Audio Engineering Society
Convention Paper 9510

Presented at the 140th Convention
2016 June 4–7 Paris, France

Visual Information Search in Digital Audio Workstations.

Josh Mycroft, Tony Stockman and Joshua. D. Reiss

Centre for Digital Music, Queen Mary, University of London
j. b. mycroft@qmul.ac.uk

ABSTRACT

As the amount of visual information within Digital Audio Workstations increases, the interface potentially becomes more cluttered and time consuming to navigate. The increased graphical information may tax available display space requirements and potentially overload visual perceptual and attentional bandwidth. This study investigates the extent to which Dynamic Query filters (sliders, buttons and other filters) can be used in audio mixing interfaces to improve both visual search times and concurrent critical listening tasks (identifying subtle attenuation of named instruments in a multi-channel mix). The results of the study suggest that the inclusion of Dynamic Query filters results in a higher amount of correctly completed visual and aural tasks.

1. BACKGROUND

As the amount of visual information in Digital Audio Workstations (DAWs) increases, the interface potentially becomes more cluttered and time consuming to navigate [1]. As a result, effective use of the interface risks becoming compromised, as it requires increased cognitive load to navigate and analyse [2]. Increased visual information on screen may also distract users from creative engagement with mixing; research by Duigan et al [3] found that many of the producers he interviewed considered the interface clutter distracting and looked for ways to minimise it so that they could focus on the mix more fully, 'untainted by edits and track organisation' (ibid. p. 168). Reducing the amount of visual information on the screen may also be better suited to the perceptual limits which are limited in capacity with 'only a few items attended at any one time and only a few properties of those items'. [4, page 6].

Typically, audio mixing workflow requires the user to switch focus from one task (such as equalisation, panning, effects setting etc.) to another in a frequent and largely iterative way [5]. In current DAW design additional channels can be displayed by scrolling navigation while additional mix information can be displayed or hidden using tiled and floating windows (e.g. equalisers, effects etc.). However, this may tax available display space requirements, obscure other (useful) screen information and potentially overload visual, perceptual and attentional bandwidth [4]. Furthermore, managing multiple windows may reach 'frustrating and counterproductive new levels [6, page 1].

In other domains, such as maps and websites, Dynamic

Query (DQ) filters are frequently used. These are User Interface (UI) objects (sliders, buttons and other filters) that facilitate rapid exploration of interfaces by real time visual display of query formulation and results [7]. By incrementally adjusting a DQ filter, users can rapidly explore and filter the information while continuously viewing the changing results. There is evidence that DQ filters may enable users to cope with information overload [8] and help keep attention focused on the primary search goal rather than being distracted by the UI [9].

While there is little use of DQ filters with DAWs, they may have the potential to ameliorate visual clutter and make displayed information more germane to a variety of different mixing requirements. For example, by allowing the user to rapidly reveal data, DQs may allow the user to discover which sections of a mixing interface is densely or sparsely populated so they can see 'where there are clusters, exceptions, gaps and outliers' [10, P.239]. This may be directly beneficial to mixing where it is useful to be able to display, for example, channels which contain a particular volume level, which are panned to a certain position within the stereo field or which have certain effects applied to them etc.

For this study DQ filters were applied both to traditional channel strip designs as well as a stage design (where numbered circles represent the channels, their x-axis represents pan position, and their y-axis represents volume). Previous work by the authors [11] has shown that this design by presenting the mix channels as an overview, can significantly improve comprehension of mix elements and their relationship to one another. The study aims to investigate whether the inclusions of DQ filters can further enhance visual search in both the stage and channel strip designs of mixing interfaces.

1.1. Participants

Thirteen participants were selected for this study, comprised of staff and students on a two-year music technology course at City and Islington College, London. All participants had at least one year's experience mixing on DAWs (with a minimum of five hours a week exposure to DAWs and mixing). Participants were 10 male, 3 female aged 17-43.

