User Experiences of Play-by-Play Captioning in Fast-Paced Live Sports

Date started: October 18, 2019
Date completed: March 23, 2023

1. INTRODUCTION

Based on the issues identified in the Understanding User Responses to Live Closed Captioning in Canada project that was conducted between 2016-2018, this new project continued the research by evaluating new ways of live captioning; comparison between two captioning conditions for live sports broadcasting; a) a conventional condition with nothing changed from what is currently in practice and available on television stations, all spoken words of announcers are captioned (Play-by-Play (PBP)) and b) the commentary announcing only condition, a modification from what is currently practiced and available on television stations (Commentary only (CO). The key issues identified as impacting captioning satisfaction for Deaf and Hard of Hearing participants (D/HoH) were 'Delay' and 'Caption Speed'. Verbatim accuracy, delay and caption speed are three key trade-offs in live captioning.

The main goal of this research was to examine the impact of a non-verbatim method for captioning fast-paced sports on the comprehension and enjoyment of the game, as well as, on the viewing behaviour of D/HoH caption users and their initial reaction to the new condition (CO) presented alongside what they are familiar with seeing (PBP).

Two 10-minute clips from live broadcasted hockey and basketball were selected for the evaluation as examples of captioning for fast-paced live sports broadcasts. Participants were able to select their preferred sport. Eye-tracking behaviour, questionnaires and interviews were used to collect data from 16 D and 11 HoH to evaluate viewing behaviour, subjective opinions about this new form of captioning, and comprehension of the game clips they watched.

This report presents the final results of the study, including eye tracking data, opinions of captioning for the game clips, and comprehension. In addition, a new method of evaluating comprehension is described. It is developed specifically for this project as prior methods reported in the literature, such as multiple choice tests or focus groups, are deemed unable to adequately reflect comprehension.

Future work includes, submitting two manuscripts to academic journals within the next three-months; one reporting the study results, and one reporting the new comprehension evaluation methodology developed during this project. In addition, we are examining the impact of this new captioning style for other live genres such as Olympic sports, and talk shows as well as investigating how automatic captioning algorithms could be exploited for this style.

2. METHODOLOGY

The purpose of the study was to investigate the user experience of real-time live captioning for fast action sports, and to explore an alternative form of captions that could potentially reduce the effort required of users to read captions that exceed 200 words per minute and keep up with the game play in fast-paced sports. An ethics application was submitted to and approved by TMU's research ethics board (including later requirements for a specific protocol for COVID-19 measures). However, all in-person research ceased between April 2020 and February 2022 with a brief reprieve in November 2021 due to COVID-19 restrictions at TMU. As a result, recruiting participants and conducting the study, which required in-person attendance due to the eye-tracking equipment, did not begin until February 2022. In this section, a description of the study protocol, recruitment, materials, and participants is provided.

PARTICIPANTS

Two groups of closed captions users, Deaf and Hard of Hearing (D/HoH), who enjoyed watching live broadcasted sports participated in the study and who were familiar with/fans of either live hockey or basketball broadcasts. Although participants were recruited for the study using a wide variety of methods (see Appendix 1), it took about five months to find sufficient numbers, particularly from the HoH community.

There were a total of 27 participants (16 D, 11 HoH) that participated in the study. Of the 16 Deaf participants there were five women, 10 men, and one transgendered participant. There were four women, seven men HoH participant. In both groups there was a wide age range as seen in Figure 1. There were 13 participants who watched basketball game (5 D, 8 HoH), and 14 participants who watched hockey game (11 D, 3 HoH).
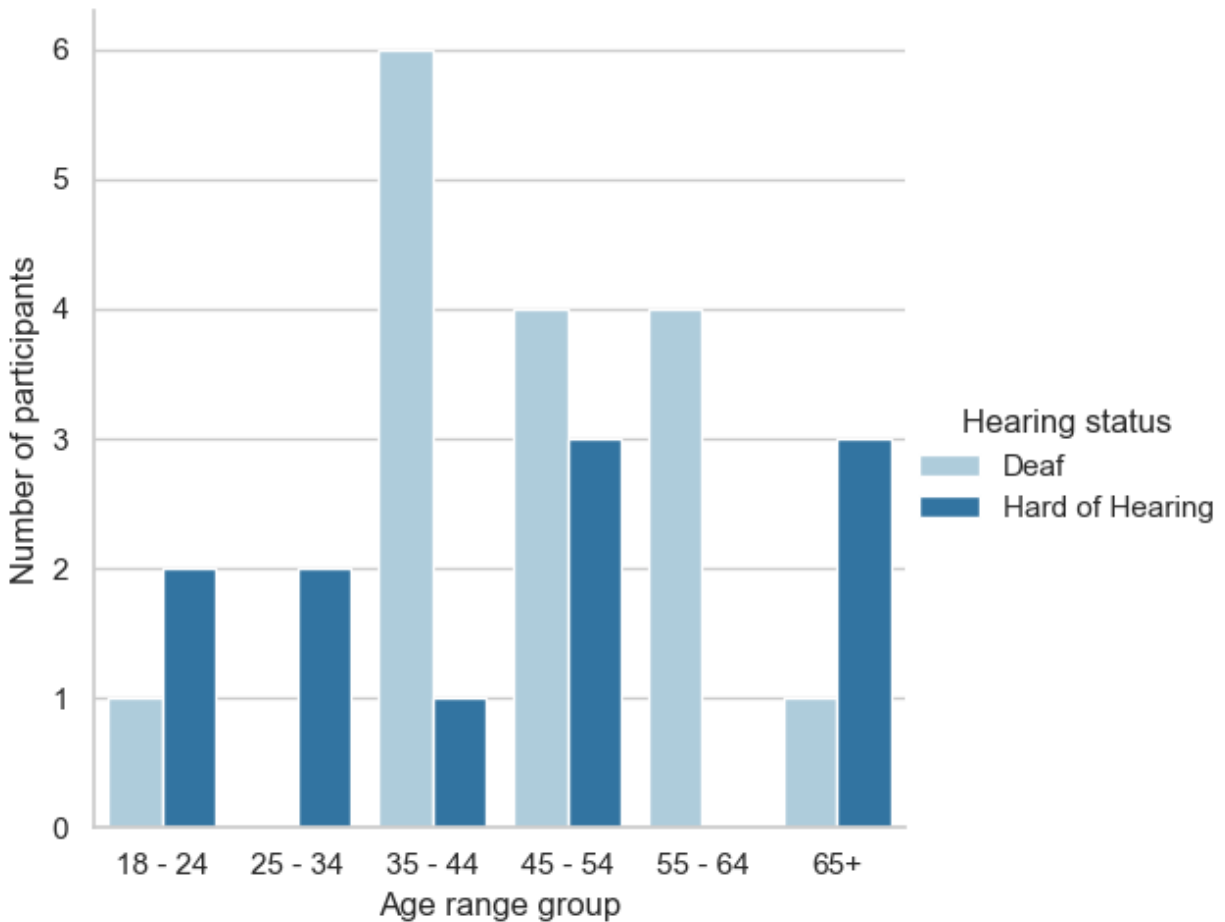
Figure 1. Number of participants in each age range group.

Hearing group	Watch more than two hours of television per day	Prefer watching		Use captions for live sports broadcasting			
		hockey	basketball	Always	Most of the time	Sometimes	Never
Deaf	13 / 16	12	4	12	2	2	0
Hard of Hearing	9 / 11	6	5	4	3	3	1

Among the Deaf participants, a majority (13 of 16) watched more than two hours of television per day. Twelve of 16 preferred to watch hockey, and four preferred basketball. Twelve of 16 always used captions for live-broadcasted sports, two used captions most of the time, and two sometimes.

Among the HoH participants, nine of 11 watched more than two hours of television per day. Six of 11 participants preferred hockey and five basketball. Four participants always used captions for live-broadcasted sports, three used captions most of the time, three sometimes and one never.

APPROACH

The study involved five main tasks: 1) consent and pre-study questionnaire (see Appendix 2); 2) view two live-broadcasted video clips of one game from their preferred sport, hockey (12.26 minutes and 8.29 minutes), or basketball (16.20 min + 18.51 min); 3) complete a mid-study questionnaire between clips; 4) complete a post-study questionnaire after all clips and 5) participate in a 15-minute conversation about the game (topics established a priori to measure comprehension).

Each clip had either conventional captions containing captioning for the game play announcing and the commentary, called the Play-by-Play (PBP) version, or captions of only the commentary portion of the announcing, called the commentary-only (CO) version. Each participant watched a PBP clip and a CO clip presented in a randomized order in order to mitigate any order-effects. Details of the video and caption characteristics are available in Appendix 3. While viewing the video clips, a participant's eye movements and dwell time for various areas of the screen (dependent variables) were tracked using a personally calibrated Tobii eye tracker. The questionnaires were available in English or in ASL.

The pre-study questionnaire (see Appendix 4) consisted of 21 questions that collected demographic data such as gender and age range as well as television viewing behaviours, and captioning use. Participants were also asked to complete a questionnaire consisting of 17 questions after viewing each clip (see Appendix 5). These questions asked about participant's experience of and reaction to the game clip and captions, and the state of flow the participant experienced viewing the clip.

Finally, participants were then invited to have a 15-minute conversation with a researcher about the game in order to assess comprehension. This method had been developed specifically for this project as we determined from the literature and experience from the previous project that multiple choice questions and focus groups were unable to assess comprehension and instead resulted in causing participants to become upset or frustrated with the research. The concept for this new method was to simulate an informal setting where two fans of the sport/game would be discussing what they just watched similar to how one might interact with family and friends after watching a game together. We selected nine topics for hockey and eight for basketball to discuss using this informal style (see Appendix 6). The information for five of the hockey questions and five basketball questions was contained in the Play-by-Play/game play portion of the clips, and information for four hockey and three basketball topics arose from the commentary.

The complete study was designed to take about one hour and 30 minutes, and participants were given an $60 honourarium plus transportation costs for their participation.

The study took place at the IMDC Usability Lab at Toronto Metropolitan University. The setup included a viewing station (television, desk, eye tracking system), and a lounge area (chairs, tables, snacks, beverages), which would be used to conduct the conversation portion of the study. Participants could have ASL interpreters or the CART system present for the study.

DATA ANALYSIS

Eye tracking data were evaluated statistically and graphically for differences between D/HoH viewers (between-subjects variable), and between PBP and CO versions (within-subjects variable). Because the different sports were different from each other in many ways, each sport viewing was treated separately for analysis. Eye tracking data were divided into two main areas-of-interest (AOI): 1) captions; and 2) game play area (see Figure 2 for the sample AOI division).

Figure 2. An example heat map of the areas of interest. The game area was positioned at the centre, since most programming will have cameras to focus at the centre.

We then assessed the time spent in each of these areas for the two types of caption versions and how often participant's eyes moved between the two areas-of-interest. The time spent other than the two defined areas-of-interest were also calculated by the difference between the total time of interest and the tracked time.

The time (i.e., duration for visit) data were normalized for the different clip durations and the eye tracker gaze detection sampling rates (see Appendix 7 for a complete outline of the various eye-tracking data elements and definitions). The questionnaire data were also analyzed statistically. The level of significance for statistical tests was set to 0.05 (i.e., $p < 0.05$). The comprehension data were divided into correct, incorrect and uncertain responses about the topics, and analyzed statistically for differences between the groups. Any conversation elements that were not related to the topics were not considered.

