

# Effect of Fidelity in Diagram Presentation

Louise Yeung<sup>1</sup>, Beryl Plimmer<sup>2</sup>, Brenda Lobb<sup>1</sup>, Douglas Elliffe<sup>1</sup>

Department of Psychology<sup>1</sup>, Department of Computer Science<sup>2</sup>

University of Auckland, New Zealand

Phone +64 9 373-7599

l.w.s.yeung@gmail.com, {b.plimmer|b.lobb|d.elliffe}@auckland.ac.nz

## ABSTRACT

The visual fidelity (fidelity) of a design diagram affects perception and design performance. Hand-drawn diagrams are more effective working documents for early design tasks such as user interface designs than the equivalent computer-prepared formal representation. However people prefer more formal representations because they feel that hand-drawn diagrams look unprofessional. Sketch-based design tools make it possible to present partially tidied designs. We have postulated intermediary levels of visual fidelity in a systematic manner and implemented these levels into a sketch tool to evaluate the effect of computerization and fidelity on perception and design performance. Our findings show that: performance decreased systematically with increased fidelity; that computer presented designs decreases performance and that performance was decreased by computerization of the hand-drawn diagrams. In contrast, user satisfaction was higher with increasing levels of fidelity. These results pose challenges to the sketch tools community and further questions for effective computer support for early design.

## Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Design, Human Factors

## Keywords

Design fidelity, sketching, sketch tools

## 1. INTRODUCTION

Sketching diagrams is a time-honoured technique for expressing ideas that is of particular value in the early stages of problem solving and design. Black[2] noticed in 1990 that students who created their initial designs using a computer tool constrained their design choices to things

that they could easily achieve with their current tool competency. In contrast, those students who created their initial designs as hand-drawn sketches did not constrain their design solution; rather they extended their tool knowledge to realize the design.

Similarly Wong[30] in 1992 related experiences with presenting her designs to colleagues and clients. If she presented a formal functional prototype she would get feedback on the visual details such as colour and font, when she wanted feedback on the overall design and functional aspects of the system. She reported that she would convert computer-produced, semi-functional prototypes to paper sketches in order to get appropriate feedback.

Both of these reports are from the very early days of end-user visual design tools. These types of tools are now widely available, yet the problems are still evident. There are thought to be a number of advantages of informal, hand-drawn diagrams. First, it is much quicker to construct a hand-drawn diagram, therefore the diagram can act as an external support for short-term memory and problem solving[11]. In contrast, creating the same diagram with a component based drawing tool takes longer and requires the user to deviate from the primary design task to the selection of components and their placement and alignment on the drawing canvas[22]. Second, sketching is quintessentially imprecise and ambiguous: the need to select a specific type of diagram component is a negative aspect of current widget-based design environments as the designer may be forced to make premature decisions[12]. With a sketch the designer can record an element and leave it semi-defined or ambiguous until later in the design process. In comparison, in many tools it is difficult to change the type of an element without completely recreating it, thus further discouraging exploration of alternatives. Lastly, a hand-drawn diagram looks incomplete whereas formal diagrams appear complete. The unfinished look seems to make it easier for both the designer and others to consider changes [30].

Studies comparing designing on paper or whiteboard with computer sketch tools and computer design tools have concluded that sketching is better for design and review phases of the process [1, 10, 22]. However, the research also shows that many people express a preference for presenting formal designs to their superiors and clients as

they believe that a formal design looks more professional than a hand-drawn design [6, 18, 21].

Computer-based sketch tools make it possible to transform, automatically and incrementally, a diagram from hand-drawn to hi-fidelity. This process of increasing fidelity has been called 'beautification' and has been investigated previously [6, 16]. However, our study is the first to define a taxonomy of visual fidelity and to isolate presentation medium from fidelity. The evaluation considers the effect of medium and intermediate levels of fidelity on the design process.

The structure of the rest of this paper is as follows: First, we review related work on diagram fidelity. Next we present a taxonomy of the dimensions of fidelity and their implementation into a sketch tool to support systematic, incremental beautification of sketches. Following this we describe an experimental evaluation of its effects on design performance and the results are presented and discussed. Our conclusions pose challenges to the sketch tools community and suggest that other design support technologies should be carefully examined.

## 2. BACKGROUND

Early observations [2, 30] of the effect of computer-based design tools were confirmed by Goel's [10] studies. He found that computer tools adversely affected the design process with less time being spent on creativity but more time devoted to tasks such as aligning objects perfectly.

A number of studies have compared combinations of paper or whiteboard with computer sketch tools and computer design tools. Bailey and Konstan [1] compared designing multi-media software on their sketch tool Demais with Authorware (multi-media software) and paper design. Demais includes embedded images, audio and video and supports interaction with the low-fidelity design. The evaluation asked participants to rank the three treatments for a number of design related sub-tasks. Pencil and paper out-ranked Demais and Authorware except for communicating behaviour, where Demais was the highest ranked. Authorware was considered the least useful in all respects. Plimmer and Apperley [22] compared Freeform, a user interface (UI) design sketch tool and traditional whiteboard for constructing UI designs. Freeform, like Demais, has an interactive mode. The study participants felt they had a better understanding of the functional requirements of the problem using Freeform. In a second study [23], participants reviewed designs rendered informally in Freeform or as formal diagrams in the Visual Studio Form designer. Significantly more substantive changes were made to the informal rendering of the designs. However the participants thought they had made more changes to the formal design as they had spent more time on tidying tasks. In contrast, Walker et al. [29] found no differences in a four-way comparison between high- and low-fidelity, paper and computer for identifying usability

problems. In this study both the computerized designs were interactive while the paper was not.

