CS6750 Assignment M5 (Summer 2021)

Titus Choi
tchoi46@gatech.edu

***Abstract*—** The task of using text messages as means of communication has become more ubiquitous in recent years. While there are various benefits to text messages, there are also pitfalls in comparison to verbal communication methods. This project explores and investigates the task of sending a message on the native iOS message application, with the goal of redesigning the interface to address a common but often detrimental scenario: *"accidental message sends"*. Collectively, the project will go through the process of defining the problem space, identifying the user types, planning & executing the needfinding plans, gathering requirements, brainstorming, creating prototypes, and performing evaluations.

1 QUALITATIVE EVALUATION

1.1 Evaluation report

The evaluation method of *survey* was utilized to evaluate the *wireframe* (*Figure 1*) previously generated (*Assignment M3*). The *survey* was conducted using the *PeerSurvey* tool (King, 2021) with twenty-five CS6750 classmates as participants, with each of the survey sessions being run identically. On that note, there were some organizational aspects of the evaluation process that could have been conducted differently: (1) Recruitment of participants outside CS6750 classmates could provide a more diverse set of data. (2) The wireframes could be separated into an individual screen to gather more data on specific components of each screen, rather than providing a singular image with all of the wireframe transitions.

1.2 Raw results

Following is a summary of the raw *survey* results: (1) The twenty-five participants covered all user types excluding the 65+ *age group*. (2) All of the participants were current users of the iMessage app. (3) The ability to avoid accidental messages in iMessage was viewed anywhere from "very easy" to "very difficult", with "somewhat easy" being the most common response. (4) Providing preventative measures for accidental messages in iMessage was view anywhere from

"very useless" to "very useful", with "somewhat useful" being the most common response. (5) All of the participants were able to identify at least one difference in the redesigned interface, with the most immediately identified differences being the "colored recipient indicator. (6) The "colored recipient indicator" and the "recipient group identifier" were generally considered to help aid in identifying recipients easier. (7) The "message submenu" and the associated "message delete functionality" were generally considered to be a useful feature. (8) The additional features were considered to be "easy to learn", "easy to understand", and "satisfactory". And lastly, (9) Some notable remarks were provided such as: "coloring could be different for colorblind", that provided additional insights on potential ways that the *wireframe* could be improved in future iterations. The unabridged raw results are provided in the *Appendix*.

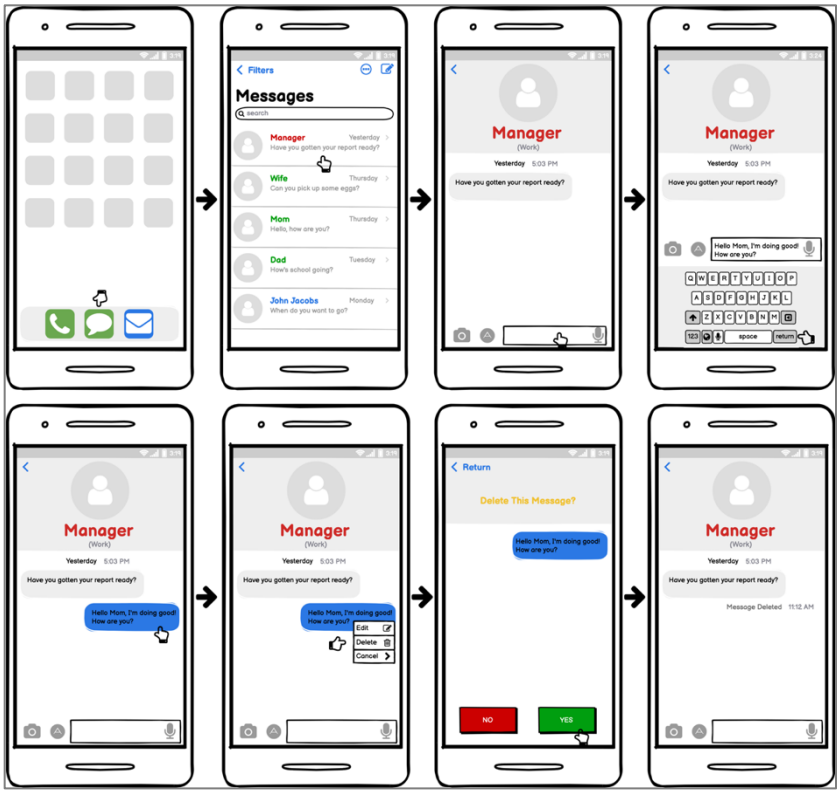

Figure 1 — Wireframe of Redesigned iMessage Interface.