RESULTS

Viewing behaviour

Viewing behaviour of participants was measured as the **total number of visits** and the **total duration for visits** in the caption area and game play area of the screen. Figure 3 depicts the overall distribution of the number of visits and duration we used in this analysis.
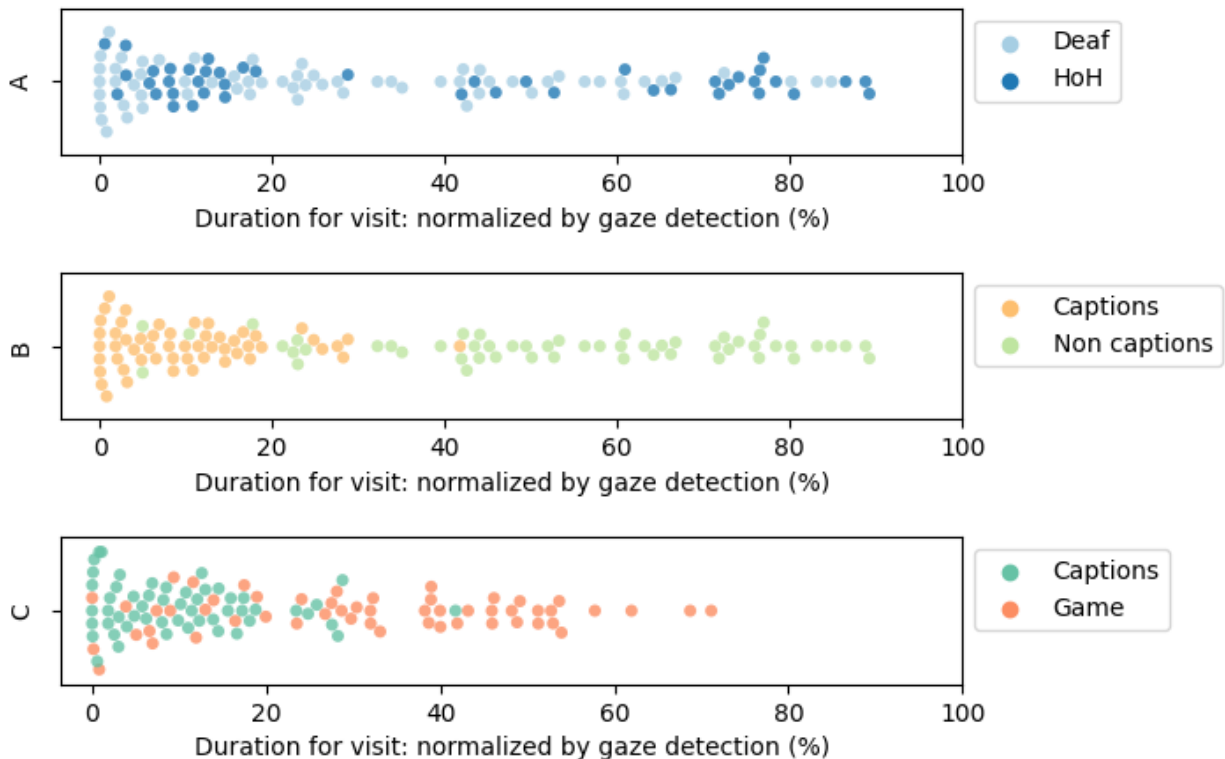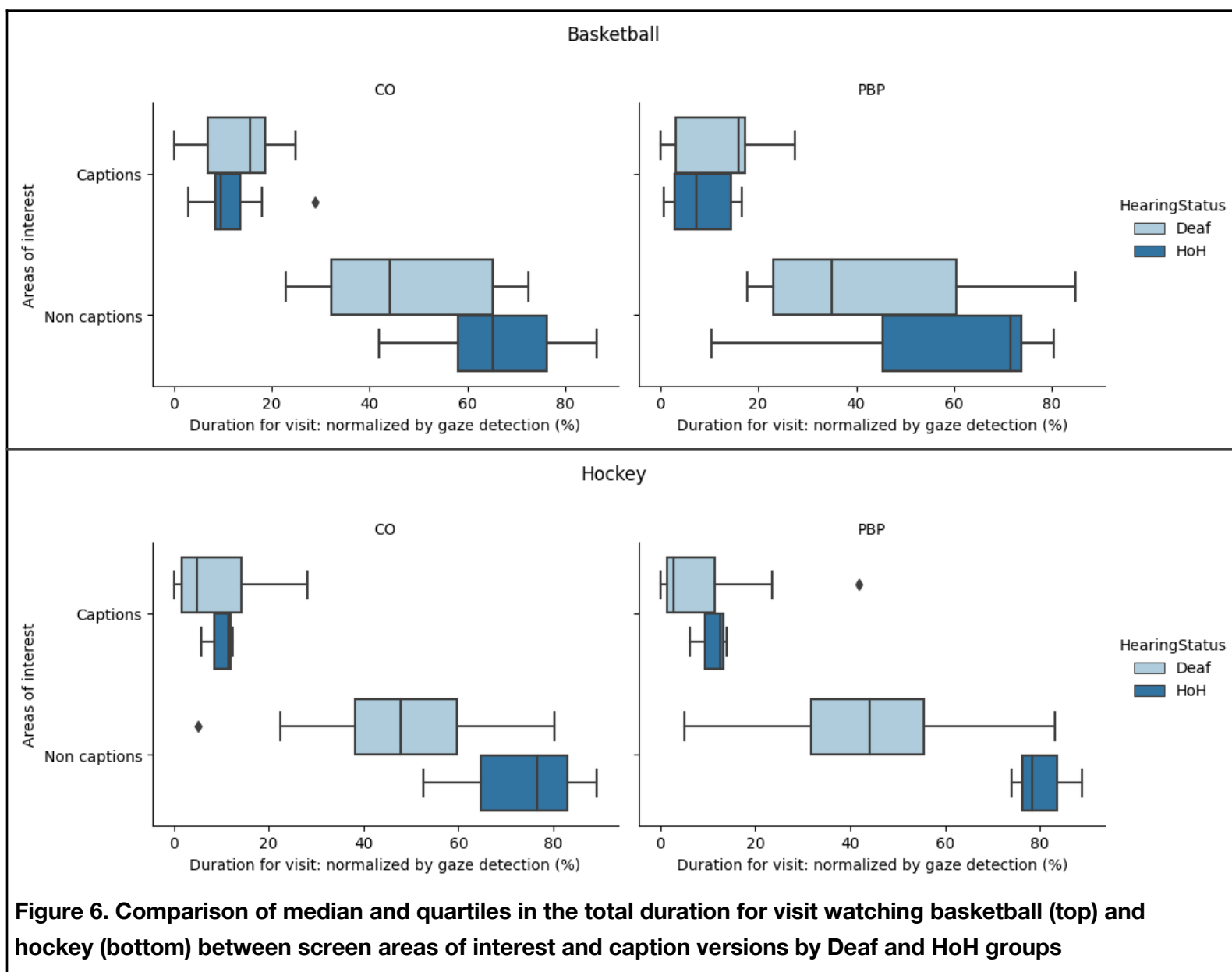

Figure 3. Swarm plot of the overall total number of visits by D/HoH participants (A), the duration for visit by the areas of interest (B), and the duration for visit by having captions as the only area of interest (C).

A Shapiro-Wilks normality test was conducted for all data and most of the data were non-normally distributed ($p < 0.05$). As a result, non-parametric statistics were used to assess differences between independent and dependent variables.

Duration for visit: Caption vs. the rest of the screen

Considering the non-AOI duration, it is possible to compare duration for visits in caption area against the area excluding the designated caption area (i.e., having caption area as the only AOI). A non-parametric Wilcoxon rank sum test was carried out to determine whether there were any differences in the total duration for visit. Figure 6 shows the comparison of median and quartiles in the total duration for visit.

Figure 6. Comparison of median and quartiles in the total duration for visit watching basketball (top) and hockey (bottom) between screen areas of interest and caption versions by Deaf and HoH groups
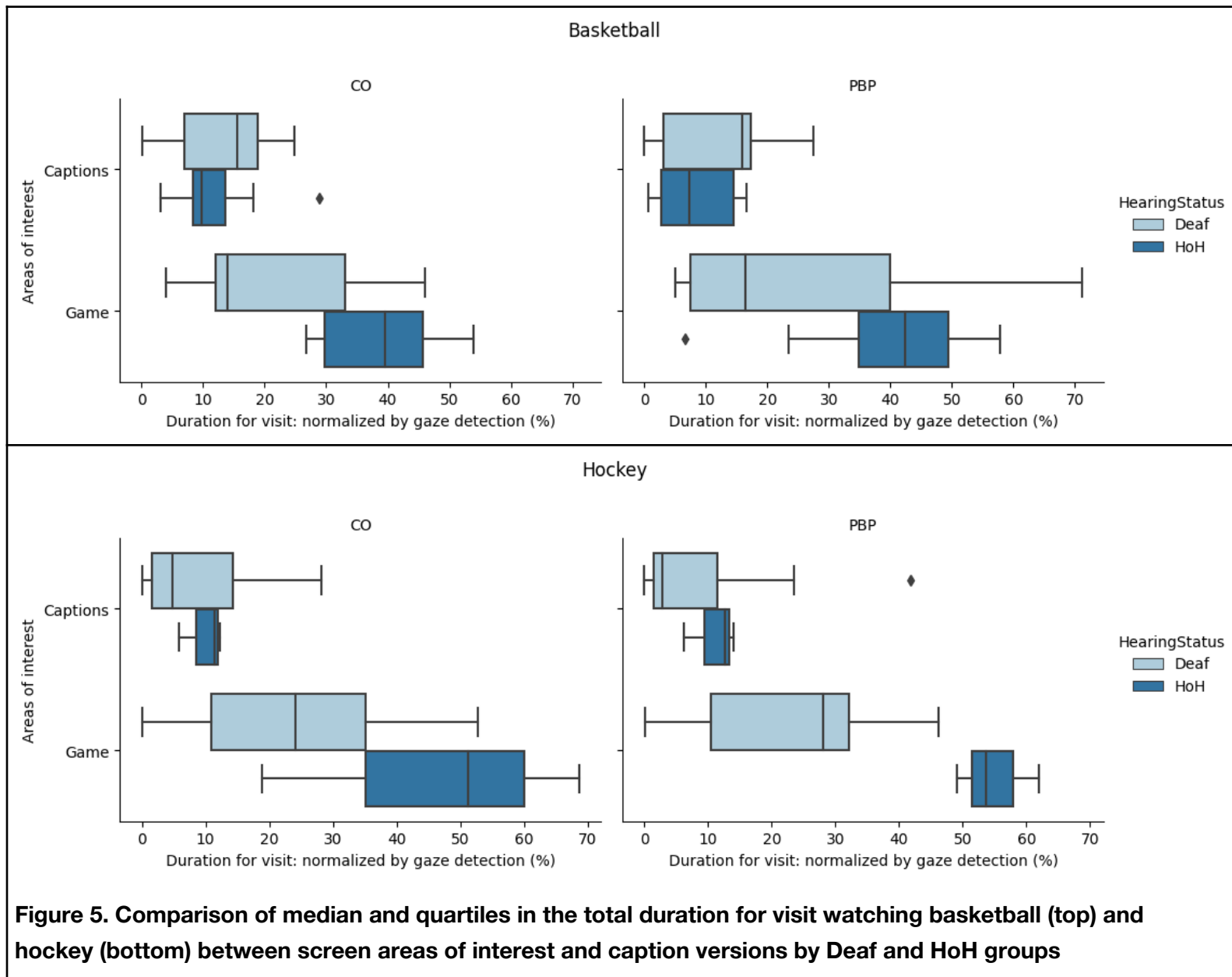
For D participants watching **hockey** in PBP condition, there was a significant difference in the duration for visit between game area (MD=44.00%, IQR=23.89%) and caption area (MD=2.79%, IQR=9.98%); Z(11) = -4.107, p < 0.05. In the CO condition, there was also a significant difference in the total duration for visit between game area (MD=47.88%, IQR=21.51%) and caption area (MD=4.85%, IQR=12.55%); Z(11) = -4.107, p < 0.05.

For HoH participants watching **basketball** in PBP condition, there was a significant difference in the total duration for visit between game area (MD=71.58%, IQR=28.50%) and caption area (MD=6.54%, IQR=11.44%); Z(9) = -3.724, p<0.05. For the CO condition, there was also a significant difference in the total duration for visit between game area (MD=65.22%, IQR=18.14%) and caption area (MD=8.57%, IQR=4.06%); Z(9) = -3.724, p<0.05.

A non-parametric Wilcoxon rank sum test was carried out to determine whether there were any differences in the total duration for visit between the areas of interest and caption versions. Figure 5 shows the comparison of median and quartiles in the total duration for visit.

Figure 5. Comparison of median and quartiles in the total duration for visit watching basketball (top) and hockey (bottom) between screen areas of interest and caption versions by Deaf and HoH groups

For D participants watching **hockey** in PBP condition, there was a significant difference in the duration for visit between game area (MD=28.13%, IQR=21.65%) and caption area (MD=2.79%, IQR=9.98%); $Z(11) = -4.107$, $p < 0.05$. In the CO condition, there was also a

significant difference in the total duration for visit between game area (MD=24.08%, IQR=24.11%) and caption area (MD=4.85%, IQR=12.55%); Z(11) = -4.107, p < 0.05.

For HoH participants watching **basketball** in PBP condition, there was a significant difference in the total duration for visit between game area (MD=45.93%, IQR=12.56%) and caption area (MD=6.54%, IQR=11.44%); Z(9) = -3.724, p<0.05. For the CO condition, there was also a significant difference in the total duration for visit between game area (MD=38.97%, IQR=15.69%) and caption area (MD=8.57%, IQR=4.06%); Z(9) = -3.724, p<0.05.
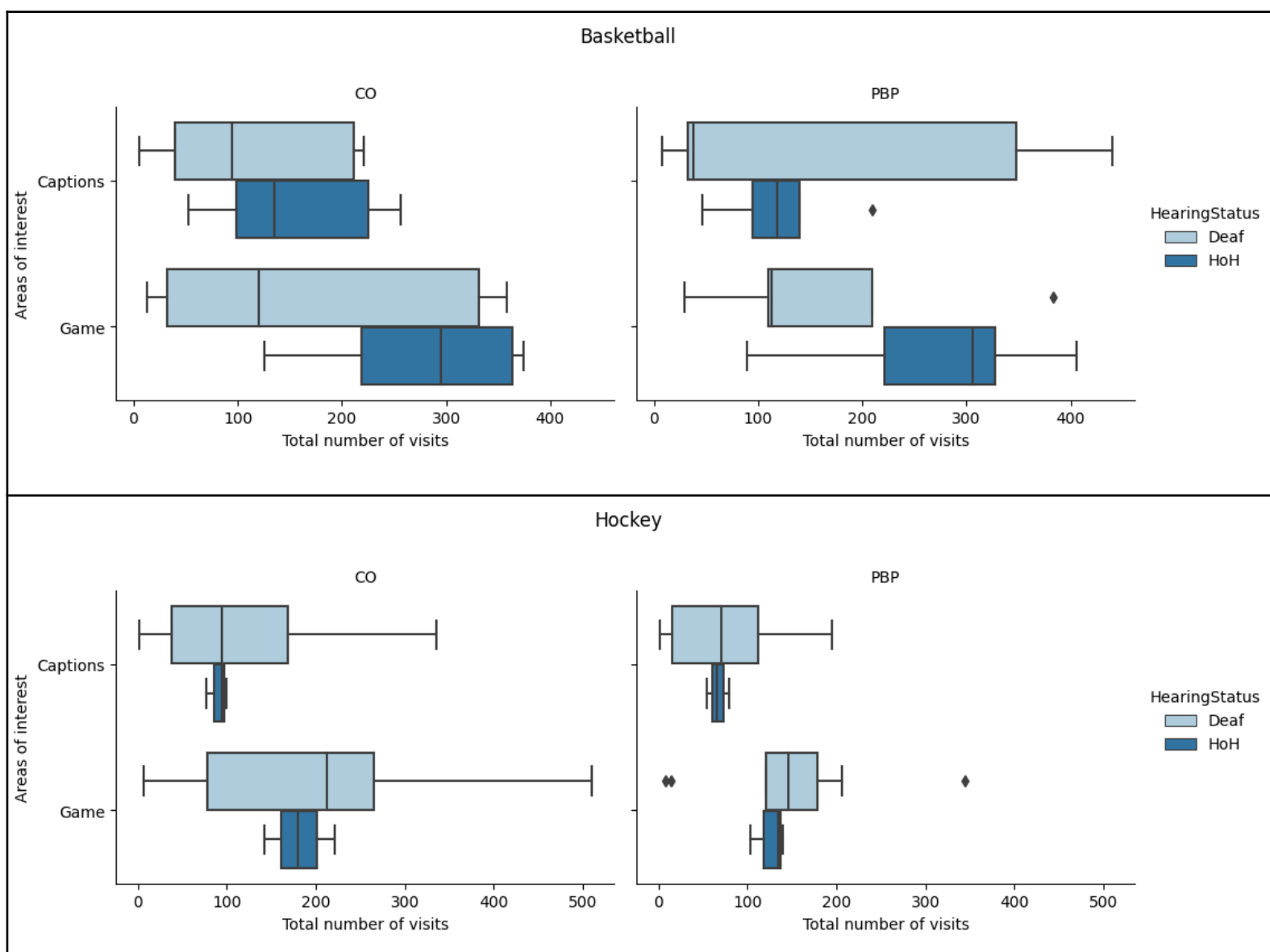
<u>Total number of visits: Caption vs. Game screen</u>