SketchXML [6] is a sketch tool that supports three levels of fidelity: low (hand-drawn, with markers in place of text), medium (widgets regular shapes) and high (widgets regular shape with labels). Pen input is immediately re-rendered to the selected level of fidelity. The evaluation study of the different levels of fidelity in this tool required the users to create two user interfaces in each fidelity. Users were questioned on their preference: Almost all the users strongly preferred the high- or medium-level fidelity and disliked the hand-drawn level. The study examined user preference but did not consider design performance.

A number of other sketch tools have incorporated some beautification. Igarashi et al. [13] assumed that pen strokes are intended to be either straight lines or regular curves, and immediately morphed pen strokes into high fidelity line segments constrained to lie at fixed angles and connections and intersections are exact.

Freeform [24] and Silk [14] take a different approach: The sketch is retained as is, but a formal interface can be generated from the sketch. Sumlow [7] keeps a synchronized formal rendering of UML class components. Denim [19], a web site design tool, immediately recognizes simple symbols such as rectangles and lines. This tool varies the degree of beautification applied according to the zoom level. For example, at storyboard level a line drawn to indicate navigation between pages is smoothed and has a dot added to the source point and an arrow to the destination point while at page level it is rendered as raw ink. Handwriting can be similarly morphed from original hand-drawn to a formal font [25].

The various comparative studies suggest that sketching and hand-sketched representations are more effective artifacts to work with at the early stages of design. However people continue to express a preference for formal representations [6, 21] and spend time on beautifying representations [3]. Bolz [3] claimed that fifty percent of the total time spent on creating formal designs on a computer is on beautification operations such as aligning and sizing the components.

To explore the effect of medium (paper or computer) and partial beautification of sketches, we developed a taxonomy of the dimensions of visual fidelity, along which attributes of design components may vary, and implemented these into a sketch tool. This environment was then used as a platform for the subsequent evaluation study.

## 3. TAXONOMY

Consider a sketched design for a user interface, such as is shown in Figure 1a, compared with a formalized representation of the same design (Figure 1e). The components used in the website design - words, textboxes, dropdown lists, radio button checkboxes and labels - are the same in both representations but the attributes of those

**Table 1 A taxonomy of attributes of design components and examples of how each may be beautified**

	Hand-drawn	Low-level Fidelity	Medium-level Fidelity	High-Level Fidelity
Smoothness				
Size				
Alignment horizontal				
Alignment vertical				
Spacing horizontal				
Fonts	Subscription Form	Subscription form	Subscription Form	Subscription Form

components - smoothness of the lines, relative sizes, horizontal and vertical alignment of the components, spacing between components and lastly, the appearance of the letters in the text - differ. The words in the sketch are irregular hand-written letters. Rather than progressively morphing the handwriting (which is technically difficult [25]) we have replaced the words with increasingly formal fonts: My Font Tool [17], Gullim and Times New Roman. The My Font Tool [17] uses a handwritten example of each letter to create a font set with the spacing between letters and words adjusted appropriately. The effect is that letters appear handwritten but each example of a letter is identical – this results in a tidier appearance. Sans-serif fonts are considered less formal in appearance than serif fonts. Gulim was chosen because it is an uncommon sans-serif font. For high-level fidelity we

selected the classic Times Roman font for its serifs, association with printed books and regular appearance.

#### 4. IMPLEMENTATION

The beautification of designs through changes to attributes of components as described above has been realized in InkKit [20]. InkKit is a software toolkit for creating sketched diagrams with a pen on a tablet PC. The InkKit recognition engine is used to identify the sketched elements such as textbox, dropdown list, radio button checkbox and label.

Systematic application of increases in the visual fidelity of the sketch can then be applied. For the purpose of this study, four levels of fidelity were proposed based on the literature on perception and reasoning to evoke a visual

continuum of fidelity. Thus a diagram can be presented at four different levels of fidelity: hand-drawn, low-level fidelity, medium-level fidelity and high-level fidelity (formal).

Ink stroke smoothing is thought to be the attribute most critical to changing the visual appearance; also it appears to interact directly with all the other attribute changes. Smoothing is applied at three levels. To smooth an ink stroke first the type of shape is identified as a line, polygon or ellipse. Using the smallest external rectangle as a guide, the bounding box, points are identified for each type of shape, end points of the line, corners for the polygon and the centre of the circle. From these points a perfect, regular shape could be scribed. However at low- and medium-level beautification the stroke is morphed from its current position to 1/3 or 2/3 of the way to the 'perfect' stroke. For high-level beautification it is morphed to the perfect stroke. Table 1 Smoothing, shows the effect of this on a rectangle, triangle and circle.