1.2. Visual Task

Three interface designs of a 24-channel mixer showing volume and pan-position were designed using Max/MSP. For all interface designs the pan and volume had a range of 12 values. For each of the three designs a version with and without dynamic query filters was included (creating six interfaces in total) so that the influence of DQ filters could be analysed for each design.

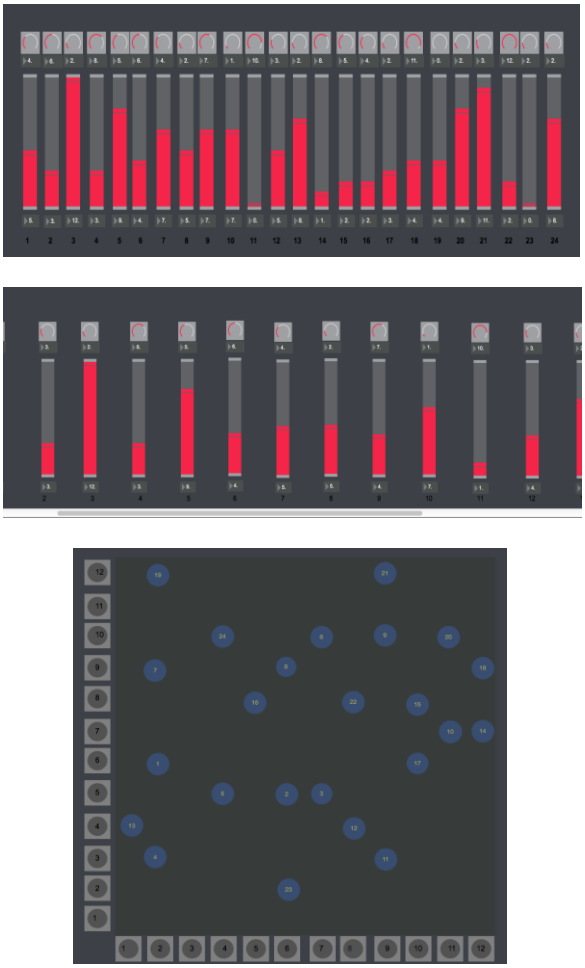

Figure 1. a) top; the mixer design with no scrolling. b) middle; the mixer design requiring scrolling navigation to view all the 24 channels. c) bottom; the stage mixer design, the numbered circles represent channels, the x-axis represents panning and the y-axis represents volume.

The designs consisted of a channel strip design with all 24 channels shown on a single page without the need to navigate (figure 1a), a channel strip mixer where scrolling navigation is required to view all 24 channels (figure 1b) and a stage design mixer presented on one page without the need to navigate (fig 1c**).** In the case of the DQ versions, the DQ filters allowed the users to query the pan position, the volume and individual channels. In the case of the stage mixer, pan position is queried by selecting the numbers on the x-axis, and volume queried by selecting numbers on the y-axis, individual channels are highlighted by clicking the numbers at the top of the screen. For the channel strip mixer designs, pan is queried using the horizontal sliders, volume queried using the vertical sliders, and individual channels selected by clicking on the channels strip numbers. The relevant channels are highlighted in the mixer displays (figure 2).