1.3 Feedback analysis

The main takeaways for improving the interface gathered were as follows: (1) While 56% of the participants answered either "somewhat easy" or "very easy" regarding the ability to avoid accidental messages on the current iMessage

interface, 76% of the same participants viewed accidental message preventative measures to be either "somewhat useful" or "very useful" if it were included in the interface, indicating that while greater than half of the participants seemed confident in their ability to avoid accidental messages, a majority of them seemed to indicate a lack in the current interface in terms of accidental message avoidance. (2) The new interface design was close enough to the currently existing interface so that it leveraged the participant's *affordances* (Joyner, 2021a) but was distinct enough so that they were able to identify the new features without assistance", showing both a high level of *learnability* and *memorability* (Joyner, 2021e). This was shown as a majority of the participants (96%) were able to distinguish the differences of the redesign without being provided the current interface, with 88% of the participants answering four or five on the Likert scale in terms of the additional features being "easy to learn" and 84% respectively for the features being "easy to understand". And lastly, (3) A majority of the participants (84%) were either "satisfied" or "highly satisfied" regarding the changes in the iMessage interface, also indicating a high level of *satisfaction* in the redesign.

One surprising factor from the analysis was that the change in the interface that the participants were predominantly able to identify first was the "colored recipient indicator" at 64%, equally matching the identifiability of the "message delete functionality". Noting that, this indicates that the efficacy of the preventative measures put in place was notably effective in drawing attention to the recipient of the message, which has a high potential of mitigating accidental messages that are caused by incorrect selection of intended recipient(s).

Lastly, while the general trends shown in the responses were as initially expected (e.g., perceived learnability, affordance-based identification), the satisfaction level of the redesign was higher than anticipated at 84%. This showed that the general perception of the redesign was overwhelmingly positive, while also addressing the core goals of the interface: to preventing accidental messages, and to provide the ability to address the accidental messages when they do occur.

1.4 Changes suggested

While the analysis of the *survey* results showed a higher *satisfaction* level than anticipated, there were some notable feedbacks that provided valuable ways the guide the next set of changes in the prototype. The first feedback was potentially adding a "face recognition function" to address accidental messages caused by

non-owners of the device (e.g., babies playing with their parent's device). The second feedback that was particularly insightful was differentiating the color of the "colored recipient identifiers" to address colorblind users. This was an insight that was not considered when creating the redesign but is a valid concern that should be addressed. Noting that, more research will be needed to determine the best method (e.g., colorblindness setting) in addressing this concern.

2 EMPIRICAL EVALUATION

2.1 Evaluation report

The *within-subjects* evaluation method was utilized to evaluate whether the *alternative hypothesis* of "the time and accuracy in which the participants identify the recipient is faster in the redesigned interface" was supported. First, the participants were shown a glimpse (1 second per treatment) of the two interfaces designs prototyped using the Balsamiq wireframing tool (Balsamiq, 2008), each with a randomized recipient identifier from a list of ten names. Then the participants were immediately asked to answer the question of "What was the recipient's name in the image shown?", as the duration and the accuracy in which they wrote down their responses were documented. As a note, the automatic transition between the image and the question was set to one second using the PowerPoint slideshow feature, and the sequence in which the two interfaces were shown was randomized per participant to minimize potential extraneous variables. A total of 10 participants consisting of co-workers and family members were recruited for this experiment, with each session being conducted via WebEx. Overall, the testing process went as expected without encountering any notable deviations. Noting that, one thing that could be done differently next time is utilizing greater variations in the sizing of the recipient identifiers to determine the optimal recipient identifier size in comparison to the screen real estate. Lastly, there weren't any notable events or occurrences that would call the results into question in the preliminary analysis.

2.2 Statistical test report

The raw results for the *within-subjects* evaluation are as follows with the rows indicating the control number, and the columns indicating the correct recipient's name, and the two interfaces with the participant's responses and duration:

Table 1 — Raw results from the *within-subjects* evaluation. Incorrect responses are labeled with an "(I)".

