').

Modern disability social research suggests two approaches to researching disability: the medical model, focusing on impairments and treatments, and the social model, focusing on abilities and inclusive practices [15], [16]. As a scientific work, despite being criticised for its potential to stigmatise disability, it is necessary for us to refer to ASD through the medical model, exploring symptoms with the use of reputable clinical literature. However, our core aim is to raise awareness of the challenges faced as a result of ASD in response to software design change, and propose improvements to inclusive design practices to reduce barriers in computer use, a goal which aligns with the principles of the social model of disability.

V. Results

A total of 11 responses from autistic subjects (two of which were incomplete) and 10 control responses were collected. All participants self-declared eligibility and group membership.

A. Quantitative Data Statistical Analysis

Non-parametric statistical methods were employed (due to the small sample size) to facilitate hypothesis testing based on our research goals, using MATLAB. Two incomplete test group responses were excluded from analysis.

To adjust the group multiplicity to $1 : 1$, to allow for fairer assumptions in non-parametric testing, three anomalous responses were excluded: a test group response with consistently very high SUS scores, a control group response with consistently very low SUS scores, and an unreliable control response where a comment described 'scrolling', indicating the browser window was too small to complete the questionnaire.

A significance of 0.05 was used for every test, with the null hypothesis *'there is no significant difficulty caused in the test group'*. A right-tailed Mann-Whitney U Test was used with test vs. control SUS scores and CMF scores on a per-interface basis, to determine the probability of a drop in test group scores in relation to control. To catch cases where both groups' scores suffered, a right-tailed Wilcoxon signed rank test was used to calculate the probability that the prior interface's scores will be larger than the current (indicating a drop).

Using these test results, through hypothesis testing we determined which changes caused significant difficulty, be it through usability or discomfort, within each group (Table II).

Averaged SUS and CMF scores per participant are presented in boxplot, scatter and violin plots in Fig. 1 (produced in

TABLE II: Research hypotheses

Change	**Test**[a]	p	H_r
1: Colour scheme	U-test: SUS	0.0287	Changing the colour scheme of an interface causes difficulty in autistic users.
2: Font face	U-test: SUS	0.0365	Changing the font face of an interface causes difficulty in autistic users.
3: Text wording	WSR: SUS WSR: CMF	0.0209 0.0289	Changing the text wording of an interface causes difficulty in autistic users.
3: Text wording	WSR: SUS	0.0295	Changing the text wording of an interface causes difficulty in control users.
6: Button order	WSR: SUS	0.0272	Changing the order of elements in an interface causes difficulty in autistic users.

[a] '*WSR*' refers to the Wilcoxon signed rank test.

MATLAB using an external package [17]). The autistic SUS scores show greater variability, and the autistic CMF median is significantly lower than the control. The autistic CMF score distribution is slightly skewed towards the lower half of the scale, whereas control CMF scores are evenly distributed.

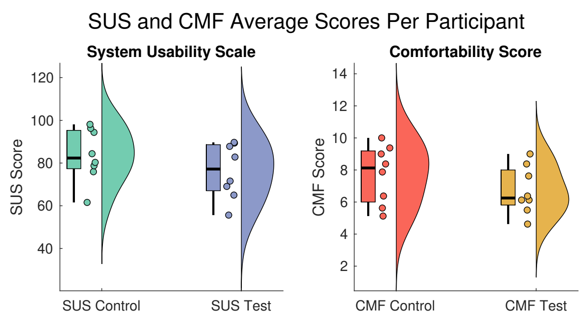

Fig. 1: Averaged SUS and CMF response data plots, per-group

B. Qualitative Data Analysis

A total of 103 (45 ASD and 58 control) interface change comments were categorised and a resultant sentiment score was calculated per-change, per-group (shown in supplementary material **III**). Responses excluded from quantitative analysis were restored, with irrelevant data omitted manually.

Sentiment towards the change of colour varied. Control responses were mostly based on colour preference without reasoning given, whereas autistic responses were based on usability impacts or emotion: "*I preferred this... the colour difference between a selected and unselected toggle was clearer*", and: "*the green was too bright and garish, I liked the blue better because it looked more relaxing*".

The font face change revealed that, despite some test and control subjects stating they could not identify what had changed, autistic usability still suffered. One autistic participant stated: "*I did not notice a change but something doesn't feel right. I feel quite strange*". Another followed suit: "*I cannot immediately identify what the change is*". Alongside the consistent SUS drop in the test group, these findings support that changing the font face impacts usability in ASD.

Autistic responses for the text wording change indicated clear discomfort consistently amongst the autistic group, with terms such as "*stressed*", "*quite uncomfortable*" and "*irritating*" used, whereas control group comments contained less severe emotional wording (only "*difficult*", "*weird*" and "*confusing*"). The test group displayed greater distress within these responses, which is reflected in the quantitative analysis.

