

AI DASH CAM PERFORMANCE BENCHMARK TESTING

Final Report

Submitted June 30, 2023

Revised August 17, 2023

Revision #2 October 23, 2025

Matthew C. Camden, Susan A. Soccolich, Christie Ridgeway, R. Lucas Parks, and Richard J. Hanowski

Virginia Tech Transportation Institute
3500 Transportation Research Plaza
Blacksburg, VA 24061



Executive Summary

Project Objective

The objective of this project was to benchmark the performance of three driver monitoring systems (DMSs): Motive DC-54, Samsara HW-CM32, and Lytx DriveCam SF400. Specifically, this study was designed to assess the following:

1. How often each of the DMSs produced a successful alert in response to six event types. A successful alert is operationally defined as the correct alert was generated when the behavior was performed.
2. User ratings in terms of video/image quality, field of view ratings, alert quality, hardware quality, and overall ranking.

Methods

The study was conducted in two phases. Phase One was an assessment to benchmark the performance of the three DMSs. This phase involved testing the ability of each system to successfully capture and alert unsafe driver behavior. One driver was allocated as the sole test-driver for the duration of testing to ensure experimental control over individual driver characteristics. The test driver executed six unsafe driving behaviors (rolling stops, no seatbelt, close following, phone calls, texting, and lap texting) under three lighting conditions (day, twilight, and night) and three windshield positions for each system. The experimental protocol was developed using each of the DMS companies' existing thresholds for each behavior. Testing took place on VTTI's Smart Road test track. In-cab alerts, each respective system's dashboard alerts, and time to alert were recorded for every trial.

Phase Two of the study, the user performance benchmarking phase, acquired feedback from heavy vehicle drivers regarding various attributes of each system's quality. This survey was distributed through connections with commercial motor vehicle (CMV) fleet safety managers and carrier representatives. One hundred eighty-eight (188) CMV drivers with an active commercial driver's license (CDL-A) participated in the survey.

Analysis

Analysis for Phase One included an assessment of in-cab alert success, dashboard alert success, and time to alert for each of the six behaviors. A logistic regression model was built for the six driving tasks to test for likelihood of alert success. The response variable was binary (alert success/alert failure), and the explanatory variables consisted of system type (Motive, Samsara, Lytx), system rotation (1, 2, 3), and time of day (day, twilight, and night). For time to alert, each task was analyzed using a Kruskal-Wallis test to test for significant differences in overall alert times. If significant, Dunn's tests with Bonferroni corrections were used to compare two systems at a time for significant differences. Dashboard notification success was summarized using frequency counts.

The survey in Phase Two consisted of three main question types: Likert scale ratings, ordering, and free response questions. For the Likert scale ratings and overall order questions, descriptive statistics were used to understand the distribution of positive evaluations across all surveyed DMS qualities. The data was analyzed using generalized linear models and logistic regression models. A content analysis was performed on the free response questions to identify recurring themes and patterns. The survey responses were further stratified based on DMS experience to compare distributions of non-Motive users' and current Motive users' ratings.

Results

The following results do not provide a comprehensive review of all analyses; rather, they provide support for key conclusions relevant to Motive. Comprehensive analyses are found in the Appendix.

In-cab alerts

Prevalence of in-cab alerts for the cell phone task, lap texting task, and close following task were calculated to understand the likelihood of a successful alert. The Motive system was 3.36 times more likely to produce an alert for the phone call task than Lytx (95% CI = 2.01, 5.58) and 2.47 times more likely than Samsara (95% CI = 1.65, 3.69). The Motive system was 7 times more likely to produce an alert for the lap texting task than Lytx (95% CI = 2.27, 21.62) and 3.50 times more likely than Samsara (95% CI = 1.58, 7.75). The Motive system was 2.89 times more likely to produce an alert for the close following task than Lytx (95% CI = 1.56, 5.34) and 3.71 times more likely than Samsara (95% CI = 1.83, 7.53). For the rolling stop task, a significant difference in successful alert distribution exists between Motive and Samsara [$\chi^2(2) = 48.75$, p -value < 0.0001]. Lytx was not evaluated for rolling stop alerts due to a known issue with false alerts. For the no seatbelt task, a significant difference in successful alert distribution exists between Motive and Samsara and Lytx and Samsara [$\chi^2(2) = 117.00$, p -value < 0.0001]. However, VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled.