Control #	**Recipient Name** (Control)	**Current Interface** (Response & Duration in Seconds)		**Redesigned Interface** (Response & Duration in Seconds)	
1	Kastor Europa	Kaston Europe (I)	75	Kastor Europa	23
2	Jeanne Fredenand	Jeane Ferdinand (I)	63	Jeanne Frednand (I)	34
3	Karl Juventas	Karl Juve… (I)	91	Karl Juventos (I)	38
4	Cosme Frieda	Cosme Frida (I)	51	Comse Frieda	36
5	Olivia Faina	Olivia Faina	32	Olivia Faina	17
6	Manu Cicero	Manu Cicero	47	Manu Cicero	23
7	Alycia Ridge	Alicia Ridge (I)	85	Alycia Ridge	31
8	Harvey Warren	Harvey Warrer (I)	77	Harvey Warren	48
9	Pauline Wilton	Pauline Winston (I)	66	Pauline Wilton	50
10	Katie Teddy	Katie Teddy	52	Katie Teddy	34

As planned previously (*Assignment M4*), the *interval/ratio* data was gathered from the *statistical test* where: (1) the *independent variable* was the text size of the recipient indicators, (2) the *dependent variable* was the average of the outcomes, (3) the *null hypothesis* was that the outcomes are equal, and lastly (4) the *alternative hypothesis* was that the outcomes are unequal. As noted above, the evaluation consisted of 10 participants consisting of co-workers and family members. Noting that, the Microsoft Excel Data Analysis tool assuming *unequal variance* was to calculate the *N values*, *means*, *standard deviations*, *variances*, *p-value*, and the *t-statistic value* for the *student's t-tests* (Joyner, 2021h)*:*

Table 2 — Statistical analysis result for response time.

Designs	**N**	**Mean** (Duration)	**Standard Deviation**	**Variance**	**P-Value** (Two-Tail)	**T-Statistic**	**T-Critical** (Two-Tail)
Current Interface	10	63.9	18.532	343.433	0.0002405728	4.51924871066	2.14478668791
Redesigned Interface	10	33.4	10.585	112.044			

Based on the results above (*Table 2*), as the *p-value* of 0.0002405728 is less than or equal to the *alpha value* of 0.05, this indicates a statistical significance between the

response times between the current and redesigned interfaces, *rejecting* the *null hypothesis* that the outcomes are equal between the two interfaces.

Table 3 — Statistical analysis result for response accuracy.

Designs	**N**	**Mean** (Accuracy)	**Standard Deviation**	**Variance**	**P-Value** (Two-Tail)	**T-Statistic**	**T-Critical** (Two-Tail)
Current Interface	10	0.3	0.48304589	0.2333333	0.0239389198	-2.4659848095	2.10092204024
Redesigned Interface	10	0.8	0.42163702	0.1777778			

Furthermore, the textual responses gathered from the participants (*Table 1*) indicated that there were also notable differences between the accuracy of the recipient's name provided. To perform the *statistical analysis* on the textual response data, the *correct responses* were represented by "1"s and the *incorrect responses* were represented by "0"s. Based on the results above *(Table 3*), as the *p-value* of 0.0239389198 is less than or equal to the *alpha value of* 0.05, this also indicates a statistical significance between the *accuracy* of the response between the current and redesigned interfaces. This finding also *rejects* the *null hypothesis* that the outcomes are equal between the two interfaces.

2.3 Feedback analysis

The *statistical test* results above (*Section 2.2*) matched the initial expectation that the increased size of the recipient identifiers will not only decrease the duration in which a user takes to identify the recipient but also increase the accuracy of the identification of the recipient. This expectation and the prototype utilized for this experiment was based on the data gathered from the *needfindin*g method of *interview* performed previously (*Assignment M2*) that suggested a "larger profile picture" and/or "larger text size" of the recipient's name would potentially prevent accidental messages caused by misidentification of the recipients. Noting that, as the potential *lurking variables* previously identified (*Assignment M4*) were controlled by ensuring that the only differences in the two interface designs were the text size of the recipient identifiers, I believe the results are based on real interface differences, and not due to *lurking variables* or *experimental errors*.

2.4 Suggested changes

Based on the results of the *empirical evaluation*, the redesigned interface prototype is very effective in ensuring that accidental messages caused by recipient

misidentification are addressed through the increase in the size of the recipient identifiers. Noting that, no notable changes seem to be warranted to the redesigned interface from this *empirical evaluation* results, and the redesigned interface should be carried forward to the next round to prototyping as is, excluding any changes warranted from the *qualitative evaluation* results above. Noting that, one change that can be explored is testing various text sizes of the recipient identifiers to determine the "sweet spot" in comparison to the screen real estate.