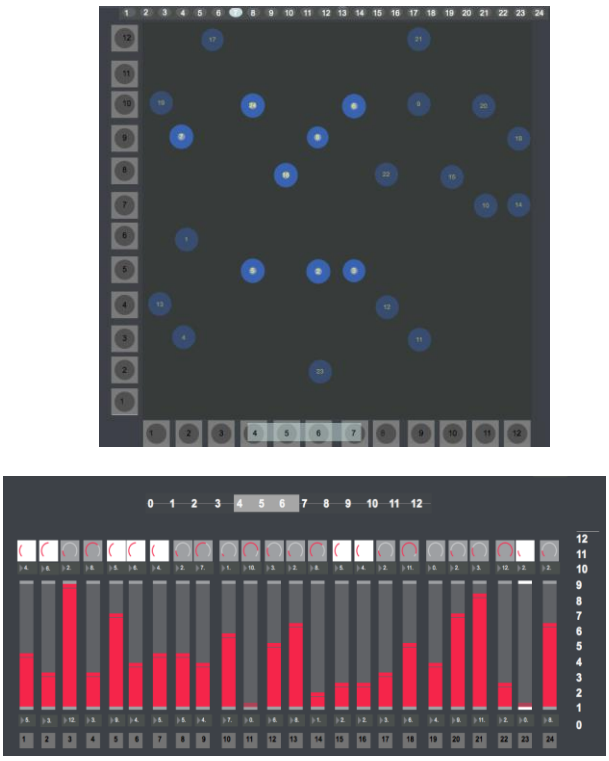

Figure 2. The stage mixer design and channel strip mixer designs with DQ functionality. The selected range highlights the relevant channels. In the stage mixer, channels panned between 4 and 7 and channel 7 are selected. In the channel strip mixer channels panned between 4 and 6 are selected.

For each interface a series of questions about the visual display was included on the screen (see table 1). Participants were required to search for the relevant information and select the correct answer from a drop down menu. When one question was answered the next would appear. There were six questions per interface, designed to test visual referencing of volume, pan and channel positions in a way that is typical of mixing workflow [12].

Participants were asked to answer as many questions as they could in the 45 seconds that the excerpt played, and as soon as the audio had finished the interface was automatically closed. Each question was asked in each interface design with the question order randomised for each participant.

1.3. Listening Task

The listening task was designed to assess whether DQ filters, by reducing visual search allowed greater resources to be given to the aural modalities, thereby increasing aural acuity [13]. The participants were played a twelve-channel audio mix (duration 45 seconds, created using Apple Loops from Logic Pro 9 and imported as 16 bit/ 44.1 KHz audio files into Max/MSP) at the same time as undertaking the visual search tasks. Each time the excerpt was played three of the instruments within the mix (namely backing vocal, snare and tambourine) were randomly attenuated by 6dB. This gain increment was chosen as it is considered an easily discernable reduction in volume [14].

The instrument attenuated in each trial was pseudo-randomised with the condition that each instrument was turned down twice for each participant (so that a direct comparison could be made between the interface designs). The point in the excerpt at which the attenuation was applied was also randomised for each participant. As soon as the excerpt had finished playing the interface which the participants were using was automatically closed and they were asked to select which instrument had been attenuated from a drop down menu with the categories; *backing vocals, snare, tambourine or couldn't tell* (this last option was included to avoid participants guessing the answer if they were unsure).

Q.1) Which channel is loudest, 3,13 or 23?
Q.2) How many channels are panned between 2 and 4?
Q.3) How many channels have volume between 11 and 12?
Q.4) Is the volume of channel ten between 1 and 3?
Q.5) What is the loudest channel panned between 1 and 3?
Q.6) What is the difference in volume between channels 3 and 7?

Table 1. Visual search questions asked per interface design.

1.4. Study Procedure

Before the study began, participants were given an opportunity to use the software and familiarise themselves with all six interface designs. Participants were also given a screening test to see if they could hear the attenuation of the specified instruments (this was done without any concurrent visual task). Participants who could not identify the attenuation would not have their results included in the study. Participants were asked to rate how easily they could hear the attenuation on a five point Likert scale (*very easy*, *easy*, *hard*, *very hard, couldn't hear*). All participants chose either very easy or easy for all three instruments suggesting that discerning audio attenuation at -6dB was well within their capabilities when there was no simultaneous visual task to conduct.

Immediately after the test a survey was given to evaluate the participant's subjective views on task completion using the various interface designs both with and without dynamic queries. Questions and results can be seen in section 2.3.

2. ANALYSIS AND RESULTS

The data for the thirteen participants was analysed for three main criteria; the amount of correctly answered visual questions, the amount of correctly identified file attenuations and an evaluation of the post-study survey.