Size is standardized using the bounding box of the components. All components of the same type are grouped and an average size calculated and normalized across the group. All members of the group are then resized so that their bounding boxes are the same size. An example of the application of this to textboxes, dropdown lists and radio buttons is shown in Table 1 Size. Theoretical and empirical research on human perception (e.g. [4, 5, 9, 15, 27]) suggests that different ratios of alignment would produce different levels of fidelity of a design: the more aligned the elements, the more formal the design will appear. Horizontal and vertical alignment are applied separately to a diagram. Both use the same

techniques. For horizontal alignment first each component is grouped with other components that are in approximately the same row (these approximations are calculated using mid-points and top/bottom extremities). Then an average bottom point is calculated and all the members of the group are moved to that horizontal position. For vertical alignment the same process is applied with grouping by column, the bottom point replaced by the left point. As this process can result in components obscuring or overlapping each other, regular spacing is applied as a part of the same process.

Changes in size and spacing are applied together on the lowest fidelity diagram as these attributes do not seem to be as critical to the visual appearance of a sketch as smoothness and alignment. Further smoothing can result in inconsistent size, alignment and spacing; these attributes are corrected as required.

Regarding size, all textboxes had an approximate height of 10mm and most had a width of 85mm. Similarly, dropdown menus had a triangle with a height and width of 50mm on the right hand side of the menu; all radio buttons had an approximate diameter of 10mm, although sizes could not be perfectly exact due to different levels of smoothness/roughness of beautified lines; labels had an approximate height of 10mm, though width of labels was uncontrollable as word length and the number of words contained in a label varied. Additionally, all labels were programmed to be spaced 50 pixels apart vertically and their associated controls (textbox, dropdown menus and radio buttons) were aligned to the labels horizontally at 30 pixels apart. Lines were smoothed systematically as shown in Table 1.

**Table 2 Application of fidelity taxonomy**

Condition	Fidelity level	Medium	Line Smoothing	Alignment	Fonts
1	Low fidelity (hand-drawn)	Paper (and pen)	0%	No elements are aligned exactly	original handwriting
2	Low fidelity (hand-drawn)	Tablet PC	0%	No elements are aligned exactly	original handwriting
3	Medium-low fidelity	Tablet PC	33%	Vertically and horizontally, every third element is aligned to the first element from the top-left of the same type	Standardized handwriting
4	Medium-high fidelity	Tablet PC	66%	Vertically and horizontally, every second element is aligned to the first element from the top-left of the same type	Gulim
5	High fidelity (computer-rendered)	Tablet PC	100%	Every element aligned exactly according to its type	Times New Roman

The appearance of writing (along with smoothing) is also thought to be critical to the visual appearance of the diagram. Therefore at each level we replace handwriting, with increasingly formal fonts as shown in Table 1. A summary of the beautification applied at each level is shown in Table 2. A casual visual scan of the example of diagrams at each level of fidelity shown in Figure 1 suggests a continuum of fidelity.

## 5. EVALUATION

We ran a counterbalanced experiment in which participants were asked to review and amend five web page designs that varied by medium and fidelity. We counted the number and type of amendments made and surveyed the participants' opinions.

### 5.2 Method

#### Participants

A within-subject repeated-measures design was used to measure the effects of fidelity on the design performance of 30 student volunteers (mean age 22.8 years, SD = 5.9), all with no previous experience with the either design tool used in the study or tablet PCs. Twenty of the students were computer science (CS) majors and of those 15 had user interface design experience. None of the non-CS students had design experience.

#### Apparatus and setting

In Condition 1, where the design was presented on paper, the participant was instructed to make changes on the original design using the blue ball-point pen provided and to use a sheet of blank A4 paper, also provided, if more space was needed. In Conditions 2-5, the designs were presented on a 15" CRT colour (LCD) screen on a Toshiba Tablet PC (Edition 2005, Intel® Pentium® M, 1600MHz, 590MHz, 512RAM, Microsoft Windows XP Tablet OS) with 1280 X 1024 pixel resolution. The participant used the tablet's stylus to draw directly on the tablet screen as if drawing on a piece of paper. Morae Recorder (2004) was used to record and save all actions performed by the participant on the computer including input from the mouse, stylus and/or keyboard, and screen captures of the information visible on the screen. The studies were carried out in a quiet office.

#### Designs and scenarios

Five form design problems were created with equivalent requirements and scenarios of use. Each problem was as equivalent as possible, in that 1) all forms served the same purpose, i.e. required users to fill in personal information; 2) element types appeared in the same order in all designs; 3) the number of each type of element was the same in all designs, and thus, 4) the total number of elements was the same in each design (58). From these problems, five form designs were made in compliance with HTML design guidelines [28] and interface and web design principles [5, 8].

We then introduced functional design errors into each of the designs, deliberately breaking guidelines and interface design principles. The number (22) and types of errors were the same across all designs. Finally the beautification variables were applied as described in Table 2. This resulted in 25 different design presentations.