A non-parametric Wilcoxon rank sum test was carried out to determine whether there were any differences in the total number of visits between the areas of interest and caption versions. Figure 4 shows the median and quartile graphs for the total number of visits.

For D participants watching **hockey** in the PBP condition, there was a significant difference in the total number of visits between game play area (MD=146, IQR=57.5) and caption area (MD=71, IQR=96.5); Z(11) = -4.107, p < 0.05. Another significant difference was also found in the total number of visits for D participants watching hockey in the CO condition, between game play area (MD=213, IQR=188) and caption area (MD=94, IQR=129.5); Z(11) = -4.107, p < 0.05.

For HoH participants watching **basketball** in the PBP condition, there was a significant difference in the total number of visits between game area (MD=296, IQR=116) and caption area (MD=106, IQR=57); Z(9) = -3.724, p < 0.05. Another significant difference was also found in the total number of visits for HoH participants watching basketball in the CO condition, between game area (MD=291, IQR=166) and caption area (MD=114, IQR=123); Z(9) = -3.724, p < 0.05

Figure 4. Comparison of median and quartiles in the total number of visits watching basketball (top) and hockey (bottom) between screen areas of interest and caption versions by Deaf and HoH groups

Due to insufficient sample size, no statistical comparative test was performed for HoH participants watching hockey. See Table 1 for all descriptive metrics of the total number of visits, and for the total duration for visit.

Table 1. Descriptive statistics for the total number of visits and the total duration for visit across areas of interests, caption versions, and hearing conditions

Sports	Hearing group	Caption version	Areas of interests	Total number of visits Median (IQR)	Normalized ratio of duration for visit Median (IQR)
Basketball	Deaf	PBP	Captions	38 (316)	15.94% (14.05%)
			Non Captions		35.06% (37.45%)
			- Game play	113 (100)	16.46% (32.6%)
		CO	Captions	94 (171)	15.61% (11.87%)
			Non Captions		44.08% (32.97%)
			- Game play	120 (300)	13.96% (21.15%)
	HoH	PBP	Captions	118 (44.5)	7.37% (11.7%)
			Non Captions		71.58% (28.50%)
			- Game play	306 (105.75)	42.44% (14.54%)
		CO	Captions	135 (126.75)	9.68% (5.2%)
			Non Captions		65.22% (18.14%)
			- Game play	295 (144.5)	39.48% (15.96%)
Hockey	Deaf	PBP	Captions	71 (96.5)	2.79% (9.98%)
			Non Captions		44.00% (23.89%)
			- Game play	146 (57.5)	28.13% (21.65%)
		CO	Captions	94 (129.5)	4.85% (12.55%)
			Non Captions		47.88% (21.51%)
			- Game play	213 (188)	24.08% (24.11%)
	HoH	PBP	Captions	66 (12.5)	12.61% (3.85%)
			Non Captions		78.39% (7.35%)
			- Game play	135 (18.5)	53.67% (6.41%)
		CO	Captions	94 (11.5)	11.45% (3.27%)
			Non Captions		76.55% (18.27%)
			- Game play	180 (39.5)	51.27% (24.89%)

In general, for both hearing groups, regardless of sports or caption conditions, participants visited the game play area more frequently than the caption area. This trend is also found in the duration for visit, where participants spent less time in the caption area compared to the game play area. For both sports, participants from both hearing groups visited more in the caption

area with CO condition than PBP condition, while the duration for visit did not show noticeable differences.

Quality factors of Closed Captioning, Overall satisfaction, Caption preferences, and Comprehension score

Caption version preference

One participant did not answer the survey questions. Therefore all questionnaire responses were out of 26 participants. Most (17/26) participants selected no preference option. From the rest of the participants, there were participants who preferred CO (7/26) more than PBP (2/26). There were more Deaf participants reporting that they preferred the CO condition (5/7) than HoH (2/7) (See Figure 7 for basketball, and Figure 8 for hockey).
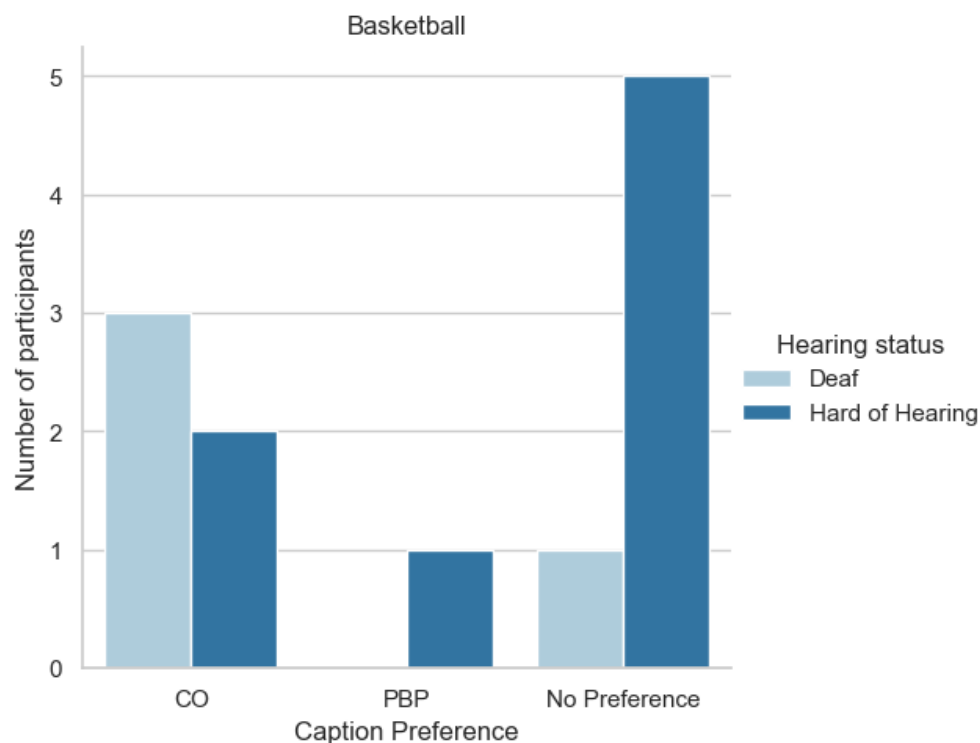
Figure 7. Frequency of preference rating after watching basketball clips for D/HoH participants.
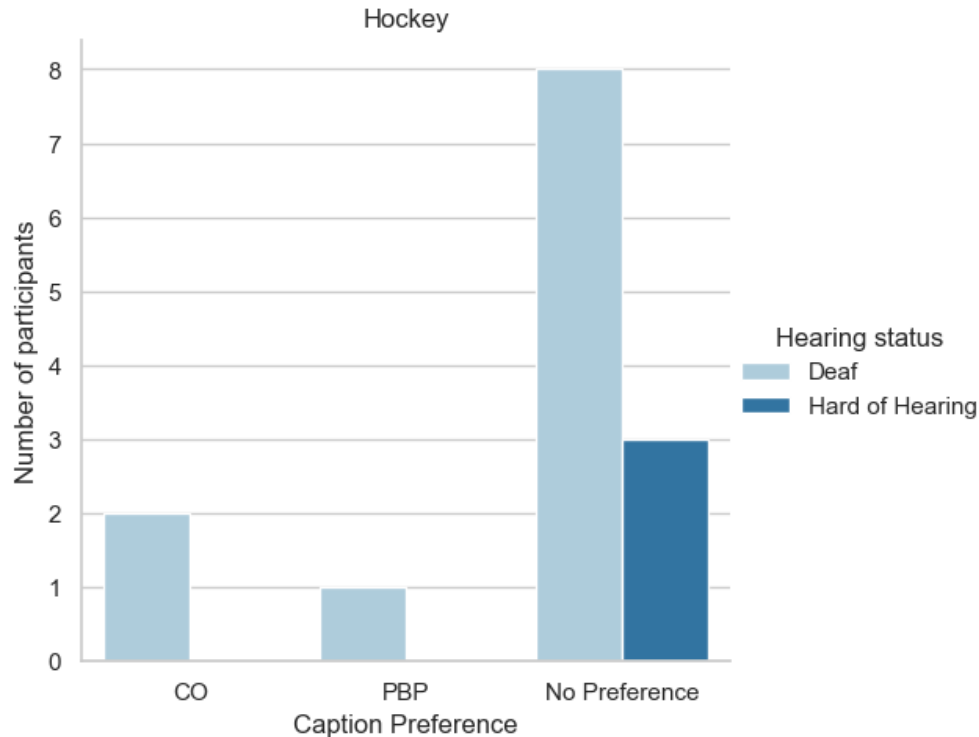
Figure 8. Frequency of preference rating after watching hockey clips for D/HoH participants.

Overall satisfaction of the captions

We also collected questionnaires about the overall satisfaction with the quality of Closed Captioning (CC) as well as viewing experience followed by its factors and perceived impact of the factors to the overall ratings. Table 2 below shows the descriptive statistics of overall satisfaction and the viewing experience with the CC quality after watching clips in different caption conditions and order.

By comparing the mean overall satisfaction with the CC quality, in both sports, CO captions received a higher satisfaction ratings than PBP captions. Participants who watched PBP after watching CO ended up rating the quality of PBP less than the quality of CO. No other particular order effect pattern seem to exist. As from the viewing experience ratings, participants who watched basketball clips seemed to have more positive experience than participants who watched hockey clips. In addition, participants were asked to answer how much of the captions help following the game. Similar to the quality satisfaction ratings, the descriptive statistics were slightly higher for CO than PBP; Basketball PBP (M=3.00, SD=1.55), Basketball CO (M=3.58, SD=1.16), Hockey PBP (M=3.14, SD=1.10), Hockey CO (M=3.50, SD=1.22). In addition to the two direct representative ratings, one of the questions we asked was about the

helpfulness of captioning in understanding/following the game (see Appendix 9 for other additional ratings). When compared the two captioning versions, the helpfulness seemed to be slightly higher when participants watched CO captioning (M=3.58, SD=1.16 from basketball, M=3.50, SD=1.22 from hockey) than PBP captioning (M=3.00, SD=1.55 from basketball, M=3.14, SD=1.10 from hockey).

Table 2. Descriptive statistics of the overall caption satisfaction and viewing experience rating after watching the clips in each caption condition and order. All ratings are out of 5, where 1 = Most Negative and 5 = Most Positive.

Sports	watched condition	Caption satisfaction Mean (SD)	Caption satisfaction Overall Mean (SD)	Viewing experience Mean (SD)	Viewing experience Overall Mean (SD)
Basketball	PBP (first)	**3.83 (0.98)**	3.18 (1.25)	4.00 (0.58)	4.25 (0.62)
	PBP (second)	**2.40 (1.14)**		4.60 (0.55)	
	CO (first)	3.60 (0.55)	3.67 (1.07)	4.40 (0.89)	4.33 (0.98)
	CO (second)	3.71 (1.38)		4.29 (1.11)	
Hockey	PBP (first)	**3.29 (1.38)**	3.00 (1.36)	3.43 (1.13)	3.36 (1.08)
	PBP (second)	**2.71 (1.38)**		3.29 (1.11)	
	CO (first)	3.57 (1.51)	3.14 (1.70)	4.00 (0.82)	3.64 (1.28)
	CO (second)	2.71 (1.89)		3.29 (1.60)	

The overall satisfaction and viewing experience ratings were also analyzed by each hearing group (see Table 3). For basketball, Deaf participants rated both PBP and CO the same, whereas HoH participants rated CO higher. For Hockey, Deaf participants rated CO higher, whereas HoH participants rated PBP higher.