2.1. Visual Task Analysis

The amount of correctly identified visual searches was analysed for each participant per interface type. From this, the mean and standard deviation were calculated for the participants' responses in the six interface types. These were used to generate Confidence Intervals (CI) at 95%, showing the range of the true population per interface type (figure 3).

The analysis revealed that participants were able to correctly identify more visual information with the DQ version of the interfaces. Furthermore, the stage DQ interface and mixer DQ interface allowed participants to find significantly more visual information than the mixer, scroll and scroll DQ interfaces.

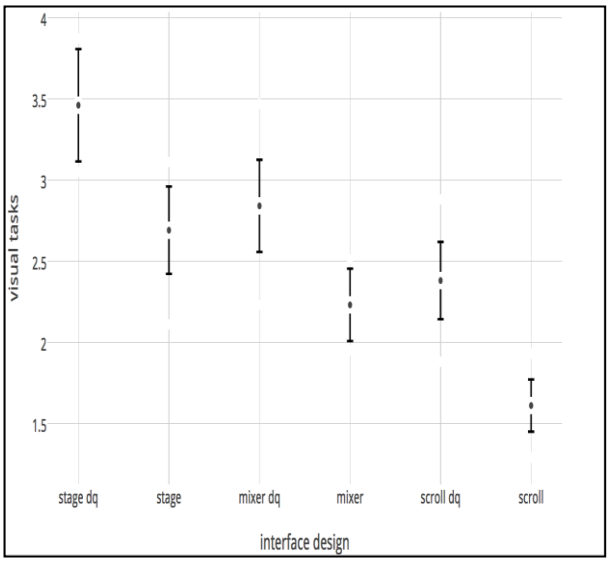

Figure 3. Visual searches successfully completed; Confidence Intervals at 95%. There is an increase in the amount of visual questions answered with the DQ versions of all the interface designs. With the exception of the mixer DQ design, the stage DQ interface yields a significantly greater amount of correctly identified visual information than any of the other interface used in this study.

2.2. Aural Task Analysis

The amount of correctly identified file attenuations were analysed for each of the thirteen participants. This was used to calculate the percentage of correct answers per interface type (table 2). A *z*-test for proportions-

dependent groups was used to determine if the percentages of correct answers from the six interfaces were significantly different from one another (table 3).

76.9	38.4	38.4	23	30.7	7.6
Stage DQ	*Satge*	*Mixer DQ*	*Mixer*	*Scroll DQ*	*Scroll*

Table 2. The Percentage of correctly identified audio file attenuations per interface type.

The analysis at 95% CI showed that the mixer DQ, stage DQ and the stage interface had significantly higher amount of correctly identified audio attenuations than the scrolling interface. Furthermore, the stage DQ produced significantly higher correct answers than the mixer interface and the scrolling DQ interface making it the most effective design in allowing the participants to discern the audio changes.

2.3. Survey Results

Following the study, participants were asked to rate their experience of using the various interfaces. The questions were designed to test their levels of comfort and their perceptions of task completion and success using the different designs with and without DQ filters. The questions asked were as follows:

- Which interface did you feel most comfortable (least stressed/ rushed) using?
- Overall how much did having the sliders help in each interface design?
- Which interface do you think helped you do the listening task best?
- Which interface do you think helped you do the visual task best?

As expected from previous work by the author [13] the scrolling interface was rated the lowest in all questions (figures 4 – 7). The mixer interface scored far more favourably than the scrolling mixer on all questions.

Stage DQ		No	No	**Yes**	**Yes**	**Yes**
Stage	No		No	No	No	**Yes**
Mixer DQ	No	No		No	No	**Yes**
Mixer	**Yes**	No	No		No	No
Scroll DQ	**Yes**	No	No	No		No
Scroll	**Yes**	**Yes**	**Yes**	No	No	
	Stage DQ	*Stage*	*Mixer DQ*	*Mixer*	*Scroll DQ*	*Scroll*

Table 3. Results of the *Z*-test for dependent groups analysis at 95% CI. 'Yes' indicates that there was a significant difference between the interfaces. The Stage DQ design had a significantly higher amount of correctly identified audio attenuations than the mixer, scrolling DQ and scrolling interfaces.