### 5.3 Procedure

A Latin square design was used to control for presentation order effects. Each participant underwent five conditions with each condition drawn from a different problem. In each condition, the participant was presented with an early-stage HTML form design and an accompanying scenario describing the form's purported use.

In Condition 1, the design was at the lowest level of fidelity and was presented on paper. In Conditions 2-5 the design was at one of four levels of fidelity, from low through medium-low and medium-high to high, and was presented on a tablet PC.

Informed consent was obtained from participants and training in the use of InkKit and task requirements was provided, followed by a guided practice session until participants demonstrated the abilities required for the task. The experiment then began and participants were provided with their first design and its associated requirements and scenario. The requirements were presented to help the participant identify whether the correct information was being sought from the end-user who would be filling in the HTML form, which in turn, guided the participant to add, delete and relocate elements/items and/or item sets appropriately. The scenario was presented to suggest whether an element was of the appropriate type (change element) and size (resize).

Participants were instructed to make as many changes as they wished to "improve the design to make it better serve its purpose". When the design was presented on the tablet PC, they were instructed to stay within the application and to use only three functionalities (draw, erase, move). Participants were given a maximum of 12 minutes per condition to complete their desired changes, because piloting suggested this was adequate for people to complete the task without undue time pressure.

A post-task questionnaire was used to measure preference for (1) design medium (paper or Tablet PC) during the experiment and in the 'real world' outside the experiment; and (2) design fidelity level, in terms of ranking of overall enjoyment of the designs from the most-liked condition (5) to the least liked condition (1).

### 5.3 Analysis

There are multiple 'good' design solutions for a form that comply with recognized design guidelines. Therefore it was anticipated that along with expected changes (i.e. fixing introduced errors), participants would make other changes to the design. All changes made at each level of fidelity were recorded (*total changes*) and categorised as

either *quality changes* (those which conformed with recognised good design guidelines) or *other changes* (those which did not). Within quality changes, *expected changes* (changes to deliberate errors) were identified. As the number of deliberate errors was the same in each design presented to the participants, the number of corrections to these errors (expected changes) made in each design was measured to allow for controlled comparisons between conditions, to explore the effects of

fidelity on design-decisions during early stages of the design process (i.e. during early prototyping).

Measuring other changes allowed assessment of the equivalence of the experimental designs used (if the number of each type of change did not differ significantly between the five designs, it could reasonably be assumed that any effect found was due to the experimental manipulation of the independent variable rather than differences between designs).