The autistic group unexpectedly showed a consistently positive reaction to the inconsistent sizing change, with comments stating that this a "*good thing*" or that they "liked" this change. In addition, one SUS comment under point *"I think there is too much inconsistency in this system"* [12]) reflected on the use of inconsistency to emphasise functionality: *"The bigger send button, although inconsistent, makes itself stand out in a good way."* There was no negative impact on any of the autistic comfortability scores for this change.

In response to the button ribbon height change, one ASD subject wrote: *"this was the biggest visual change aside from the colour change so I had to re-visualise what I was meant to be doing."* This was unexpected as the task flow remained the same and the change was uniform. While distress is not shown, it does show that the autistic user was working from a mental visualisation of the interface.

One autistic user left a detailed account of their experience with the change of button order, stating their frustration around *"pointless change"* (see supplementary material **IV**).

Two autistic questionnaire responses were abandoned. In one case, the user was exposed to the text wording change and said: *"the change of names on the interface stressed me as it didn't follow the instructions"*. The wording of the task deliberately did not change to simulate the user re-learning the new terminology, without a guide to rely on. Upon completing the next change test, the participant wrote: *"one change of words can often throw me"* and then closed the questionnaire. The text wording change caused significant impacts in the user to the extent where they could not continue, further supporting the research hypothesis built on this test.

VI. Reflections, Limitations and Future Works

A. Assessing the Effectiveness of our Methods

While our sample size of 11 ASD and 10 control subjects was sufficient to show significance in non-parametric testing, a larger sample size would improve statistical power.

As a spectrum disorder, autistic individuals' symptoms can vary in manifestation and sensitivity. The DSM-5 [1] provides three "*severity levels*" of ASD. "*Inflexibility of behaviour*" is cited across all three levels, however as the questionnaire was distributed via a university's disability support services, it is likely that it reached only severity level 1 ("*requiring support*") autistic individuals, as individuals at higher severity levels are less likely to partake in university education due to the cognitive impairments experienced [1]. Despite this, our findings remain applicable as a minimum case, with reactivity to change increasing with severity level. We recommend that, in future studies, a broad range of ASD severity levels be targeted, and warn of the limitation of using a university support service.

Researchers must also consider *alexithymia* (the inability to put feelings into words) [13], a potential trait of ASD which may impact the reliable completion of usability questionnaires. We had attempted to mitigate this when designing the comfortability score by leaving intermediary scale labels blank, to allow for subjective scoring. This may become more important when relying on self-indication of usability from autistic subjects at higher severity levels, which must be considered.

We also remind that, due to the distress that may be caused by exposure to change, autistic participants may abandon future studies at a higher rate. Despite this, these abandoned responses may still hold value, as we have shown in our qualitative analysis, and we urge researchers to record and actively examine abandoned responses to see such reactions which otherwise may be overlooked.

B. Assessing the Effectiveness of Measures Used

Due to the nature of the SUS questions as a usability evaluation [12], a minority of participants commented more on the functionality of the design as opposed to the introduced change, despite a prompt given to focus on the change introduced. A slight re-wording of the SUS questions to clearly refer to the change introduced would benefit future studies, especially where autistic subjects are included, who are more likely to interpret instructions literally [1].

Using the dataset prepared for statistical analysis, by transforming the data to average-response-per-participant SUS-CMF pairs, non-parametric correlation checking was performed to test whether usability and comfortability scores significantly correlate. Based on the results in Table III, a positive correlation is shown between SUS and CMF scores within the autistic group and the combined groups, and is stronger within the ASD group, showing that usability may have a stronger association with comfortability in autistic users. See supplementary material **V** for data plots with linear fit lines and confidence bounds showing SUS-CMF correlation between and across groups.

TABLE III: SUS-CMF Spearman rank correlation test results

Group	ρ	p	H_r
Autistic	0.8383	0.0131	There is an SUS-CMF correlation in the test group.
Control	0.6429	0.0962	There is not an SUS-CMF correlation in the control group.
Combined	0.7094	0.0021	There is an SUS-CMF correlation in the combined dataset.

Collecting qualitative data allowed for further insight into the impact of change on the participants, supported our quantitative findings and provided feedback as to the methods used. We recommend that further change impact studies collect such responses to obtain this depth that is otherwise unobtainable through quantitative measures. We also note that, as our qualitative measures were optional, their results are not representative of every user, but of individual cases. This does not mean that they are without value, however.