Again in-line with the authors' previous work (ibid), the results suggest that removing the scrolling navigation not only improves listening and visual task completion but also improves the respondents' subjective experience of using the interface. The stage metaphor was rated favourably on all measures. This is especially notable given the novelty of the design to the majority of participants. Indeed, the stage interface had the highest amount of respondents rating it as the interface they felt most comfortable using **(**figure 4). Furthermore, when asked which interface they thought had helped them to successfully complete both the visual task and listening task, the majority of participants named the stage DQ (figures 6 and 7).

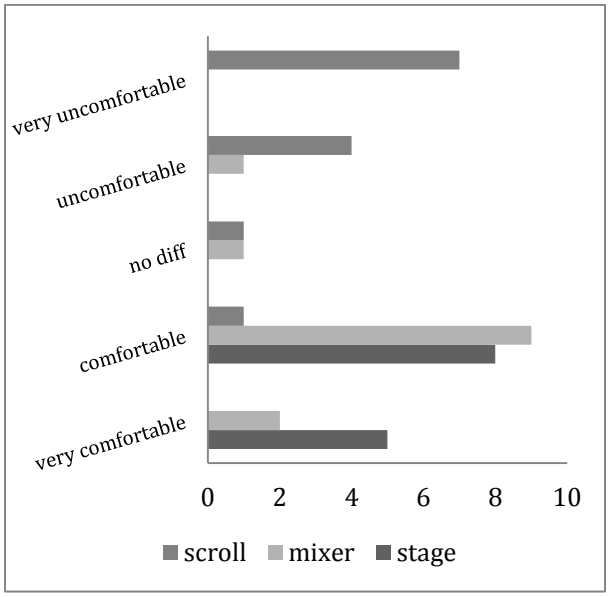

Figure 4. Results from the question 'Which interface did you feel most comfortable using'? The stage and mixer fare more favourably than the scrolling interface.

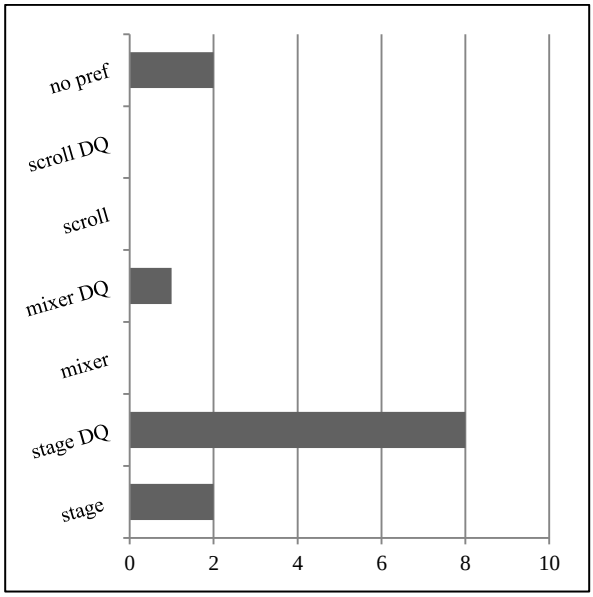

Figure 6. Results from the question 'Which interface do you think helped you do the listening task best'? The stage DQ was perceived as the most effective.