Table 3. Descriptive statistics of the overall satisfaction and viewing experience rating after watching the clips in each caption condition.

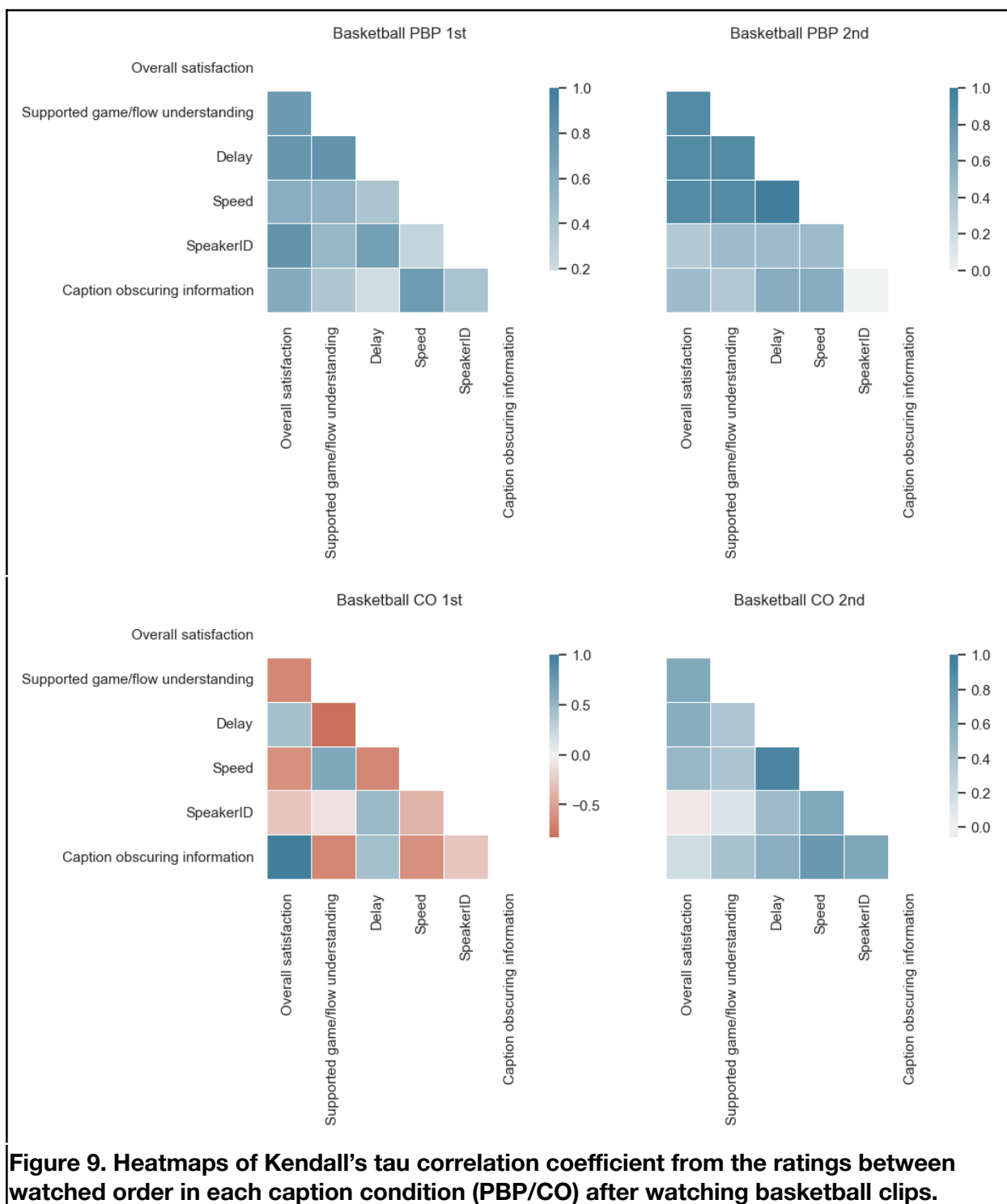| Sports	watched condition	Hearing group	Caption satisfaction Mean (SD)	Viewing experience Mean (SD)
Basketball	PBP	Deaf	3.75 (0.96)	4.50 (0.58)
		Hard of Hearing	2.86 (1.35)	4.13 (0.64)
	CO	Deaf	3.75 (0.96)	4.25 (0.96)
		Hard of Hearing	3.63 (1.19)	4.38 (1.06)
Hockey	PBP	Deaf	2.91 (1.38)	3.45 (1.13)
		Hard of Hearing	3.33 (1.53)	3.00 (1.00)
	CO	Deaf	3.36 (1.75)	3.73 (1.27)
		Hard of Hearing	2.33 (1.53)	3.33 (1.53)

Correlation between overall satisfaction ratings and the quality factors of the captions
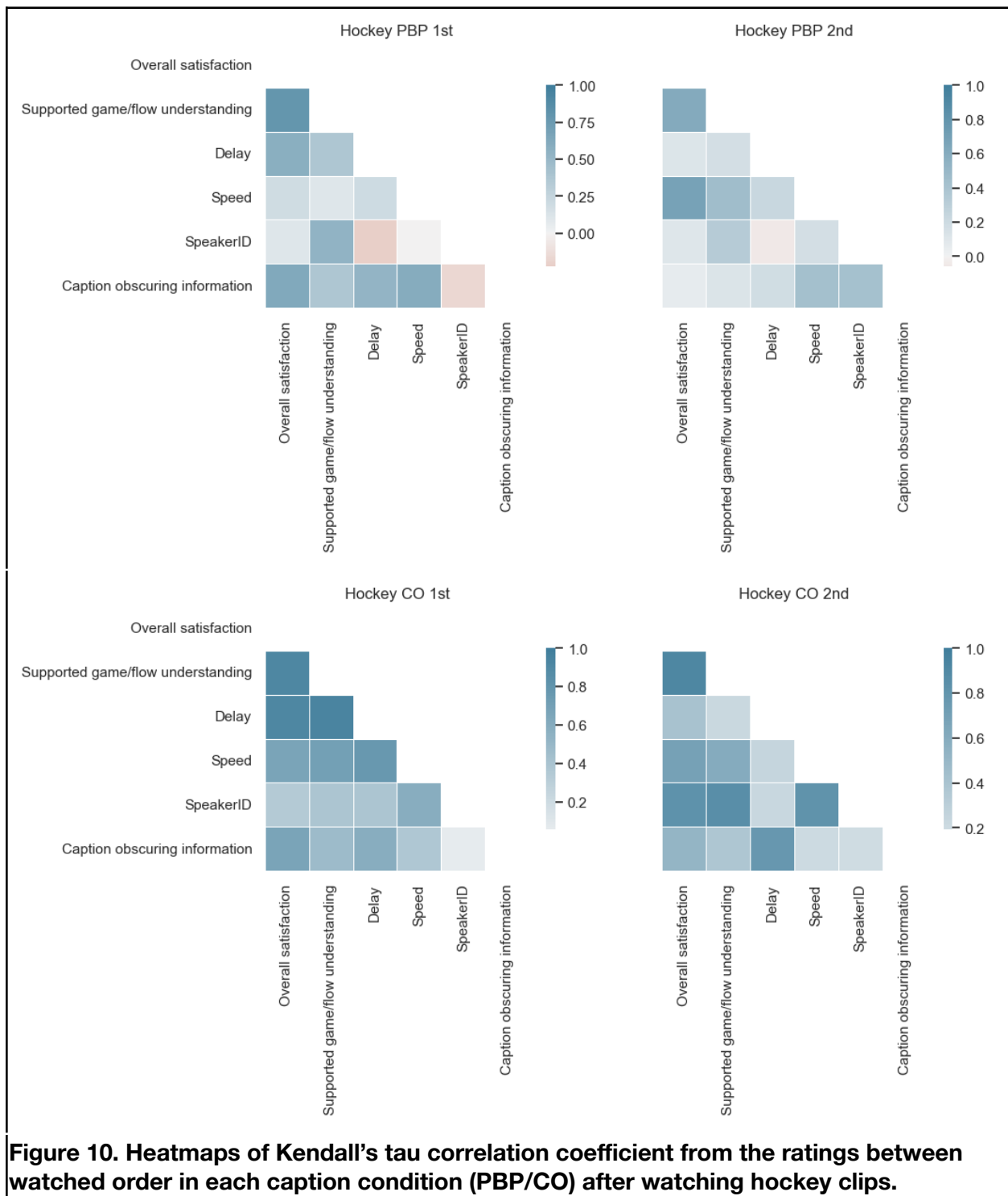
Kendall rank correlation tests were performed to evaluate the relationships between questions in the mid and post questionnaires and the different orders of clip viewing to assess order effect (see Figure 9 for correlation heatmap after watching basketball, and Figure 10 for hockey). The questionnaire consisted of 5-point Likert scale type responses, with 1 = Most Negative to 5 = Most Positive response categories. The descriptive statistics for all questionnaire responses were added in the Appendix 9. Overall, the captioning quality satisfaction ratings and the visual experience ratings had a statistically significant correlation ($t > 0.55$, $p<0.05$), where the correlation was slightly higher with captioning quality satisfaction ratings and the quality factor ratings, than with visual experience ratings.

From watching basketball, participants seems to rate similarly on PBP condition regardless of the watched order, whereas the rating pattern seem to be changed when they watched CO condition. From watching hockey, participants seems to rate similarly on both conditions regardless of the watched order.

Table 4, Table 5, Table 6, and Table 7 shows statistically significant Kendall's Tau values for all pair combinations that were used to generate the heatmaps. Pairs other than displayed showed no statistically significant correlation.

Overall, the ratings of the overall caption quality satisfaction seem to be associated with the factors of the captioning; delay, speed, speaker identification, and caption positioning.

Figure 9. Heatmaps of Kendall's tau correlation coefficient from the ratings between watched order in each caption condition (PBP/CO) after watching basketball clips.

Figure 10. Heatmaps of Kendall's tau correlation coefficient from the ratings between watched order in each caption condition (PBP/CO) after watching hockey clips.

Table 4 Kendall's tau correlation coefficient of the selected questionnaire rating responses after watching basketball with PBP captioning first (top) and second (bottom).

	Overall satisfaction	Supported understanding	Delay influence	Speed influence	Speaker ID enabled	Caption positioning
Delay influence	0.783 (p=0.049)	0.801 (p=0.037)				
Speaker ID enabled	0.818 (p=0.045)					
Caption positioning					0.402 (p=0.0328)	
Supported understanding	0.889 (p=0.037)					
Delay influence	0.889 (p=0.037)	0.889 (p=0.037)				
Speed influence	0.889 (p=0.037)	0.889 (p=0.037)	1.00 (p=0.019)			

Table 5 Kendall's tau correlation coefficient of the selected questionnaire rating responses after watching basketball with CO captioning first (top) second (bottom).

	Overall satisfaction	Supported understanding	Delay influence	Speed influence	Speaker ID enabled	Caption positioning
Caption positioning	1.00 (p=0.046)					
Speed influence			0.939 (p=0.010)			
Caption positioning				0.759 (p=0.043)		

Table 6 Kendall's tau correlation coefficient of the selected questionnaire rating responses after watching hockey with PBP captioning first (top) and second (bottom).

	Overall satisfaction	Supported understanding	Delay influence	Speed influence	Speaker ID enabled	Caption positioning
Supported understanding	0.915 (p=0.009)					
Delay influence	0.915 (p=0.009)	0.944 (p=0.006)				
Speed influence		0.707 (p=0.047)	0.766 (p=0.031)			
Supported understanding	0.915 (p=0.009)					
Speaker ID enabled	0.824 (p=0.019)	0.857 (p=0.014)		0.814 (p=0.022)		
Caption positioning			0.778 (p=0.032)			

Table 7 Kendall's tau correlation coefficient of the selected questionnaire rating responses after watching hockey with CO captioning first (top) and second (bottom).

|	Overall satisfaction	Supported understanding	Delay influence	Speed influence	Speaker ID enabled	Caption positioning
Supported understanding	0.915 (p=0.009)					
Delay influence	0.915 (p=0.009)	0.944 (p=0.006)				
Speed influence		0.707 (p=0.047)	0.766 (p=0.031)			
Supported understanding	0.915 (p=0.009)					
Speaker ID enabled	0.824 (p=0.019)	0.857 (p=0.014)		0.814 (p=0.022)		
Caption positioning			0.778 (p=0.032)			

Each participant had a short 15-minute post-study interview with a researcher, where the intention was to have a casual conversation atmosphere that would emulate a natural conversation setting. Simple topics relating to the comprehension of the game content (such as defense/offense, referee calls, and players' actions) were discussed and later video analysis was conducted on participant responses and a determination made on participant comprehension via correct, incorrect, or uncertain answers to the topic questions in the interview. Different numbers of questions were asked to each participant because of three reasons: 1) there were different genres of video clips that participants watched. There were a total of eight questions for Basketball, whereas, Hockey had nine questions total; 2) the interview was focused on making the atmosphere friendly and casual and not a test setting, so it would not be natural to make sudden topic changes; and 3) the interview time was limited.

There were two question types: CO and Gameplay. Gameplay questions asked about the gameplay specifically, which expected participants to be able to answer them from watching the clips. Then, CO questions asked about the information which was displayed on CO captioning only. Thus, participants were expected to answer the questions only after watching clips with the CO captions. There were no statistically significant correlations between comprehension scores and overall quality ratings across caption conditions.

Figure 11 and Figure 12 displays the mean and standard deviation of the comprehension scores between hearing groups for each sports clip watched, split by question types.