3 EVALUATION SUMMARY

The results from the *qualitative* and *empirical* evaluations each indicate different areas of focus in the next iteration through the design life cycle. First, the results from the *qualitative evaluation* introduced the potential changes of: (1) adding a "face recognition function" to address accidental messages caused by non-owners of the device (e.g., babies playing with their parent's device), and (2) adding a "colorblindness function" to address the colorblind users that might not be able to distinguish the colored recipient identifiers in the redesign. Second, while the results from the *empirical evaluation* did not warrant specific changes, differing increases in text sizes can be tested as noted above (*Section 2.4*) to determine the optimal recipient identifier size to screen size ratio. The two plans for the next iterations through the design life cycles are explored in more detail below:

3.1 Description of additional understanding desired

Based on the results of the *qualitative evaluation* (Joyner, 2021f), there are a few questions about the user that could be investigated through additional *needfinding* (Joyner, 2021b) exercises. First, for the "face recognition function" mentioned above, the questions that can be asked are: (1) What percentage of users utilize facial recognition on their iOS devices, (2) What percentage of iOS device users have devices that support facial recognition functionality, (3) Would the iMessage users actually utilize facial recognition functionality, (4) Do users find facial recognition based message send locking feature useful, and (5) Do users find facial recognition based message send locking feature easy to learn. Second, for the "colorblindness function" mentioned above, the questions that can be asked are: (1) What percentage of iMessage users have colorblindness, (2) What type of colorblindness is represented in the iMessage user population, (3) What are some methods utilized in the industry that addresses these concerns. For both of these

functionalities, the *needfindings* of *surveys, interviews,* and *evaluation of existing user interfaces* can be conducted to gather more understanding about the user.

3.2 Evaluation of additional design alternatives

Based on the insights from the *empirical evaluation* (Joyner, 2021g), there are a series of additional design alternatives that might be explored in a second iteration through the design life cycle. As noted above (*Section 2.4*), design alternatives with differing text sizes of the recipient identifiers can be explored to determine the "sweet spot" in comparison to the screen real estate. While the *empirical evaluation* results showed a statistical significance between the increases in the text size of the recipient identifier to its impact on the accuracy and duration (*Section 2.2*), it did not determine the optimal text size that would provide the benefit without negatively impacting the overall usability of the application. Noting that, additional design alternatives that utilize varying text sizes can be generated to determine the optimal design to move forward with.

3.3 Brainstorming prototype revisions

While the results of both the *qualitative* and *empirical* evaluation were generally positive, the additional understanding desired (*Section 3.1*) and the additional design alternatives outlined (*Section 3.2*) would need to be iterated further before moving on to the next level of fidelity. Noting that, the next iteration through the design life cycle would stay in the fidelity levels of *textual prototype*, *verbal prototype*, and *wireframe* (Joyner, 2021c & 2021d) until no additional significant insights are found. And once enough data is gathered, potential higher fidelity levels such as *digital prototypes* (Babich, 2017) can be generated to bring the redesigned interface closer to launch.

3.4 Planning next evaluation type to employ

As the next iteration through the design life cycle would result in redesigned *textual prototype*, *verbal prototype*, and *wireframe*, the *qualitative* and *empirical* evaluations would be employed again as the combination of *qualitative* and *empirical* evaluations result in a robust set of data that provides generalizable conclusions based on provable advantages. On that note, since more iteration are planned at the current time as noted above (*Section 3.1 & 3.2)*, the current prototype is not yet ready for rigor of a true empirical evaluation and will need to be conducted once no additional significant insights are found from the next round of iteration.