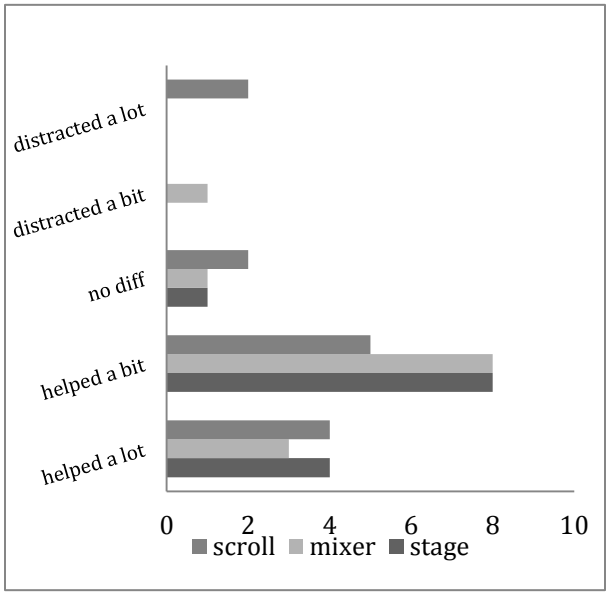

Figure 5. Results from the question 'Overall how much did having the sliders help'? The majority of the participants found them to help.

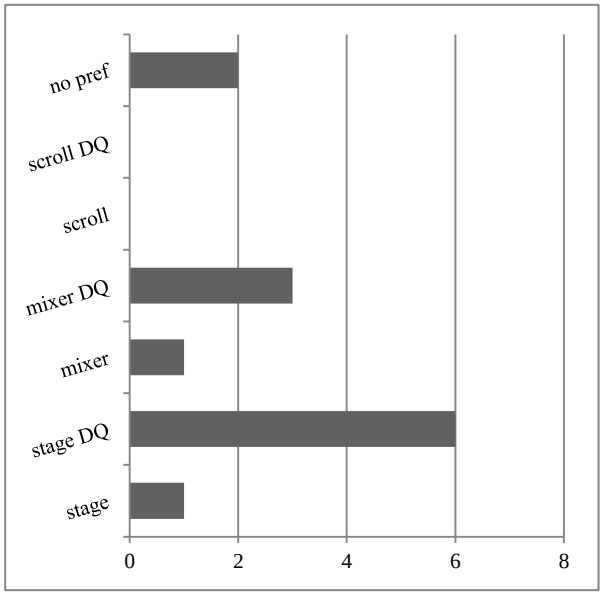

Figure 7. Results from the question 'Which interface do you think helped you do the visual task best'? The stage DQ and mixer DQ are perceived as being most effective. No respondents chose either of the scrolling designs.

3. CONCLUSIONS

The results of the study show that DQ filter interfaces resulted in a higher amount of correctly completed visual search and aural acuity tasks compared to versions of the same interface without them. In keeping with the author's previous work [11, 13,15] the results suggest that interfaces which reduce working memory load and navigation increase not only visual search times but also concurrent critical listening tasks. This finding is also in agreement with other research that found a link between increased visual load and decrease in aural acuity [e.g. 16, 17, 18].

On a more subjective level, the inclusion of DQ sliders was met favourably, with the majority of participants perceiving them as 'helping a lot' compared to non-DQ designs. This was especially the case with the stage design interface. This may be due to the fact that in this design channel numbers were not arranged sequentially, but rather distributed in a more random configuration. DQ filters appear to have been helpful in this regard by allowing the random distribution to be examined according to the users requirements, resulting in an improvement in both visual search and critical listening compared to the non-DQ version.

While results for the mixer DQ were significantly greater than the scrolling interface, as discussed above, the best results were found in the stage DQ design. Though there have been some implementations of the stage design [19,20,21,22] to the best knowledge of the authors they have not incorporated DQ filters. However, their inclusion within the stage design, may allow the user to further benefit from the overview of mix information.

Although DQ filters are found in many other display types (websites, maps etc.) they are not commonly used in DAWs. However, all interfaces used in the study yielded improved listening test results when DQ filters were incorporated, (figure 5 and table 2). By this measure, their inclusion in mixing interfaces may help users in visual search task, while allowing them to remain focused on the audio elements of the mixing process.