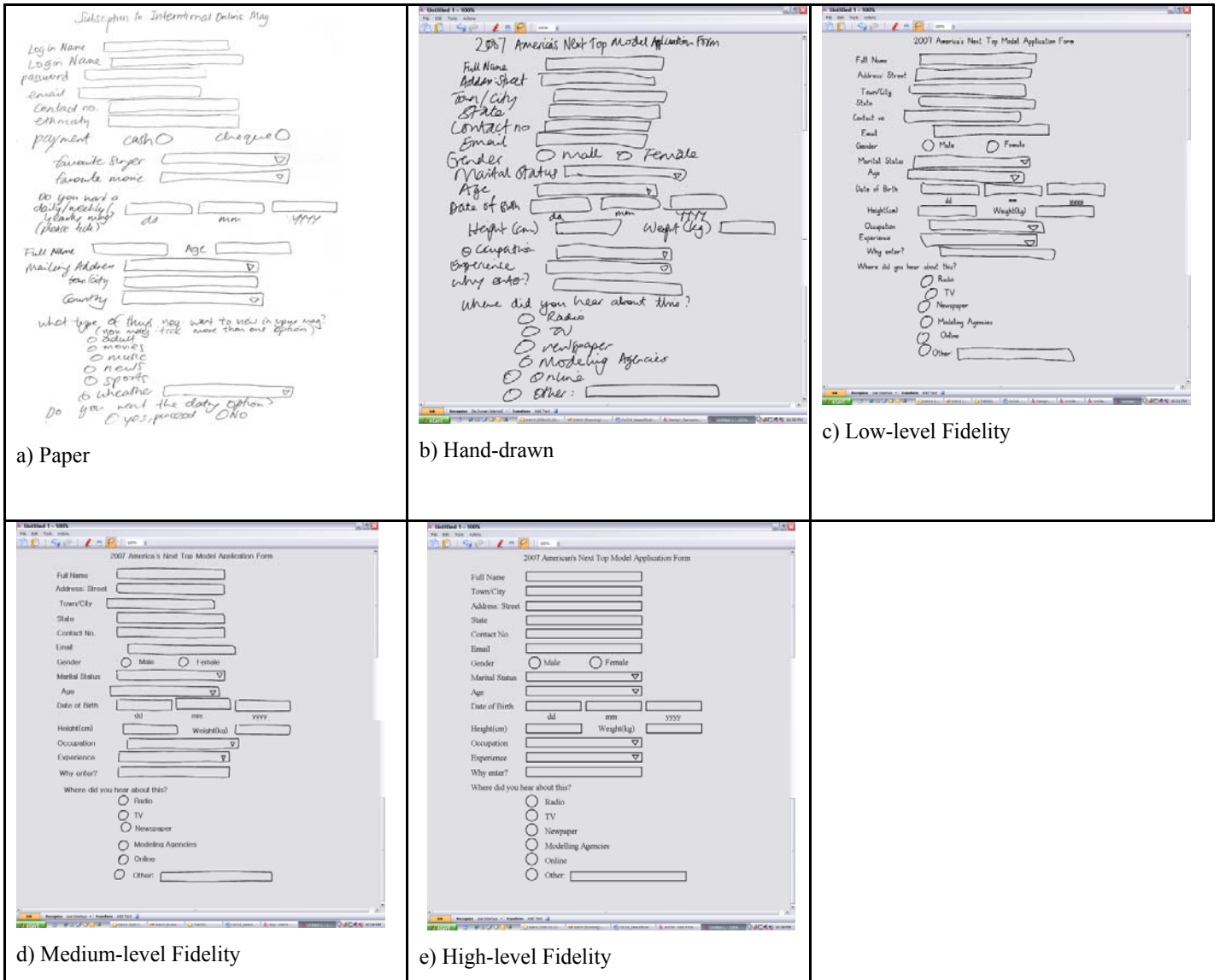


Figure 1 User interface designs at four levels of fidelity

## 6. RESULTS

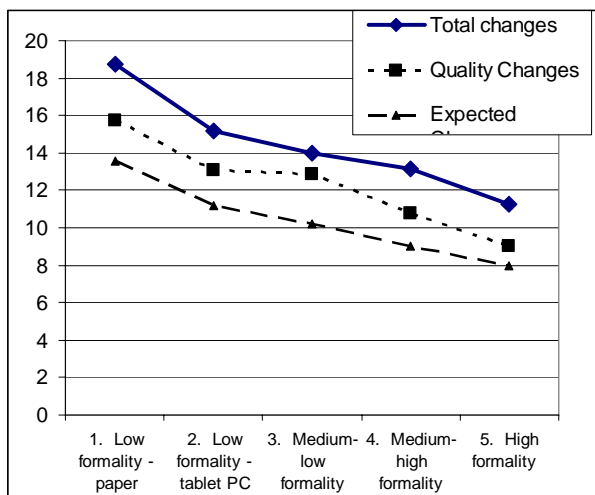
A significant main effect of fidelity on **total** changes was revealed by one-way repeated measures ANOVA (Wilk's Lambda = .265,  $F_{4,26} F(4, 26) = 17.99, p < .001$ , multivariate partial Wilk's Lambda = .265,  $F(4, 26) = 17.99, p < .001$ , multivariate partial  $\eta^2 = .74$ ) (Figure 2). A strong significant linear trend was also found, ( $F(1, 29)$

= 59.59,  $p < .001$ , partial  $\eta^2 = .67$ ), over the mean value of total changes made at each level of fidelity. A weaker, but significant, cubic trend was found, ( $F(1, 29) = 8.529, p < .01$ , partial  $\eta^2 = .23$ ), suggesting that overall, the number of total changes made was the highest when participants were presented with the low-fidelity design

on paper, and decreased systematically as fidelity increased.

Variability between subjects decreased systematically across Conditions 1-5, from low to high fidelity (SD range = 6.57-3.51). Pair-wise comparisons (with Bonferroni adjustment for multiple comparisons) showed that the *difference* in number of changes made increased as the level of fidelity decreased. The total number of changes made to the low fidelity design presented on paper was significantly higher than to all other designs, while the total number of changes made to the highest fidelity design was significantly lower than to all other designs ( $p < .05$ ).

Interestingly, even though there were two low fidelity conditions, one presented on paper and one presented on the tablet, the total number of changes made still differed significantly between these conditions – the mean difference was 3.57. While there was no significant difference in the mean number of total changes between medium-high fidelity and medium-low fidelity conditions or between medium-low fidelity and low fidelity on the Tablet PC, the total number of changes made at low fidelity was significantly higher than at medium-high fidelity.



**Figure 2 Mean total changes (bold), quality changes (dotted) and expected changes (dashed) across levels of fidelity**

### 6.1 Quality changes

The number of quality changes also showed a significant main effect of fidelity (one-way repeated-measures ANOVA,  $F(4, 116) = 31.763, p < .001$ , partial  $\eta^2 = .48$ ). A significant linear trend was found across mean quality changes at each level of fidelity, ( $F(1, 29) = 76.91, p < .001$ , partial  $\eta^2 = .73$ ), indicating that the number of quality changes made was the highest when participants were presented with the low-fidelity design on paper, and decreased systematically as fidelity increased in the other conditions. No significant quadratic or higher-order trends were found.

Variability between subjects decreased systematically across Conditions 1-5 from low to high fidelity (SD range = 5.5-3.54). Pairwise comparisons (with Bonferroni adjustment for multiple comparisons) showed that the number of quality changes made to the low fidelity design presented on paper was significantly higher than to all other designs ( $p < .05$ ). A significantly higher number of quality changes was made to the low fidelity design presented on the tablet PC than to all other tablet-presented designs except medium-low fidelity (Condition 3) design ( $p < .05$ ).