4 REFERENCES

1. Babich, N. (2017). Prototyping 101: The Difference between Low-Fidelity and High-Fidelity Prototypes and When to Use Each. Retrieved July 22, 2021, from https://blog.adobe.com/en/publish/2017/11/29/prototyping-difference-low-fidelity-high-fidelity-prototypes-use.html.
2. Balsamiq. (2008). Balsamiq. Rapid, Effective and Fun Wireframing Software: Balsamiq. Retrieved June 26, 2021, from https://balsamiq.com/
3. Joyner, D. A. (2021a). 2.5: Design Principles and Heuristics – 6. Affordances [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14034/slides/70792
4. Joyner, D. A. (2021b). 3.3: Needfinding and Requirements Gathering – 16. Surveys [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14030/slides/70692.
5. Joyner, D. A. (2021c). 3.5: Prototyping – 5. Verbal Prototypes [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14036/slides/70855
6. Joyner, D. A. (2021d). 3.5: Prototyping – 8. Wireframing [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14036/slides/70858
7. Joyner, D. A. (2021e). 3.6: Evaluation – 4. 5 Tips: What to Evaluate [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14039/slides/70915.
8. Joyner, D. A. (2021f). 3.6: Evaluation – 7. Qualitative Evaluation [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14039/slides/70918
9. Joyner, D. A. (2021g). 3.6: Evaluation – 11. Empirical Evaluation [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14039/slides/70922
10. Joyner, D. A. (2021h). 3.6: Evaluation – 14. Quantitative Data and Empirical Tests [video file]. Retrieved from https://edstem.org/us/courses/5912/lessons/14039/slides/70925
11. King, B. (2021). PeerSurvey. Retrieved June 13, 2021, from http://peersurvey.cc.gatech.edu/

5 APPENDIX

The following are the *raw* results from the *qualitative evaluation* conducted. This *raw* data is solely intended to be referenced by the analyses above in *Sections 1*. The *raw data* result of the survey responses is comprised of responses provided by a total of twenty-five participants through the *PeerSurvey* tool. Because of the length of the data, the results below are displayed in the original CSV format.

response,Q1,Q2,Q3,Q4,Q5,Q6,Q7,Q8,Q9,Q10,Q11,Q12,Q13,Q14,Q15,Q16,Q17

1,18 - 29,Yes,Very easy,Somewhat useful,color coding.,color coding.,Colored recipient indicator (name),Ok,3,2,1,3,3,3,Neutral,I don't agree with the functionality of delete. Currently the delete functionality removes the message from your phone but is not and should not be capable of delete off someone elses phone.,This functionality of delete already exists as well as the submenu when you fill in a contacts company.

2,40 - 49,Yes,Somewhat easy,Somewhat useful,withdraw a sent message,"letting him/her know ""message delete""",Message delete functionality,Ok,3,3,4,4,5,5,Satisfied,I think it's no big deal.,"Maybe add a face recognition function. If identifying the user not being the owner, block the message function. Moms tend to let their babies to play with their phones, but your design does not address messages accidentially sent by a baby. "

3,50 - 64,Yes,Very difficult,Very useful,"I noticed the option to edit, delete and cancel message. ",The color schema to personalize the name,Colored recipient indicator (name);Recipient group identifier;Message submenu;Message delete functionality;All of the above,Ok,5,5,5,5,5,5,Highly satisfed,,

4,18 - 29,Yes,Somewhat easy,Somewhat useful,I have not seen the delete message feature before,Colored contact names,Colored recipient indicator (name);Message delete functionality,Ok,5,3,3,5,5,4,Satisfied,,

5,18 - 29,Yes,Somewhat easy,Somewhat useful,The extra large work contact,Deleting messages,Colored recipient indicator (name);Recipient group identifier,Ok,3,4,5,5,5,5,Satisfied,,Coloring could be different for colorblind

6,30 - 39,Yes,Somewhat difficult,Somewhat useful,The Contact info at the top is bigger. ,I see an edit option for an already sent message,Colored recipient indicator (name);Message submenu;Message delete functionality,Ok,4,4,5,5,5,5,Satisfied,,

7,18 - 29,Yes,Somewhat easy,Somewhat useful,Being able to cancel or edit a message,"The Manager is tagged with work, also their name is highlighted red",Message submenu;Message delete functionality,Ok,5,4,4,4,5,5,Satisfied,,

8,18 - 29,Yes,Somewhat easy,Somewhat useful,"New colors, Deleting message",None,Colored recipient indicator (name);Message delete functionality,Ok,4,3,4,5,4,4,Satisfied,,

9,18 - 29,Yes,Neither easy nor difficult,Somewhat useful,I'm not sure you can edit messages in the current interface.,None,Colored recipient indicator (name);Message submenu;Message delete functionality,Ok,5,3,5,5,5,5,Highly satisfed,,