4. LIMITATIONS AND FUTURE WORK

In this paper, the use of DQ filters dealt with three attributes of the mix; channel selection, panning and volume. In further work it may be useful to investigate other mix elements, such as equalisation, effects and dynamic processing. Future studies will investigate how these attributes of a mix can be linked to DQ filters and to which extent their inclusion impacts on the speed and accuracy of visual search within the mixing interface display.

5. REFERENCES

[1] Golkhe, K; Hlatky, M; Heise, S; Black, D; Loviscach, J. Track Displays in DAW Software: Beyond Waveform Views. In: Proc. Audio Engineering Society, London, 2010.

[2] Lam, H. A Framework of Interaction Costs in Information Visualization. IEEE Trans. on Vis. and Comp. Graphics, 14, 2008.

[3] Duignan, M; Noble, J; Biddle, R. Metaphors for Electronic Music Production in Reason and Live. In 6th Asia-Pacific Conference on Computer-Human Interaction, 2004.

[4] Rensink, R. The Management of Human Attention in Visual Displays. In Human Attention in Digital Environments. Edited by Claudia Roda. Cambridge University Press, 2012.

[5] Coppinger, R. http:// www. theproaudiofiles.com/ the Complete Guide to Mixing Vocals. Accessed November 2015.

[6] Hutchings, D., Czerwinski, M., Smith, G., Meyers, B., Robertson, G . Display space usage and windowmanagement operation comparisons between singlemonitor and multiple monitor users. In Proceedings of AVI , 2004

[7] Li, Q., North, C. Empirical comparison of dynamic queries sliders and brushing histograms, Proc. IEEE Information Visualization (2003), 147-15.

[8] Shneiderman, B. Dynamic queries for visual information seeking. IEEE Software 11(6): 70-77.1994

[9] Mendelzon .Visualizing the World Wide Web. In Proc. of the working conference on advanced visual interfaces (AVI), 1996.

[10] Card, S., Mackinlay, J., Shneiderman, B. Readings in Information Visualization: Using Vision to Think. Morgan Kaufmann Publishers Inc., San Francisco, CA.1999.

[11] Mycroft, J., Stockman, T and Reiss J.D. 'Audio Mixing Displays; The Influence of Overviews on Information Search and Critical Listening,' International Symposium on Computer Music Modelling and Retrieval (CMMR), 2015.

[12] Moylan, William. Understanding and Crafting the Mix: The Art of Recording. 2nd Edition (Oxford, Focal Press) 2007.

[13] Mycroft, J., Reiss, J.D., Stockman, T. "The Influence of Graphical User Interface Design on Critical Listening Skills," Sound and Music Computing (SMC), Stockholm, July/Aug., 2013.

[14] Everest, F.A . Critical Listening Skills for Audio Professionals. Artist Pro Publications., 1998.

[15] Mycroft, J, Reiss, J. D. and Stockman, T. 'The effect of differing user interface presentation styles on audio mixing'. International Conference on the Multimodal Experience of Music (ICMEM), Sheffield, 23-25 March, 2014.

[16] Colavita FB. Human sensory dominance. Percept Psychophys 16(2):409–4122(9):e943.1974

[17] Sinnett, S., Costa, A., & Soto-Faraco, S. Manipulating inattentional blindness within and across sensory modalities. Quarterly Journal of Experimental Psychology, 2007.

[18] Soto-Faraco, S., & Spence, C. Assessing the role of attention in the audiovisual integration of speech. Information Fusion, 11(1), 4-11, 2010.

[19] V. Diamante, "Awol: Control surfaces and visualization for surround creation." Technical report, University of Southern California, Interactive Media Division, 2007.

[20] J. P. Carrascal and S. Jorda, "Multitouch interface for audio mixing," in Proc. of NIME, pp. 100–103, 2011.

[21] P. Gibson, The Art Of Mixing: A Visual Guide To Recording, Engineering, And Production. ArtistPro Press, 1997.

[22] Gelineck, S , and Korsgaard, D. Stage Metaphor Mixing on a Multi-touch Tablet Device. AES 137th Convention 2014 October 9–12 Los Angeles, USA.