As with total changes, the number of quality changes made differed significantly between the low fidelity designs presented on paper and on the tablet PC – the mean difference was 2.68.

### 6.2 Expected changes

Similar trends were found for expected changes (correction of deliberately introduced errors). One-way repeated measures ANOVA revealed a significant main effect of fidelity on the number of expected changes made to the designs,  $F(4, 116) = 29.28, p < .001$ , partial  $\eta^2 = .50$ . A significant linear trend in the mean number of expected change was found, ( $F(1, 29) = 92.70, p < .001$ , partial  $\eta^2 = .76$ ), across levels of fidelity. Figure 2 shows that, overall, participants made the most expected changes to the low-fidelity design on paper, and the number of expected changes decreased as fidelity increased. No significant quadratic or higher-order trends were found.

Between subject variability was less for expected changes than for total and quality changes (SD range = 4.24-3.25) and did not decrease systematically across Conditions 1-5 from low to high fidelity (SD range = 5.5-3.54). Pair-wise comparisons (with Bonferroni adjustment for multiple comparisons) revealed that participants made significantly more expected changes to the low fidelity paper design than to all designs presented on the Tablet PC ( $p < .05$ ). The difference in number of expected changes was greater as the level of fidelity of the Tablet PC designs increased; participants made significantly fewer expected changes when they were presented with the high fidelity design, compared to designs with medium-low fidelity, low fidelity on the Tablet PC and low fidelity on paper. However, no significant difference in number of expected changes was found between high fidelity and medium-high fidelity; between medium-high fidelity and medium-low fidelity; or between medium-low fidelity and low fidelity on the Tablet PC. As with total changes and quality changes, the number of expected changes made differed significantly between the low fidelity design on paper and the low fidelity design on the tablet, – the mean difference was 2.37.

### 6.3 Order effects and design equivalence

Although a counterbalanced design was used to control for order of design presentation, it has been suggested that internal validity may yet be reduced by order effects. In

the present study there was some variability observed in individual subjects' performance suggesting that order effects may have played a role in contributing to the combination of linear and cubic trends. Therefore data were grouped according to order of presentation and, as the small sample size rendered statistical analysis inappropriate, the number of each type of change was plotted as a function of level of fidelity. Although there was some evidence of a small effect when conditions were presented in the orders 54321 and 12345, the linear trends reported above were still clear.

Equivalence of the experimental designs used was assessed using ANOVA and post-hoc comparisons to compare the number of total, quality and expected changes across all five conditions. Although differences were found across all levels of fidelity, suggesting a possible confounding effect of differences between designs, the linear effect of fidelity on all types of change nevertheless remains clear and evident in Figure 2.

#### 6.4 Preference for design fidelity and medium

As can be seen in Figure 3, questionnaire data indicated that mean preference for working on designs presented on the PC increased as fidelity increased; however the low fidelity design presented on paper was preferred to that presented on the Tablet PC.

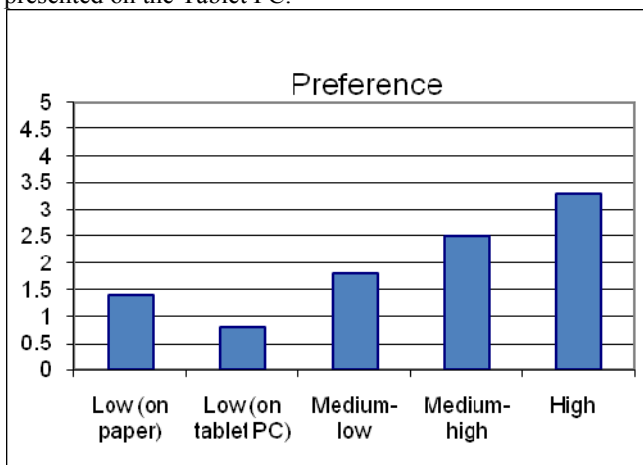


Figure 3 Mean rank and standard deviation of preference for design fidelity.

The majority of participants (17/30, 57%) ranked the high fidelity design as 'most liked' while only 5 participants (17%) ranked it 'least liked'. The majority of participants (20/30, 67%) ranked the low fidelity design on the Tablet PC as 'least liked', while the low fidelity on paper design was ranked as 'least liked' by only five participants (17%) and was most frequently ranked (14/30, 47%) as the second least liked. Thus when working on low fidelity designs, participants generally preferred paper to Tablet PC presentation. It is unlikely that this can be attributed to overall preference for design medium alone, as participants were nearly equally divided in terms of preference for design media during the experiment, with

13 (43%) preferring paper, 15 (50%) preferring the Tablet PC and two (6%) indicating indifference.

Questionnaire data also indicated that preference for design fidelity levels were based on design appearance for the majority of participants (21/30, 70%), on effort required to improve it for 11/30 participants (37%) and on the degree of fun or stimulation it provided for 7/30 participants (23%). Preference in the 'real world' outside the experiment was to use paper in the early stages of design and then switch to using a computer with Photoshop, Visual Basic.Net or other popular applications for 13/30 participants (43%); to use a computer only for 7 participants (23%); the Tablet PC only for 3 participants (10%); paper and pen only for 3 participants (10%); paper then the Tablet PC for one participant (3%) while two participants (6%) expressed no preference.