C. Assessing the Effectiveness of the Testing Framework

While some participants described the interface as "*simple*" and "*basic*", it served its purpose: to deliver controlled, modular design changes. Researchers embarking on a deeper design change impact study may wish to develop a more complex interface, with refined functional design elements and complex workflows to test, but should be cautious to ensure that findings are generally applicable and that non-tested elements are controlled.

It is important to note that our study explores only individually introduced, static design changes. The change of animations within an interface, the introduction of multiple changes, and assessing resilience in terms of change magnitude in ASD are all areas for future work; however, based on our findings here, we predict that such scenarios will cause additional discomfort within autistic users.

VII. Forming our Heuristics

Through our analyses, we have successfully identified changes which caused usability or comfortability scores to significantly suffer in our test and control groups. From these, we present the following heuristics for reducing the negative impacts of interface design change in autistic users:

1) Keep the interface colour scheme consistent.
2) Use the same font face.
3) Retain existing text wording (unless the resulting functionality has changed).
4) Ensure consistent design element ordering (do not make unnecessary ordering changes).

VIII. Conclusions

Through systematically recording the impacts of controlled user interface design changes on autistic and neurotypical users, analysing our data and asserting our research hypotheses, we have successfully met the research goals of our study: we have shown that autistic users do indeed react to change in interface design negatively (to a greater extent than neurotypical users), that these negative impacts impair usability and inflict discomfort, and have produced introductory heuristics based on our results for reducing the negative impacts of interface design change in autistic users.

These heuristics, alongside the presented need for research in this area, our method for design change impact testing, including the introduction of the comfortability score, results, considerations around involving autistic users in HCI research, and suggested avenues for further exploration are the contributions of this research to the HCI discipline.

References

[1] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Arlington(VA): American Psychiatric Association, 2013.

[2] M. Riba, S. Sharfstein, and A. Tasman, "The American Psychiatric Association," *International Psychiatry*, vol. 2, pp. 18–20, 7 2005.

[3] M. Gomot and B. Wicker, "A challenging, unpredictable world for people with autism spectrum disorder," *International Journal of Psychophysiology*, vol. 83, no. 2, pp. 240–247, 2012.

[4] J. A. Kientz, G. R. Hayes, M. S. Goodwin, M. Gelsomini, and G. D. Abowd, "Introduction," in *Interactive Technologies and Autism, Second Edition*. Cham: Springer International Publishing, 2020, pp. 1–13.

[5] K. E. Vaniea, E. Rader, and R. Wash, "Betrayed by updates: how negative experiences affect future security," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 2671–2674.

[6] D. Franzmann, "User expectations of hedonic software updates," in *Proceedings of the 21st Pacific-Asia Conference on Information Systems*, 2017.

[7] M. Fleischmann, M. Amirpur, T. Grupp, A. Benlian, and T. Hess, "The role of software updates in information systems continuance — an experimental study from a user perspective," *Decision Support Systems*, vol. 83, pp. 83–96, 2016.

[8] M. Alzahrani, A. L. Uitdenbogerd, and M. Spichkova, "Human-computer interaction: Influences on autistic users," *Procedia Computer Science*, vol. 192, pp. 4691–4700, 2021.

[9] D. Çorlu, Ş. Taşel, S. G. Turan, A. Gatos, and A. E. Yantaç, "Involving autistics in user experience studies: A critical review," in *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. New York, NY, USA: Association for Computing Machinery, 2017, pp. 43–55.

[10] World Wide Web Consortium, "Cognitive accessibility at W3C," w3.org, 2022, accessed: 2024-04-16. [Online]. Available: https://www.w3.org/WAI/cognitive/

[11] N. Pavlov, "User interface for people with autism spectrum disorders," *Journal of Software Engineering and Applications*, vol. 07, pp. 128–134, 01 2014.

[12] J. Brooke, "SUS: A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, 11 1995.

[13] J. Poquérusse, L. Pastore, S. Dellantonio, and G. Esposito, "Alexithymia and autism spectrum disorder: A complex relationship." *Front Psychol*, vol. 9, p. 1196, 2018.

[14] H. v. Helmholtz, "The young-helmholtz theory of color vision, 1860." in *Readings in the history of psychology*, W. Dennis, Ed. East Norwalk, CT, US: Appleton-Century-Crofts, 1948, pp. 199–205.

[15] M. Anderson-Chavarria, "The autism predicament: models of autism and their impact on autistic identity," *Disability & Society*, vol. 37, no. 8, pp. 1321–1341, 2022.

[16] J. A. Haegele and S. Hodge, "Disability discourse: Overview and critiques of the medical and social models," *Quest*, vol. 68, no. 2, pp. 193–206, 2016.

[17] P. Karvelis, "daviolinplot - violin and raincloud plots," Jul. 2024. [Online]. Available: https://doi.org/10.5281/zenodo.12749045