Overall, just comparing the mean values, HoH group had higher comprehension scores (M=64.55%, SD=30.23% for Basketball, M=91.67%, SD=12.91% for Hockey) than Deaf group (M=67.50%, SD=20.29% for Basketball, M=73.20%, SD=28.51% for Hockey), while comprehension scores after watching Hockey clips were slightly higher.

In Basketball, the descriptive statistics suggest slightly higher Gameplay comprehension scores from both D group (M=68.00%, SD=27.97%) and HoH group (M=73.13, SD=24.34) than CO comprehension scores (D group (M=63.33, SD=24.72) and HoH group (M=66.67, SD=35.64).
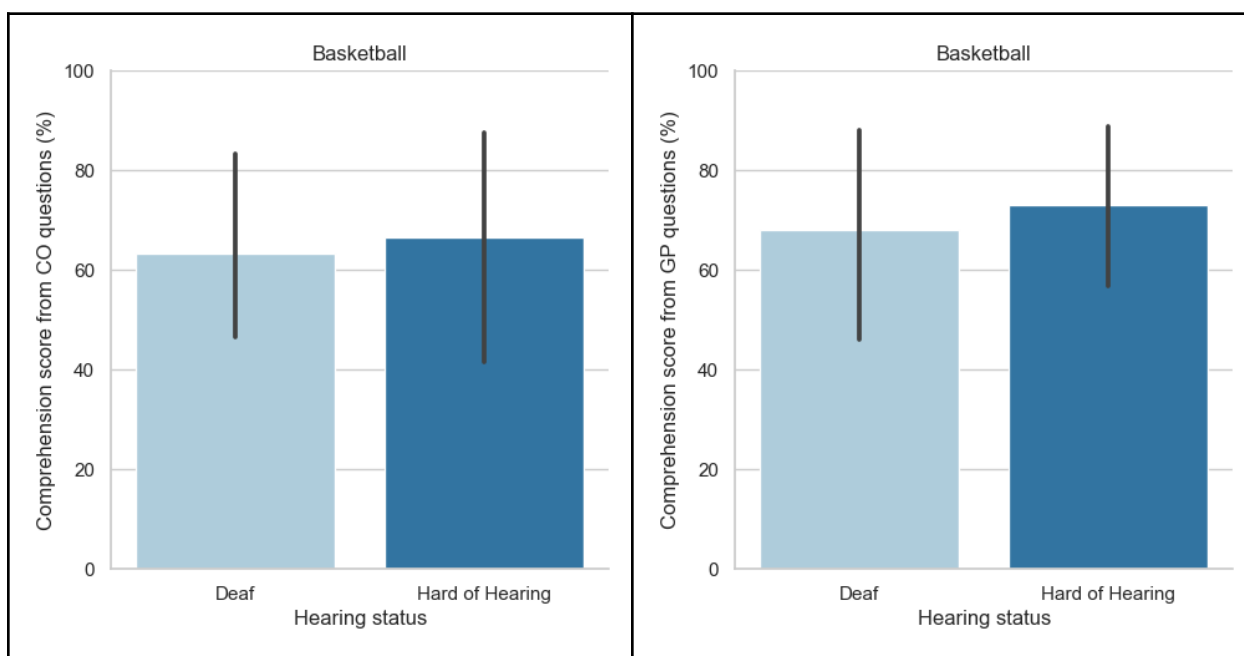
Figure 11. Comprehension score from CO and Gameplay questions for each hearing group from watching basketball.

In Hockey, the descriptive statistics suggest slightly higher Gameplay comprehension scores from both D group (M=88.64%, SD=14.16%) and HoH group (M=100.0, SD=0.0) than CO comprehension scores (D group (M=65.15, SD=32.02) and HoH group (M=83.33, SD=24.87).
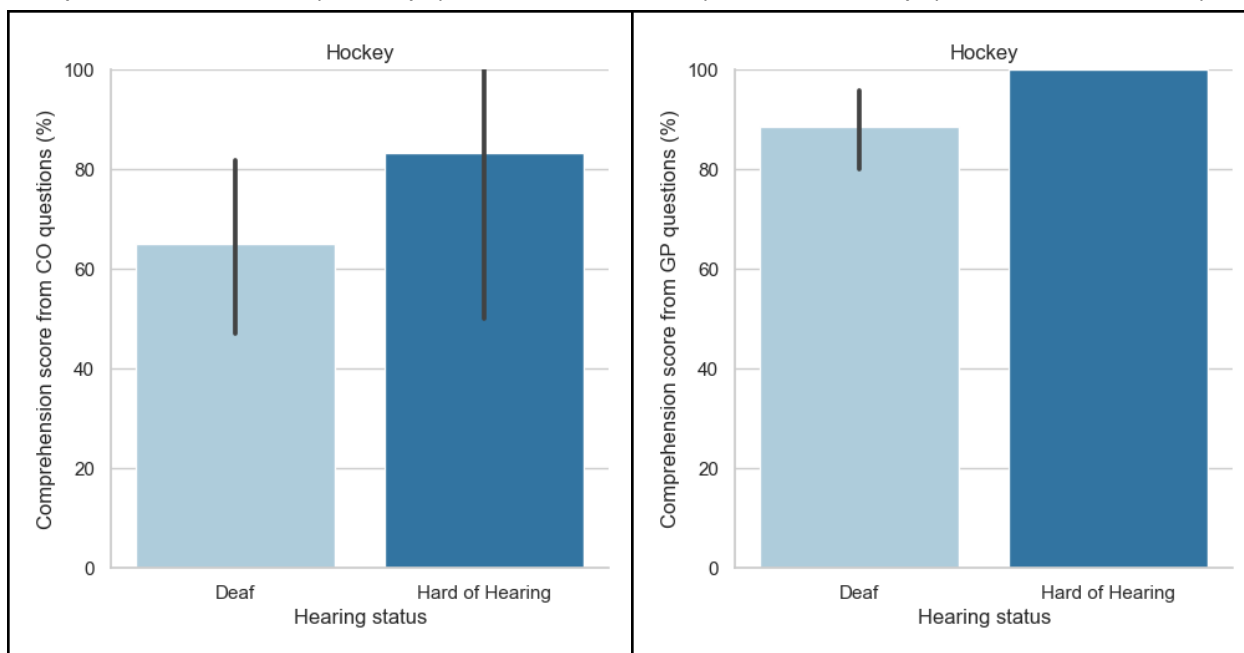
Figure 12. Comprehension score from CO and Gameplay questions for each hearing group from watching hockey.

For the Hockey group, a Mann-Whitney U test indicated that comprehension scores from Gameplay questions were statistically higher than the comprehension scores from CO

questions, (Mann–Whitney $U = 143.00$, $P < 0.05$, two-tailed). There were no statistically significant differences in comprehension scores between CO questions and Gameplay questions for the Basketball group.

DISCUSSION

From comparing the data collected using the eye tracking system, we found that there were no significant differences between watching the two caption conditions. However, when comparing the descriptive values, it was possible to infer that participants visited more to the CO captions than PBP captions. There seemed to be high variances amongst all data, suggesting the need for a further study with larger sample size. Having larger sample sizes would support our comparisons between the eyetracking indicators and caption conditions as it is expected to have smaller variances and convergence of data spread.

There were also similarities between the ways participants watched both conditions. For example, in both instances, participants' spent longer watching outside of the caption area than within the caption area. This was expected behaviour given the fast paced visual nature of basketball and hockey sports, requiring a certain level of concentrated attention and focus that cannot be maintained while reading lengthy captions or text. Another interpretation of the relatively little time spent on the caption area (in relation to the game screen area) could come from the rationale some participants observed for using captions; to *"recall the names/plays that they missed"*. The reading of player names requires comparatively less time then complete phrases/comments made by announcers.

Furthermore, while there were a small number of participants who reported their preferred caption condition, the majority of the participants were uncertain and had no preference relating to the caption conditions. Participants reported the PBP captions were *"too distracting"*, *"stressful"*, *"seriously overwhelming", "a lot of work for my [his] eyes"*, *"a lot of visual back and forth [looking between the captions and game screen]"*, *"on-going"*, *"too much"*, "feel[ing] a little stressed because I do not want to miss the important information before it disappears", "like reading a book and when I looked up I would miss the captions" and "the amount of reading I had to do was over the top", resulting in participants avoiding the captions and going *"straight to watching the game"*.

When comparing the overall satisfaction ratings and caption support ratings in context understanding, we could find the difference between CO and PBP captions in both sports. Many participants seemed to be satisfied more with the quality of CO captions than PBP captions regardless of the watched order or context. This finding was fortified with the participants comments: One participant said with a big smile and applause sign [in ASL] that

she could actually watch and *"respond to the game"* with the CO captions. Similarly, a participant notes *"you were actually able to watch the game and look at the captions at the same time"*. Even when participants identified both as being acceptable forms of captioning, they did appreciate CO for being *"less distracting"*.

Some participants even felt that the *"language [was] enough"* to *"see everyone moving around the ice"* suggesting a text translation was not needed for these sports. One participant summed up what many were expressing in his/her mid and post study questionnaire: *"I could concentrate on the game and not worry that I was missing out on something by trying to keep up with both"*, *"It was at an eye level that even when the play was starting I could finish reading and still see what was happening in my peripheral [vision]."* This participant also expressed his caption preference rational spontaneously and unrequested during the interview:

> *"I would prefer the second clip's captioning [referring to CO]. It was good because I understood the game. I do not need the Play-by-Play, when live play is happening the commentators are not even talking about that, so I get distracted easily. I like [CO] because there were no captions during the live play and because I can't hear, I cannot tell if they are talking, and I could catch up on what was going on when live play was dead (replays and timeouts).*

Not all comments however, favored the CO captions. For example, one participant preferred PBP because she *"knew that everything was on it"* and that she *"wasn't missing anything"*. Another participant also preferred PBP as she found she just *"missed certain words"* and she could locate all the missed words in PBP and not in CO. This might be account for many participants describing their expectation of a delay. Participants were used to having delayed captioning, and they use it to catch up with the content they missed from watching visuals. In contrast, CO captioning did not display all spoken words and delays were reduced, which did not allow the participants to use their habit of using captioning. Participants' negative comments towards CO, suggests there is still research needed to better understand user viewing techniques to compensate for problems in CCs and the impact of CO captions on flow and speaker identification even if it has solved for the issue of game immersion and reduced distraction for D/HoH viewers. Participant comments relating to flow say *"[there] wasn't anything connecting the comments together"*, it was *"[hard to] follow along"*, and she had to do *"all the work [connecting comments] herself"*. In terms of the PBP captions, participants expressed how they *"missed [player IDs] on the CO captions"* as he was used to those caption hints when he *"wants to know a player's name"*.

Given the user familiarity with, and in some instances habituation to the status quo captioning condition (PBP), the participants can have change aversion to the CO captioning. Change aversion refers to the negative short-term reaction to changes in a service, which is natural

psychological reaction when users are used to an existing products or services (Sedley & Müller, 2013; Ram & Sheth,1989; Zajonc, 1968; James, 1890). Therefore, having any supporting comments or positive feedback on CO captioning would be interpreted as encouraging evidence for the potential utility of CO captioning. The participants' comments that illustrate their bias towards the status quo included comments that identified CO as an error or too difficult to adjust to given their expectation of a constant flow of caption text without pauses: "I think the captioning is off", "missing some actions, names or plays", "missing words made the captions confusing", "was a little bit off or delayed", "There were some words that were missing but I went with it", "I liked the second one [PBP] (…) That way I knew everything was on, I knew that I wasn't missing everything", "not captioned accurately. Still lots of missing information on what happened in background, history of scoring and lack of explaining about moments", "skipped [words] …confused", and contradicts there expectation of what captions should be: "overall it's the responsibilities of the cc captioners to hear and follow what the broadcasters are saying".

In some instances there was also all-or-nothing thinking, identifying conditions by extremes, for example for the CO condition, stating: *"they were terrible"*, *"lousy"* and *"my eyes were forced into reading the captions"*, whereas, for the PBP condition: *"Perfect viewing for me"* and *"was set up right for me to keep track of what's going on with the game"*. Not all comments were bias though, for example, one person stated: *"I didn't feel annoyed when I watched this clip with different closed captions at all"*. Any new form of captioning condition, such as CO used in this study, would disrupt the status quo, which we would expect participants to be hesitant in their acceptance of the change.

Participants also mentioned the quality factors of Closed Captioning, including synchronization delay, presentation speed, omissions of words, and spelling mistakes. The mentioned factors commonly appear in general live captioning, thus the judgment criteria seemed to be also projected onto the overall satisfaction ratings not only from the video clips they watched. Participant comments contributed insights into aspects of captioning and quality factors that they felt more strongly/confident about. A repeated quality factor raised by participants (regardless of gender or hearing status), was how captions were blocking the actual game play (or the visuals) and interfered with the game play viewing. Based on the statistical analytics, we found a correlation between the above quality factors and the overall quality satisfaction ratings. While trivial, the correlation was also found between the overall rating and the helpfulness of captioning to understand the flow and context. This supports the survey findings from 2016, as the preferred factors seemed to re-appear from our study as important and influential factors to determine the overall perceived quality of captioning.

In relation to the comprehension, there seemed to be no particular relationships associated with the captioning conditions and the comprehension scores. The overall comprehension

score was higher for HoH participants than D participants, which may be an additional evidence that supports previous studies relating to reading comprehension. When comparing the Gameplay questions and CO specific questions, from watching Hockey specifically, participants seem to have more difficulty in answering CO questions. While CO captions provide the information to correctly answer questions, participants only spent a small fraction of time reading captions. As described above, participants seem to have the 'fear of missing out', when they don't watch the visuals (non-caption area), thus we may interpret the difference in comprehension scores in the duration for visit in the caption area.