## 7. DISCUSSION

The visual fidelity or fidelity of user interface designs is known to affect practice and satisfaction when working with the designs[6,10,22]. Visual fidelity is a combination of the smoothness of lines, the relative sizes, alignment and spacing of elements and appearance of writing fonts. In this paper we have identified different factors of fidelity to create a taxonomy of visual fidelity. We made some reasoned decisions about when and how to vary each to increase the visual fidelity of diagrams systematically and steadily.

We then carried out a controlled, counter-balanced experiment with two hand-drawn versions of the design, one on paper and one on a tablet PC, two intermediary partially beautified designs and one fully beautified design, giving five experimental conditions.

Overall, there was an effect of fidelity on the number of changes made: specifically, the number of changes made decreased as the designs' level of fidelity increased, there was a negative relationship between fidelity and design performance. The results suggests that fidelity plays an important role in affecting the subjects' performance, especially on decisions on making changes to improve the designs presented in terms of functionality and usability. The medium of the design (paper or computer) also had an effect on the number of changes; there were significantly more changes to the paper version of the designs. User preference ranked highest on the highest fidelity realization, declining as fidelity decreased. However low-fidelity paper was preferred to the lowest fidelity computer rendering.

There are limitations on these results due to the scope of the study. The participants were a small and heterogeneous sample of students. Experimenter bias was possible in the five forms designed and levels of fidelity created and their intervals. There may have been varying difficulty in improving the five designs (an experimental confound). This last argument can be viewed as one of the uncontrollable limitations in the study, even though all



objective measures were taken to make each design as similar as possible. InkKit, like other sketch tools, was (and is) still in its development stage. There are two particular functions that differ from paper – first, mode changes are required to move between creating and editing ink – equivalent to changing from pencil to eraser; and second, support is provided for selecting and moving/resizing ink that has no paper equivalent. Finally because of the small sample size no statistical analyses of subjective measurement of design tool preference during the experiment was possible.

Other experiments have compared computer-based sketch tools with high-fidelity computer-based tools and low-fidelity tools (paper or whiteboard) [1, 16, 22, 29]. These computer-based sketch tools have included functional support for ‘executing’ the low-fidelity prototype providing support that is not available on paper. This is the first study to isolate visual fidelity on the sketch tool. The general trend of high fidelity resulting in fewer changes is evident in these studies too. However while [22, 29] found no difference in performance between whiteboard/paper and sketch tool, the sketch tool supported functional interaction which could have mitigated the negative effect that we have identified. Some of this effect could be due to lack of Tablet PC experience, however as no order effect was identified it is more likely to be due to interference from the technology.

Our findings on preference are consistent with Coyette et al. [6] where it was found that as fidelity increased across computer construction of a design, so did satisfaction. Our study replicates and extends Coyette et al.’s in that these did not include a paper alternative.

There is incongruity here: users’ preference is in direct contradiction to the most effective presentation. While we found a significant decrease in performance as fidelity increased, users preferred the higher fidelity rendering.

Furthermore, the performance difference between the two hand-drawn designs suggests that computer technology negatively affects the design process. Earlier work hypothesized that it was the visual fidelity that affected the design process, suggesting tasks such as component selection were distracters. This study suggests that even a minimalist computer interface is a distracter. A possible explanation for this could be that the study participants were novice Tablet PC users, however as no order effect was identified it is more likely that the computer environment will continue to be a distracter.

Computer tools offer some clear advantages for editing, storing and transmitting data. Computer supported sketching can also facilitate interactive sketches, while Rettig [26] recommended ‘playing computer’ with paper designs. No rigorous evaluations have been carried out directly to compare interactive computer-based sketches with ‘Wizard of Oz’ techniques in isolation.

## 8. CONCLUSION

This study examined the aspects of visual fidelity to define a taxonomy of design attributes that contribute to design fidelity. We then applied these attributes systematically to simple UI designs to compare the affect of partial beautification to designer reviews. The results show that each increase in fidelity resulted in a decrease in performance measured by changes made to the designs. Additionally computerization of the hand-drawn design adversely affected performance.

These results are consistent with previous studies that compared the end-points of hand-drawn and fully beautified designs. However it is the first study to consider intermediary levels of fidelity on design performance. This study and Coyette et al.’s [6] work measuring users satisfaction have found similar trends with increasing satisfaction as fidelity increases. This contradiction of performance versus satisfaction will continue to challenge teachers and designers alike.

There are many more aspects of computerization of design which warrant careful comparison. Two which directly follow from this study is the affect of intended audience and of functional interaction with the design. Does who the design is for (e.g. self, colleague or client) affect fidelity preference? Is Wizard of Oz functional interaction as effective, or more effective than computer supported interaction of either low or high fidelity designs? There are also continuing challenges to the sketch tools community to provide a computer supported sketching environment that compares favourably with paper for basic drawing.