10,18 - 29,Yes,Neither easy nor difficult,Very useful,the yes and no buttons,None,Colored recipient indicator (name);Recipient group identifier,Ok,4,4,4,4,4,4,Satisfied,,

11,30 - 39,Yes,Very difficult,Very useful,delete message,none,Colored recipient indicator (name),Ok,5,5,5,5,5,5,Highly satisfed,,

12,18 - 29,Yes,Somewhat easy,Somewhat useful,The abiity to delete numbers,Larger manager name,Colored recipient indicator (name);Recipient group identifier;Message delete functionality,Ok,4,2,3,5,4,3,Satisfied,,

13,40 - 49,Yes,Somewhat difficult,Somewhat useless,the manager in red. They confirm button.,the manager in red. The confirm button.,Colored recipient indicator (name),Ok,4,4,3,4,4,4,Satisfied,,

14,18 - 29,Yes,Neither easy nor difficult,Neither useful nor useless,I noticed the delete option. Never knew it existed.,The delete function. The yes and no. Not sure about the message deleted.,Colored recipient indicator (name);Message delete functionality,Ok,5,4,4,4,3,3,Satisfied,,

15,18 - 29,Yes,Somewhat easy,Somewhat useful,"ability to edit, or cancel a message after it's been sent","confirmation of ""message deleted"" with its time stamp",Message submenu;Message delete functionality,Ok,4,4,4,4,4,4,Satisfied,,

16,30 - 39,Yes,Neither easy nor difficult,Neither useful nor useless,Nothing special,None,None of the above,Ok,2,2,3,4,4,4,Dissatisfied,I like the current design. ,Wireframe is hard to compare. Need real interface to compare.

17,40 - 49,Yes,Somewhat difficult,Somewhat useless,"I don't have the current interface in front of me (I don't use it that often), so I can't say for sure, but I believe the Edit, Delete, and Cancel buttons are different",None,Colored recipient indicator (name);Message submenu;Message delete functionality,Ok,4,4,5,4,4,4,Highly satisfed,,

18,30 - 39,Yes,Somewhat difficult,Very useful,Option to delete and edit.,Highlight of names in different colors.,All of the above,Ok,5,5,5,5,5,5,Highly satisfed,,

19,18 - 29,Yes,Somewhat easy,Very useless,i see that there is an edit/delete/cancel dropdown list. this doesnt exist in the current interface as far as i know,n/a ,Message delete functionality,Ok,2,2,5,5,5,5,Satisfied,,

20,30 - 39,Yes,Somewhat easy,Neither useful nor useless,the ability to delete messages with a delete button,None,All of the above,Ok,4,4,4,4,3,3,Neutral,how do I know if the message was actually deleted to the sender? ,"If I delete on my phone, does it delete on the recipients phone?"

21,40 - 49,Yes,Somewhat easy,Somewhat useful,Yes ,Yes ,Message submenu;Message delete functionality,Ok,3,4,4,5,4,4,Satisfied,NA,NA

22,18 - 29,Yes,Somewhat easy,Somewhat useful,The color name,"The (Work) tag, the ability to edit message, the huge bar on top, different appearance for deleting message (which I'm assuming would delete the one they received too)",Colored recipient indicator (name);Recipient group identifier;Message submenu;Message delete functionality;All of the above,Ok,4,2,5,5,5,5,Neutral,"The name and picture are way larger than they are in the existing interface, which reduces the space that is used for viewing the actual texts",

23,30 - 39,Yes,Somewhat easy,Somewhat useful,Can edit a message i've already sent rather than just delete it,The format of deleting messages is different (two buttons),Message delete functionality,Ok,4,4,4,4,4,4,Satisfied,,

24,18 - 29,Yes,Somewhat easy,Somewhat useful,"The colors, but other than that it looks really similar",None,Colored recipient indicator (name),Ok,3,4,4,4,4,4,Satisfied,,

25,30 - 39,Yes,Neither easy nor difficult,Very useful,The ability to delete an accidentally sent text,"The colors of different categories of people - e.g., red for people you definitely don't want to send an accidental text to",Colored recipient indicator (name);Message submenu;Message delete functionality,Ok,4,4,5,5,4,4,Satisfied,,"An accidental text sent to a family member or friend could still be very bad (e.g., if you happen to send a text to your mom complaining about your mom that you meant to send to your sister). Just to say I'm not sure the color coding is always accurate"