Summary of Findings

The following insights emerged from the analyses in the study:

1. Participants spent longer time watching non-caption areas than the time spent on reading captions, regardless of the caption versions. There were indications of participants spending more time on CO captions for certain cases. However, there were no statistically significant differences between CO and PBP found from the eyetracking data collected in this study.

2. There is no discernible difference in captioning preferences between the CO and PBP captions, although slight differences were found in different cases of sports and audience groups. Further research with more participants is needed to find more statistical evidence regarding the caption preferences other than simple descriptive statistics.

3. The quality factors of Closed Captioning, synchronization delay, presentation speed, speaker identification, and caption positioning were found to be important in determining the perception of quality satisfaction in CC.

4. General comprehension seems to be higher for HoH participants. However, participants seemed to have a little difficulty over answering CO specific questions.

CONCLUSIONS AND RECOMMENDATIONS

In this report, we investigated the caption viewers' eye-behaviour from watching fast-paced sports. Participants in two hearing conditions, Deaf and Hard of Hearing, were invited to watch two randomly selected sport clips with two different conditions (colour commentary only and Play-by-Play captioning).

In conclusion, it is noted that, the analytical results did not reveal any particular differences between preferences of caption versions. This suggests that there were no preferred caption

version reported or detected from eye tracking record. While participant comments and overall average caption quality satisfaction rating seemed to favor CO condition, the variability of data suggests the need for further research with a larger sample size. It is anticipated to have a convergence in preference, comprehension impact, and behavioural trend by having a larger sample size. In addition, given that the participant reactions to CO captioning should be interpreted based on the bias they might have towards the status-quo PBP captioning, having CO captioning with positive comments only found its potential impact when participants get used to watch this new captioning condition. The findings suggest another longitudinal study where participants can have sufficient time to adapt the CO captioning, then investigate the impact and difference from watching the two captioning again.

Thank you everyone for being a part of this project. We hope it will positively impact the stakeholders who are working to make captioning more enjoyable, accessible and accurate in the near future.

References

Sedley, A., & Müller, H. (2013). Minimizing change aversion for the google drive launch. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 2351-2354).

James, W. (1890). *Habit*. H. Holt.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of personality and social psychology*, *9*(2p2), 1.

Ram, S., & Sheth, J. N. (1989). Consumer resistance to innovations: the marketing problem and its solutions. *Journal of consumer marketing*, *6*(2), 5-14.

Appendices

Appendix 1: Recruitment materials

Facebook Page: https://www.facebook.com/Livesportsresearch

Project Live Captioning Canada Web Page:
http://www.livecaptioningcanada.ca/about.PBP.html

Web recruitment page: http://www.livecaptioningcanada.ca/PBP-invitation.html

Appendix 2: Consent

1. You will have the opportunity to participate in a study involving closed captioning for fast-paced live sports broadcasts for either a hockey or basketball game. You will watch two 10-minute clips with captions where one clip contains captions for the Play-by-Play descriptions and the accompanying commentary, and the second one has only captions for the commentary. You will watch these clips in a living room setting with a couch, chair and television.

2. The researchers are interested in how you consume live sports broadcasts and your opinions of the captions for the clips.

3. While watching the clips your eye movement will be tracked with eye-tracking equipment that is set up on a table with the TV. You will also be asked to answer a pre/post/midway questionnaire about your captioning viewing habits and your opinions about the captions for the two clips. You will also be asked to have a short conversation with another person about the games you watched.

4. Participation is voluntary and you can stop at any time.

5. Everything you say and do will remain confidential within the research team.

Investigators

Principal Investigator: Deborah Fels, PhD., P.Eng., Ryerson University (416) 979-5000 ext. 557619 or dfels@ryerson.ca

Maria Karam, PhD, Ryerson University, maria.karam@ryerson.ca

Christie Christelis, TSI Global, christie@tsiglobalnet.com

Margot Whitfield, Ryerson University, margot.whitfield@ryerson.ca, ext. 557110

Nazaret Fresno, nazaret.fresno@utrgv.edu

Pablo Romero-Fresco, P.Romero-Fresco@roehampton.ac.uk

Student investigator:

Jenny Leung, Ryerson University, jenny3.leung@ryerson.ca

Tatyana Kumarasamy, tkumarasamy@ryerson.ca

Study Title: Understanding user experiences of Play-by-Play captioning in fast-paced live sports

You are being asked to participate as a research participant in a research study. Before you give your consent to be a volunteer, it is important that you read the following information and ask as many questions as necessary to be sure you understand what you will be asked to do.

Purpose of the Study: The purpose of this study is to investigate users' experiences and opinions about captioning for fast-paced live sports. The nature of captioning for live sport is particularly challenging for deaf and hard-of-hearing viewers because it is sometimes difficult to follow what is going on. Captioning for this genre involves Play-by-Play descriptions as well as descriptions of the commentary provided by the announcers. We want to better understand how Deaf or Hard of Hearing (D/HoH) viewers consume live sports and their opinions of the issues, challenges and potential captioning options.

Inclusion Criteria: Participants eligible to be included are D (ASL or oral deaf)/HoH users of captions and live sports television viewers. In particular, participants must have at least one year of experience watching live hockey or basketball on television. We are recruiting 20 people per user group who are interested in this topic.

Exclusion Criteria: D/HoH people who are either not interested in live sports on television, who are not interested in either hockey or basketball and who have only watched sports for less than one year. Hearing people are also excluded.

Description of the Study

Following the consent process and consent form signing, you will be asked to carry out five main activities:

1) complete a pre-study questionnaire on paper or online (in English or ASL) that has 21 questions;

2) watch live sports clip 1 for about 10 minutes;

3) complete a mid-study questionnaire on paper or online (in English or ASL) that has 17 questions,;

4) watch live sports clip 2 that is a continuation of the game from the first clip for another 10 minutes;

5) have a 10-15-minute discussion about what happened in the game with a researcher who has also watched the clip, similar to what you might do when watching a game with a friend or family member;

6) complete a post-study questionnaire on paper or online (in English or ASL) that has 19 questions.

One of the sports clips will have captions for the Play-by-Play action, as well as the commentary, accompanying those descriptions and the other clip will have only the commentary captioned. Which one you see first will be randomly selected. The entire study including the consent process will take about 1 hour and 20 minutes.

While you are watching the game clips, an eye-tracking system will be set up on the table with the television in order to track your eye movements. No equipment will be placed on your body or head. We will need to calibrate the system to your eyes which will require a 5-minute eye movement exercise with the system, which will be placed near the television.

The entire study will be audio/video recorded in order for the researchers to analyze the data at a later date.

Risks or Discomforts: The risks associated with the study are minimal. You might feel uncomfortable or fatigued while sharing your opinions in the questionnaires or in the discussion with the researcher. If you feel tired or uncomfortable, you can take a break to rest or discontinue participation in the study either temporarily or permanently. You may feel uncomfortable being videotaped. The video recorder will be turned on during the pre-study questionnaire so that you can become used to it being on. If that does not help, you can discontinue the session or leave the study completely. The researcher may take written/typed notes in addition to any video or audio recording in order to further document the session and in case the video equipment fails. Only members of the project team will have access to all of the data, which will be stored in encrypted folders on the Ryerson Google Shared Drive at Ryerson. If you feel tired or uncomfortable, you may take a break or discontinue participation in the study either temporarily or permanently.

Benefits of the Study: There is no direct benefit to you in participating in this study, however, the information you provide will help us better understand the issues, concerns and opinions of D/HoH viewers of fast-paced live sports. This may in turn help government regulators and broadcasters improve captioning for live sports that are better suited to D/HoH needs.

Confidentiality: All data that you provide us will remain confidential and only be shared with the researchers listed on this form. No individual data will be shared with broadcasters or government agencies. Data will only be presented in summary form for publications and presentations. We will not use any recordings in any public setting. If you want the data to be removed, it will be. You have the right to review/edit the recordings or transcripts as well.

Data storage: All data will be secured on the Ryerson Google Shared Drive that encrypts all data at Ryerson University. These data will be deleted from the Ryerson Google Shared Drive or shredded if paper consent was signed, after five years. All data will be stored in its original raw format (as video, audio and text) on this Drive.

Dissemination of results: Interim and final reports from this study will be available on the IMDC website at www.imdc.ca. You may also request an individual copy of these reports or any papers published by contacting Deborah Fels at dfels@ryerson.ca

Costs and/or Compensation for Participation: There are no costs associated with your participation. You will be given an honorarium of $60 for your participation and reasonable travel costs up to $20 if incurred. If you withdraw, you will still receive the honorarium.

Voluntary Nature of Participation: Participation in this study is voluntary. Your choice of whether or not to participate in this study will not influence your future relations with Ryerson University, TSI Global, the Canadian Hard of Hearing Society or the Canadian Association of the Deaf. If you decide to participate, you are free to withdraw your consent and to stop your participation at any time without any penalty. At any particular point in the study, you may restrict the data you share with us or stop participation altogether. The final date to withdraw your data will be once we begin analyzing the data on Oct. 26, 2020.

Questions about the Study: We sincerely appreciate your cooperation. If you have any questions or concerns, please do not hesitate to call Deborah Fels at 416-979-5000 ext. 7619 or another one of the principal investigators.

Questions/Concerns about Participant's Rights: This research has been approved by the Ryerson University Research Ethics Board with Protocol #: 2020-066. The Research Ethics Board Chair may be contacted at (416) 979-5000 ext. 5741 or at rebchair@ryerson.ca should there be any complaints or concerns about the participant's rights as a participant, c/o Office

of Research Services, Ryerson University EPH-241, 350 Victoria St., Toronto, ON M5B 2K3, Tel: (416) 979-5042.

Agreement:

Your consent below indicates that you have read the information in this agreement, have had a chance to ask any questions you have about the study and know that your participation is entirely voluntary. Your consent also indicates that you agree to be in the study and have been told that you can change your mind and withdraw your consent to participate at any time.

You have been told that by consenting to participate in this study you are not giving up any of your legal rights.

I consent to audio recording (Yes or No)

I consent to video recording (Yes or No)

I consent to participate in the study (Yes or No)

Appendix 3: Video clip and captioning characteristics

Video Lengths
- NHL video total = 20.55 minutes (12.26 mins and 8.29 mins)
- NBA Video total = 35.16 minutes (16.20 mins and 18.51 mins)

Presentation speed (Words-per-minute)
- NBA-1 PBP: ~ 144 wpm
- NBA-1 CO: ~ 168 wpm
- NBA-2 PBP: ~ 129 wpm
- NBA-2 CO: ~ 138 wpm

- NHL-1 PBP: ~ 159 wpm
- NHL-1 CO: ~ 203 wpm
- NHL-2 PBP: ~170 wpm
- NHL-2 CO: ~ 217 wpm

Caption position
- NBA-CO: Lower screen
- NBA-PBP: Top of screen
- NHL-CO: Lower screen
- NHL-PBP: Lower screen

Average synchronization delay
- NHL CO: ~ 2 seconds
- NHL PBP: ~4 seconds

- NBA CO - 2 seconds
- NBA PBP - 4 seconds

Appendix 4: Pre-study questionnaire

The purpose of this questionnaire is to collect demographic information such as age and hearing status, as well as your current television viewing habits for live sports. There are 21 questions in this survey and it should take about 10 minutes to finish. Thank you in advance for participating in this study.

Study ID (Please use the participant number that was provided in your email from the researcher)

1. Are you?

- ☐ Female
- ☐ Male
- ☐ Prefer not to say
- ☐ Other:

2. In which age group are you?

- ☐ 18 - 24 years
- ☐ 25 - 34 years
- ☐ 35 - 44
- ☐ 45 - 54
- ☐ 55 - 64
- ☐ 65+

3. What is the highest level of education you have completed?

- ☐ Grade school
- ☐ Some high school
- ☐ Graduated high school
- ☐ Some college/university
- ☐ Graduated college/university
- ☐ Some post-graduate university
- ☐ Graduated post-graduate university
- ☐ Rather not say

4. What best describes your current working status?

- ☐ Full time employment (26+ hours per week)

☐ Part time employment (up to 25 hours per week)

☐ Self-employed / small business owner

☐ Student

☐ Unemployed

☐ Retired

☐ Other:

5. Which of the following best describes your hearing status?

☐ Deaf

☐ Hard of hearing

☐ Other:

6. At what age did you become deaf or hard of hearing?

☐ From birth

☐ Under the age of 2 years old

☐ From the age of 2 to 5 years

☐ From the age of 6 to 18 years

☐ From the age of 19 to 34 years

☐ From the age of 35 to 54 years

☐ From age of 55 years and above

7. Do you wear a hearing aid while watching TV?

☐ Yes

☐ No

☐ Sometimes

8. Which of the following best describes your vision capability?

	Always	Sometimes	Never
I wear glasses or contact lenses			
I have difficulty reading captions on a TV screen			
I have difficulty seeing the images on a TV screen			
I use a magnifier for reading captions on a TV screen			

I wear glasses or contact lenses			
I have difficulty reading captions on a TV screen			
I have difficulty seeing the images on a TV screen			
I use a magnifier for reading captions on a TV screen			

9. I am best at expressing myself in… Please select only one.

☐ English
☐ American Sign Language
☐ I express myself equally well in English and American Sign Language

10. Which of the following apply to you…? Please select all that apply.

☐ English is my first language
☐ I find it easy to read English
☐ I can understand American Sign Language
☐ I can understand signing in a language other than American Sign Language

11. How would you describe the speed at which you read? Please select only one.

☐ I am a fast reader
☐ My reading speed is about average
☐ I am a slow reader
☐ I really have no idea
☐ Live sports television viewing

12. How many hours a day do you typically watch television?

☐ Less than 1 hour
☐ 1 - 1.9 hours
☐ 2 - 2.9 hours
☐ 3 - 3.9 hours
☐ 4 or more hours

13. Which types of live sports have you watched in the past year? Check all that apply.

☐ Hockey

☐ Basketball

☐ Football

☐ Baseball

☐ Soccer

☐ Rugby

☐ Curling

☐ Tennis

☐ Figure Skating

☐ Other:

☐ Summer Olympics

14. Of these two sports, which one do you prefer watching live on television

☐ Hockey

☐ Basketball

15. For how long have you been watching live hockey or basketball programs?

☐ Less than 6 months

☐ 6 months to less than 1 year

☐ 1 year to less than 2 years

☐ 2 years or more

16. Which hockey team are you a fan of?

☐ Maple Leafs

☐ Blackhawks

☐ Both teams

☐ Neither

17. Which basketball team are you a fan of?

☐ Toronto Raptors

☐ Dallas Mavericks

☐ Both teams

☐ Neither

18. How often do you use captions when watching live sports?

☐ Never

☐ Hardly ever

☐ Sometimes

☐ On most occasions

☐ Always

19. How important is it to you that live sports have closed captions available?

(1=Not at all important - 7=Very important)

20. When you watch live sports, are you typically…? Please select only one.

☐ Alone

☐ With people who are deaf or hard of hearing

☐ With people who can hear

☐ With a mix of people with different hearing abilities

21. Which of the following devices have you watched live sports programs on in the past six months (i.e. wherever you have watched live sports, not only at home)? Please select all that apply.