## 9. REFERENCES

- 1 Bailey, B. P., Konstan, J. A., Are Informal Tools Better? Comparing DEMAIS, Pencil and Paper, and Authorware for Early Multimedia Design, in *proc CHI 2003*, ACM, (2003), 313-320
- 2 Black, A., Visible planning on paper and on screen: The impact of working medium on decision-making by novice graphic designers, *Behaviour and information technology*, 9, 4, (1990), 283-296
- 3 Bolz, D., Some aspects of the user interface of knowledge based beautifier for drawings, in *proc Intelligent user interfaces '93*, (1993), 45-52
- 4 Boutsen, L., Humphreys, G. W., Axis-alignment affects perceptual grouping: Evidence from simultanagnosia, *Cognitive Neuropsychology*, 17, 7, (1999), 655-672
- 5 Brink, T., Gergle, D., Wood, S. D., Design web site that works: Usability for the web, Morgan Kaufmann Publishers, (2002),

- 6 Coyette, A., Kieffer, S., Vanderdonckt, J., Multi-Fidelity Prototyping of User Interfaces, in *proc Interact*, (2007), 75-82
- 7 Damm, C. H., Hansen, K. M., Thomsen, M., Tyrsted, M., Supporting Several Levels of Restriction in the UML, in *proc UML 2000*, (2000), 518-525
- 8 Fowler, S., Stanwick, V., *Web application design handbook: Best practices for web-based software*, Morgan Kaufmann Publishers, (2004),
- 9 Gibson, J. J., *The perception of the visual world*, Houghton Mifflin, (1950),
- 10 Goel, V., *Sketches of thought*, The MIT Press, (1995),
- 11 Goldschmidt, G., Serial sketching: Visual problem solving in design, *The backtalk of self-generated sketches*, *Cybernetics and Systems*, 23, (1992), 191-219
- 12 Gross, M., Do, E. Y. L., Ambiguous intentions: a paper-like interface for creative design, in *proc UIST '96*, ACM, (1996), 183-192
- 13 Igarashi, T., Matsuoka, S., Kawachiya, S., Tanaka, H., Interactive Beautification: a technique for rapid geometric design, in *proc UIST 97*, (1997), 105-114
- 14 Landay, J., Interactive sketching for user interface design, in *proc Chi '95 Mosaic of Creativity*, Doctoral Consortium, (1995), 63-64
- 15 Locher, P. J., Stappers, P. J., Overbeeke, K., The role of balance as an organizing design principle underlying adults' compositional strategies for creating visual displays, *Acta Psychologica*, 99, 2, (1998), 141-161
- 16 McCurdy, M., Connors, C., Pyrzak, G., Kanefsky, B., Vera, A., Breaking the Fidelity Barrier: An Examination of our Current Characterization of Prototypes and an Example of a Mixed-Fidelity Success., *CHI 2006*, ACM, (2006),
- 17 Microsoft Corporation, My Font Tool for Tablet PC, <http://www.microsoft.com/windowsxp/downloads/powertoys/tabletpc.mspix>, accessed on 20 May 2008
- 18 Newman, M. W., Landay, J. A., Sitemaps, storyboards and specifications, in *proc Designing interactive systems: processes, practices, methods and techniques*, ACM, (2000), 263-274
- 19 Newman, M. W., Lin, J., Hong, J. I., Landay, J. A., DENIM: An Informal Web Site Design Tool Inspired by Observations of Practice, *Human-Computer Interaction*, 18, 3, (2003), 259-324
- 20 Plimmer, B., Freeman, I., A Toolkit Approach to Sketched Diagram Recognition, in *proc HCI, eWiC*, (2007), 205-213
- 21 Plimmer, B. E., Using Shared Displays to Support Group Design; A Study of the Use of Informal User Interface Designs when Learning to Program, *Computer Science*, University of Waikato, (2004),
- 22 Plimmer, B. E., Apperley, M., Evaluating a Sketch Environment for Novice Programmers, in *proc SIGCHI*, ACM, (2003), 1018-1019
- 23 Plimmer, B. E., Apperley, M., INTERACTING with sketched interface designs: an evaluation study., in *proc SigChi 2004*, ACM, (2004), 1337-1340
- 24 Plimmer, B. E., Apperley, M., Software for Students to Sketch Interface Designs, in *proc Interact*, (2003), 73-80
- 25 Pomm, C., Werlen, S., Smooth morphing of handwritten text, in *proc AVI '04*, ACM, (2004), 328-335
- 26 Rettig, M., Prototyping for tiny fingers, *Communications of the ACM*, 37, 4, (1994), 21-27
- 27 Rock, I., In defence of unconscious inference, Stability and constancy in visual perception: Mechanisms and process, Wiley, (1977), 321-74
- 28 Shelly, G. B., Cashman, T. J., Woods, D. M., *HTML: Complete concepts and techniques* (Third Ed.), Shelly Cashman Series, (2005),
- 29 Walker, M., Takayama, L., Landay, J. A., High-fidelity or low-fidelity, paper or computer medium?, in *proc Human Factors and Ergonomics Society*, (2002), 661-665
- 30 Wong, Y. Y., Rough and ready prototypes: Lessons from graphic design, in *proc Human Factors in Computing Systems CHI '92*, (1992), 83-84