☐ A TV set

☐ A personal computer, laptop or notebook computer

☐ A tablet computer

☐ A smart phone

Appendix 5: Post-Study questionnaire

The purpose of this questionnaire is to gather participant opinions and reflections on the live video clip that you just watched. There are 19 questions in this questionnaire. Thank you in advance for your participation.

1. Overall, how would you rate your viewing experience when watching the video clip (1=It was terrible. I didn't enjoy it at all - 5=It was great. I enjoyed it very much)

2. Overall, how satisfied are you with the quality of the live closed captioning in the video segment that you just watched? (1=Totally dissatisfied - 5=Completely satisfied)

3. Please explain why you gave the ratings from questions 1 and 2.

4. To what extent did the captions help you follow the game? (1=They did not help me at all, I wanted to turn them off - 5=They helped me follow the game. I couldn't have watched the game without them.)

5. To what extent did the images help you follow the game? (1=They did not help me at all - 5=They helped me follow the game. I couldn't have watched the game without them)

6. I did not notice time passing (1=Not at all - 7=Very much)

7. I had no difficulty concentrating on what was happening in the game (1=Not at all - 7=Very much)

8. I was totally absorbed in watching in (1=Not at all - 7=Very much)

9. I felt that my viewing experience was under control (1=Not at all - 7=Very much)

10. I was completely lost in the game/video clip (1=Not at all - 7=Very much)

12. Which of the following factors, if any, prevented you from enjoying the game (check all that apply)

   ☐ The paraphrasing changed the ideas, making the content difficult to understand.

   ☐ Missing words made the captions confusing.

   ☐ Spelling and grammar errors.

   ☐ Game captions were cut off when there was a switch to a commercial break.

   ☐ I did not have any problems with any of these factors.

   ☐ Other:

13. How did the delay between the spoken words and captions appearing on the screen influence your ability to understand what was going on in the video segment? (1= The delay was so bad that the content was almost impossible to follow - 5= I could follow exactly what was going on without difficulty).

14. How did the speed at which captions appeared on the screen influence your ability to read the captions, while at the same time giving you time to watch the images on the screen? (1= The captions were too fast. I couldn't keep up - 5= I could read the captions and watch the content without difficulty).

14. To what extent did the captions allow you to identify who was speaking at each stage in the program segments (1= I couldn't keep track of who was saying what at all - 5= I could follow exactly who was saying what without difficulty).

15. To what extent did the placement of the captions on the screen obscure information that you would have wanted to see? (1= The placement of the captions was terrible in general. Critical information was blocked - 5= The captions didn't block anything important on the screen).

16. What did you dislike about the way that the sports clip that you watched was captioned?

17. What did you like about the way that the sports clip that you watched was captioned?

Question for after the second video clip is watched

18. Which style of captioning did you prefer

- ☐ Play-by-Play only
- ☐ COentary only
- ☐ Both were acceptable

19. Briefly, explain your preference.

Appendix 6: Topics for conversation, Hockey and Basketball

After watching the videos participants were asked to take part in a conversation with the interviewer and the sessions were video recorded. The interviewer embedded certain questions based on the sport and condition the participant was in, within the conversation. This was done in a voluntary and natural way, and was used as a way to guide conversation. Specific questions about the video clips were asked to gain an understanding of participants' general knowledge of the game, these were labeled as Play-by-Play (PBP) questions. In

addition, the interviewer asked commentary only (CO) questions that could only be answered correctly if the participant read the commentary. For each of the questions researchers looked for specific answers to the questions that were established before conducting conversations. This was used as a way to grade the general understanding of participants and their use of the captioning on the screen.

Questions and sample answers

Basketball (PBP)

1. Overall, what was particularly noticeable about the game that you watched?

The Raptors made a great comeback, they worked very well and it was an exciting game.

2. What did you think of the Leaf's (or Raptors) defensive/offensive play?

The Raptors defense was great, they were getting rebounds and the defense men were mostly bench players and they played phenomenal. Lowry helped the offense, he is a top shooter, and got many 3 pointers.

3. What did you think of the Chicago's (or Dallas for basketball game) defensive/offensive plays?

Chicago had a pretty good offense, they lacked in their defense. They were not able to get to the ball fast enough and it seemed like Lowry was too smart for them and played on their weaknesses.

4. Which player(s) stood out for you and why?

Lowry, Boucher, Porzingis, Brunson, Jefferson

5. What did you think about the "reffing" for the game?

The reffing was not great they were making unusual calls and missing crucial ones

Basketball (CO)

Condition 1

1. What did Lowry (number 7) say about the comeback in the 2nd half? Why or why did it not make sense?

This moment right here. It was a team effort, the bench players …acknowledges bench players and how they were shining not just about him but about the team

2. Kyle Lowry got 30 points, 18 points in the scored 36 against Milwalkee on Nov. 2

If participant remember commentators saying this then it is correct

Condition 2

1. Do you think that the call against boucher for a foul against Brunson was a good call

Dallas didn't get a penalty and were unhappy there was a penalty called, Raptors were right to challenge it. Overall, what was particularly noticeable about the game that you watched?

Hockey (PBP)

1. **Overall, what was particularly noticeable about the game that you watched?**

Toronto lost, they were not playing well, however, chicago is a strong team because they have a great goalie

2. **What did you think of the Leaf's defensive/offensive play?**

leafs defense and goalie were weak, they had a better offense but no match for chicagos' defense

3. **What did you think of the Chicago's (or Dallas for basketball game) defensive/offensive plays?**

Chicago had a strong defense, and their goalie made many saves. Their offense was also strong, good at attacking and getting the puck up the ice to score.

4. **Which player(s) stood out for you and why?**

Anderson, Crawford, Kane, Toews, Nylander and Matthews

5. **What did you think about the "reffing" for the game?**

The refing for the game was fair, no major calls that were missed but there were some here and there, favoring Chicago slightly

Hockey (CO)

Condition 3

1. **Do you think Kane (number 88) made history from this part of the period? Why or why not**

Kane was trying to get his 1000 goal, but he didn't get it in this period maybe he did after. He is definitely going to make history just not sure in this game we have to watch the next period

2. **Does Patrick Kane (number 88) do well against Toronto? Why or why not**

Kane did well, a great leader, a great team player. He works hard to make his team successful not just himself

3. **Do you think Andersen (number 31) should be pulled for the next part of the period? Why or why not?**

Yes he should be pulled out, he was off his game he let so many goals in during the period

Condition 4

1. **Chicago and Toronto meeting for the 653rd time**

the answer is correct if they said this is correct or remember seeing this in the captioning

Appendix 7: Details of Tobii Eye tracking data collected

The eye tracking study provided us with the following variables that were used in our exploration of captions for this research. While we are only using a few of these results, we are including these in this report for completeness, as we are aiming to provide a comprehensive report of the complete methodology used in this research, towards supporting future studies aiming to expand on or explore different areas of closed captioning using eye tracking systems.

Fixation: A fixation is when a participant's gaze was detected in an area of interest for longer than about 250 milliseconds.
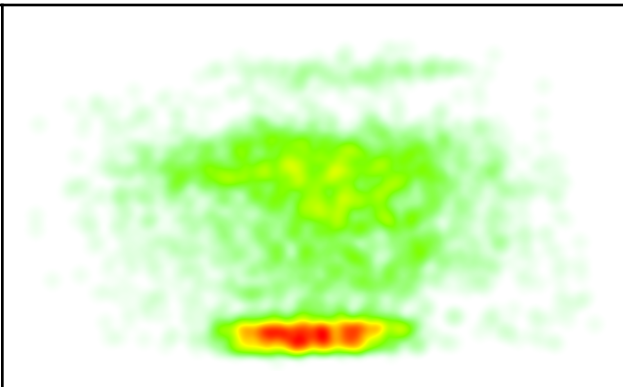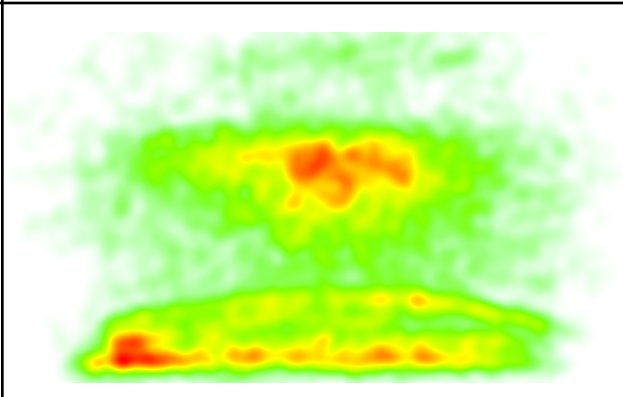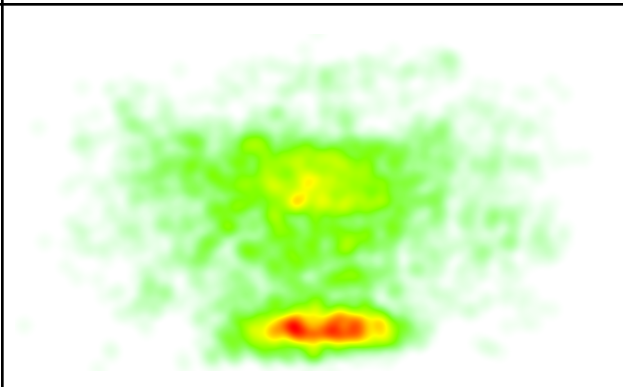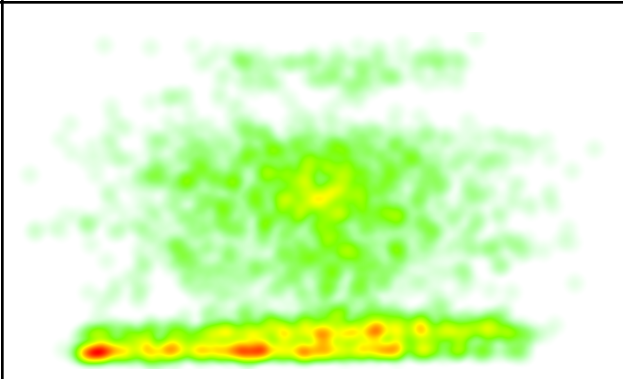
Saccade: A saccade is when eyes are detected moving between areas of interest, usually lasting less than 100 milliseconds or 1/10th of a second.

Visit: A visit corresponds to all the data between the start of the first fixation inside the defined areas of interest (AOI) to the end of the last fixation in the same AOI. **From the first fixation inside the AOI until the last fixation inside the AOI, all data is considered as part of the AOI visit (even saccades, blinks or invalid gaze data)**. AOI visit metrics allow you to measure statistics based on visits inside an AOI.

Gaze Samples: The percentage is calculated by dividing the number of correctly identified eye-tracking samples by the theoretical maximum. An eye tracker with a 50 Hz sampling frequency generates 50 samples per second. If the software can use all samples to calculate gaze points, the value in the Gaze Samples column would be 100%. However, this percentage is rare, because some samples are always lost due to the participant blinking, or looking away from the monitor in the case of a screen-based eye tracker. Blinking usually causes around 5-10% data loss during a recording.
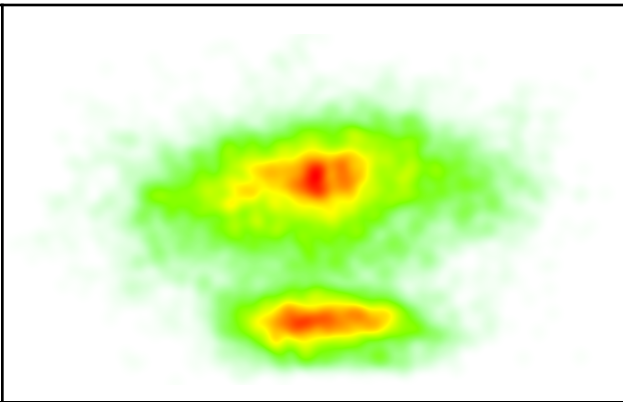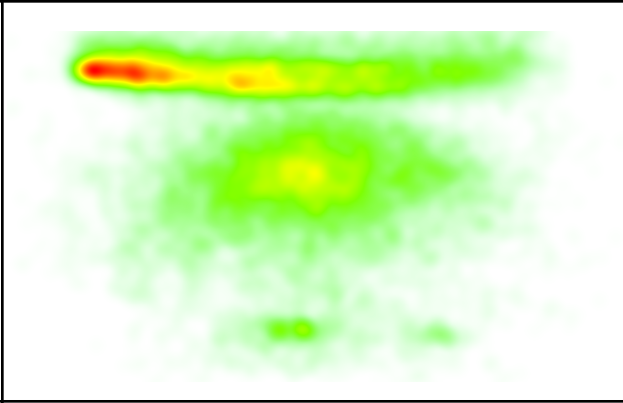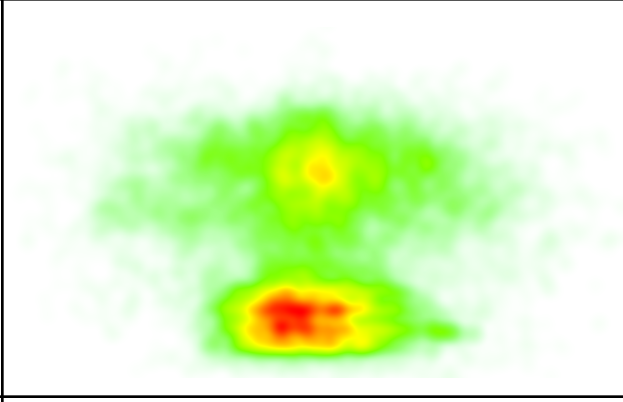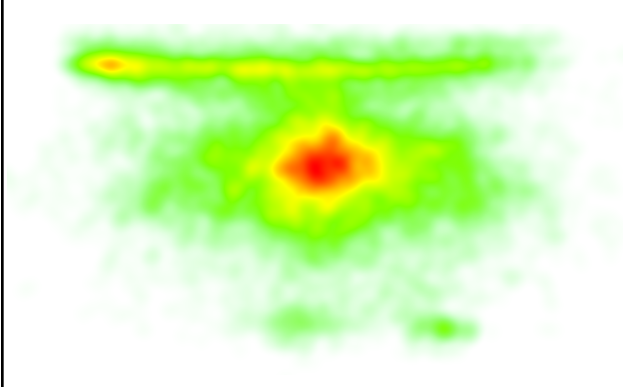
In addition, we are using the eye tracking data results to determine where the user's eye gaze, fixation and saccades are during the game time for the following conditions and recordings:

Example heat maps for Hockey:

<u>Hockey 1-CO</u>	
<u>Hockey 1-PBP</u>	
<u>Hockey 2-CO</u>	
<u>Hockey 2-PBP</u>	

Example heat maps for Basketball:

<u>Basketball 1-CO</u>	
<u>Basketball 1-PBP</u>	
<u>Basketball 2-CO</u>	
<u>Basketball 2-PBP</u>	

Appendix 8: Mean and standard deviation of the total number of visits and the total duration for visit

Sports	Hearing group	Caption version	Areas of interests	Total number of visits Mean (SD)	Normalized ratio of duration for visit Mean (SD)
Basketball	Deaf	PBP	Captions	173.20 (204.48)	12.80% (11.22%)
			Non Captions		44.20% (28.04%)
			- Game play	169.00 (135.73)	28.03% (27.78%)
		CO	Captions	114.20 (98.25)	13.24% (9.80%)
			Non Captions		47.36% (21.10%)
			- Game play	171.00 (164.15)	21.78% (17.26%)
	HoH	PBP	Captions	118.88 (48.30)	8.27% (3.28%)
			Non Captions		59.10% (24.06%)
			- Game play	275.75 (98.50)	38.94% (16.63%)
		CO	Captions	154.75 (77.93)	12.24% (7.93%)
			Non Captions		65.26% (14.71%)
			- Game play	280.50 (90.67)	39.23% (10.63%)
Hockey	Deaf	PBP	Captions	77.73 (66.61)	9.09% (12.98%)
			Non Captions		43.73% (21.45%)
			- Game play	147.27 (91.08)	22.84% (15.68%)
		CO	Captions	116.27 (100.43)	9.62% (10.15%)
			Non Captions		46.83% (21.04%)
			- Game play	202.73 (144.07)	24.68% (17.21%)
	HoH	PBP	Captions	67.00 (12.53)	10.92% (4.12%)
			Non Captions		80.1% (7.56%)
			- Game play	126.00 (20.08)	54.94% (6.50%)
		CO	Captions	90.33 (11.93)	9.85% (3.55%)
			Non Captions		72.82% (18.55%)
			- Game play	181.00 (39.51)	46.32% (25.25%)

Appendix 9: Descriptive statistics of the responses from the questionnaire

Below table displays the descriptive statistics of the questionnaire responses (1 = Most negative - 5 = Most positive).

Sports	watched condition	Captions helped follow the game Mean (SD)	Captions helped follow the game Overall Mean (SD)	Images helped follow the game Mean (SD)	Images helped follow the game Overall Mean (SD)
Basketball	PBP (first)	3.50 (1.38)	3.00 (1.55)	4.43 (0.53)	4.33 (0.65)
	PBP (second)	2.40 (1.67)		4.20 (0.84)	
	CO (first)	3.60 (1.14)	3.58 (1.16)	4.40 (0.89)	4.50 (0.67)
	CO (second)	3.57 (1.27)		4.57 (0.53)	
Hockey	PBP (first)	3.29 (0.95)	3.14 (1.10)	3.29 (1.38)	3.29 (1.49)
	PBP (second)	3.00 (1.29)		3.29 (1.70)	
	CO (first)	3.57 (1.27)	3.50 (1.22)	4.43 (0.79)	4.07 (1.00)
	CO (second)	3.43 (1.27)		3.71 (1.11)	

Sports	watched condition	Hearing group	Captions helped follow the game Mean (SD)	Images helped follow the game Mean (SD)
Basketball	PBP	Deaf	4.00 (1.41)	4.25 (0.25)
		Hard of Hearing	2.43 (1.40)	4.38 (0.74)
	CO	Deaf	4.25 (0.96)	4.25 (0.50)
		Hard of Hearing	3.25 (1.16)	4.63 (0.74)
Hockey	PBP	Deaf	3.09 (1.04)	3.00 (1.48)
		Hard of Hearing	3.33 (1.53)	4.33 (1.15)
	CO	Deaf	3.73 (1.27)	4.00 (1.00)
		Hard of Hearing	2.67 (0.58)	4.33 (1.15)

Sports	watched condition	Delay influenced understanding the video Mean (SD)	Delay influenced understanding the video Overall Mean (SD)	Caption speed influenced understanding captions Mean (SD)	Caption speed influenced understanding captions Overall Mean (SD)	Caption helped identify speakers Mean (SD)	Caption helped identify speakers Overall Mean (SD)	Caption position didn't obscure information Mean (SD)	Caption position didn't obscure information Overall Mean (SD)
Basketball	PBP (first)	3.29 (0.95)	3.25 (1.22)	3.57 (0.98)	3.50 (1.17)	2.43 (1.51)	2.75 (1.71)	4.14 (1.21)	3.67 (1.30)
	PBP (second)	3.20 (1.64)		3.40 (1.52)		3.20 (2.05)		3.00 (1.22)	
	CO (first)	4.00 (1.00)	3.67 (1.37)	4.00 (0.71)	4.17 (0.72)	4.20 (0.84)	3.83 (0.94)	4.60 (0.55)	4.42 (0.79)
	CO (second)	3.43 (1.62)		4.29 (0.76)		3.57 (0.98)		4.29 (0.95)	
Hockey	PBP (first)	3.29 (1.25)	3.36 (1.08)	3.86 (0.38)	3.71 (0.73)	3.14 (1.57)	2.71 (1.38)	3.86 (1.21)	3.57 (1.22)
	PBP (second)	3.43 (0.98)		3.57 (0.98)		2.29 (1.11)		3.29 (1.25)	
	CO (first)	3.43 (1.27)	3.64 (1.15)	3.00 (0.82)	3.29 (1.07)	2.57 (1.27)	2.57 (1.22)	3.86 (1.35)	3.93 (1.33)
	CO (second)	3.86 (1.07)		3.57 (1.27)		2.57 (1.27)		4.00 (1.41)	

Sports	Caption condition	Hearing group	Delay influenced understanding the video	Caption speed influenced understanding the video	Caption helped identify speakers	Caption position didn't obscure information
Basketball	PBP	Deaf	3.75 (0.96)	3.50 (1.29)	2.75 (1.50)	4.00 (1.41)
		Hard of Hearing	3.00 (1.31)	3.50 (1.20)	2.75 (1.91)	3.50 (1.31)
	CO	Deaf	2.75 (0.50)	4.25 (0.50)	3.25 (0.96)	4.00 (0.82)
		Hard of Hearing	4.13 (1.46)	4.13 (0.83)	4.13 (0.83)	4.63 (0.74)
Hockey	PBP	Deaf	3.18 (1.08)	3.64 (0.81)	2.64 (1.29)	3.27 (1.19)
		Hard of Hearing	4.0 (1.0)	4.0 (0.0)	3.0 (2.0)	4.67 (0.58)
	CO	Deaf	4.0 (1.0)	3.45 (1.13)	2.73 (1.27)	4.00 (1.34)
		Hard of Hearing	2.33 (0.58)	2.67 (0.58)	2.00 (1.00)	3.67 (1.53)

Below tables displays the descriptive statistics of the questionnaire responses (1 = Not at all - 7 = Very much).

Sports	watched condition	I did not notice time passing Mean (SD)	I did not notice time passing Overall Mean (SD)	I had no difficulty concentrating on what was happening in the game Mean (SD)	I had no difficulty concentrating on what was happening in the game Overall Mean (SD)	I was totally absorbed in watching in Mean (SD)	I was totally absorbed in watching in Overall Mean (SD)	I felt that my viewing experience was under control Mean (SD)	I felt that my viewing experience was under control Overall Mean (SD)	I was completely lost in the game/video clip Mean (SD)	I was completely lost in the game/video clip Overall Mean (SD)
Basket ball	PBP (first)	3.29 (1.60)	3.08 (1.78)	4.71 (1.80)	4.58 (1.83)	4.71 (1.60)	5.08 (2.02)	5.14 (1.35)	5.00 (1.76)	2.86 (1.77)	3.00 (2.26)
	PBP (second)	2.80 (2.17)		4.40 (2.07)		5.60 (2.61)		4.80 (2.39)		3.20 (3.03)	
	CO (first)	4.00 (2.35)	3.75 (2.05)	5.60 (2.61)	4.92 (2.50)	5.80 (0.45)	5.50 (2.02)	6.80 (0.45)	5.92 (1.78)	2.20 (2.68)	3.17 (2.62)
	CO (second)	3.57 (1.99)		4.43 (2.51)		4.57 (2.23)		5.29 (2.14)		3.86 (2.54)	
Hockey	PBP (first)	4.86 (1.46)	4.00 (2.00)	5.00 (2.16)	3.79 (2.19)	5.57 (1.27)	4.93 (2.06)	5.57 (1.72)	5.36 (1.65)	2.14 (1.77)	4.07 (2.37)
	PBP (second)	3.14 (2.19)		2.57 (1.51)		4.29 (2.56)		5.14 (1.68)		6.00 (0.58)	
	CO (first)	2.86 (2.04)	3.57 (1.95)	3.14 (2.27)	3.93 (2.34)	5.86 (1.68)	5.86 (1.23)	6.00 (1.41)	5.71 (1.38)	3.71 (2.43)	2.79 (2.04)
	CO (second)	4.29 (1.70)		4.71 (2.29)		5.86 (0.69)		5.43 (1.40)		1.86 (1.07)	

Sports	Watched condition	Hearing condition	I did not notice time passing Mean (SD)	I had no difficulty concentrating on what was happening in the game Mean (SD)	I was totally absorbed in watching in Mean (SD)	I felt that my viewing experience was under control Mean (SD)	I was completely lost in the game/video clip Mean (SD)
Basketball	PBP	Deaf	4.50 (2.38)	4.50 (2.38)	4.50 (2.38)	4.50 (2.38)	1.75 (1.50)
		Hard of Hearing	3.25 (1.58)	4.63 (1.69)	5.38 (1.92)	5.25 (1.49)	3.63 (2.39)
	CO	Deaf	3.25 (2.63)	4.25 (2.36)	6.00 (0.82)	5.75 (1.26)	1.25 (0.50)
		Hard of Hearing	4.00 (1.85)	5.25 (2.66)	5.25 (2.43)	6.00 (2.07)	4.13 (2.75)
Hockey	PBP	Deaf	3.82 (1.72)	3.73 (2.10)	4.45 (2.07)	5.18 (1.78)	3.91 (2.26)
		Hard of Hearing	4.67 (3.21)	4.0 (3.0)	6.67 (0.58)	6.0 (1.0)	4.67 (3.21)
	CO	Deaf	3.73 (2.00)	4.27 (2.37)	6.09 (0.94)	5.73 (1.42)	2.64 (2.11)
		Hard of Hearing	3.0 (2.0)	2.67 (2.08)	5.0 (2.0)	5.67 (1.53)	3.33 (2.08)