Dashboard Alerts

For Motive, "In-Cab Alert or Dashboard Alert" was the most frequent categorization across all task types whereas the most frequent option for Samsara and Lytx was "No In-Cab or Dashboard Alert" for all task types except no seatbelt where Lytx's most common alert type was "In-Cab Only." It is also important to note that the Lytx system was not evaluated for rolling stop tasks as the alert was deactivated by the company for all customers during the duration of the study. Additionally, VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled.

Time to Alert

The time to alert results assessed the time between task start time and a successful alert. There were significant differences in time to alert for phone call, 45-degree texting, close following, and no seatbelt tasks between Motive and at least one other system. Motive had a significantly faster time to alert for the phone call task and 45-degree texting task when compared to Lytx. Motive also had significantly faster time to alert for the close following task when compared to Samsara. Lastly, Motive had a significantly faster time to alert for the successful no seatbelt trials when compared to Lytx.

Survey Likert Ratings

The logistic regression model indicated a significant difference for the likelihood of being ranked 4 or higher for nighttime video quality between Motive and both Lytx and Samsara. A significant difference was also found for day image quality between Motive and Lytx, for night image quality between Motive and Samsara, and for seat belt audio alert quality between Motive and both Lytx and Samsara.

For nighttime video quality, Motive was 1.72 times more likely to be rated a 4 or higher compared to Lytx [95% CI = (1.13, 2.62)] and Motive was 2.96 times more likely to be rated a 4 or higher compared to Samsara for nighttime video quality [95% CI = (1.89, 4.65)]. For daytime image quality, Lytx was 1.55 times more likely to be rated 4 or higher compared to Motive for daytime image quality [95% CI = (1.03, 2.33)]. For nighttime image quality, Motive was 2.55 times more likely to be rated 4 or higher compared to Samsara [95% CI = (1.54, 4.24)]. For seatbelt audio alert quality Lytx was 1.65 times more likely to be rated 4 or higher compared to Motive [95% CI = (1.07, 2.54)] and Samsara was 2.05 times more likely to be rated 4 or higher compared to Motive [95% CI = (1.34, 3.15)].

Survey Overall Rankings

The logistic regression model was used to assess the likelihood for one system to be ranked first when compared to the other two. A significant difference was found for daytime and nighttime video quality between Motive and both Lytx and Samsara. Additionally, a significant difference was found for nighttime image quality between Motive and Samsara, cell phone audio alert quality between Motive and Lytx, and close following audio alert quality between Motive and both Lytx and Samsara.

For daytime video quality, Motive was 1.75 times more likely to be ranked first compared to both Lytx and Samsara ([95% CI = (1.16, 2.66)] and [95% CI = (1.15, 2.66)] respectively). For nighttime video quality Motive was 1.64 times more likely to be ranked first compared to Lytx [95% CI = (1.08, 2.49)] and Motive was 1.83 times more likely to be ranked first compared to Samsara [95% CI = (1.20, 2.78)]. Additionally, Motive was 1.80 times more likely to be ranked first compared to Samsara for nighttime image quality [95% CI = (1.18, 2.75)]. For cell phone audio alert quality, Motive was 1.79 times more likely to be ranked first when compared to Lytx [95% CI = (1.17, 2.73)]. For close following audio alert quality, Motive was 1.56 times more likely to be ranked first compared to Lytx [95% CI = (1.01, 2.41)] and 1.73 times more likely to be ranked first compared to Samsara [95% CI = (1.12, 2.68)].

Content Analysis

The content analysis identified key themes across participant responses for video quality, image quality, audio alert quality, and hardware quality. Video and image clarity was the most commonly cited topic under video and image quality questions. Users preferred visual media that had clear imagery and no blur, glare, or graininess. For audio alert quality, users mentioned volume most often. Many comments noted the alerts were too quiet, whereas some were too loud. It is important to note that participants of the survey used personal devices to take the survey, so this topic could potentially be a result of the volume settings on the participants' device. Lastly, for hardware, participants placed the most emphasis on the size of the units. The commercial drivers were concerned about the obstruction of a dashcam unit and preferred smaller units that least inhibited the view out of the front windshield.

Conclusions

While the study provided a scientific benchmarking effort between the three DMSs, it is important to note the study assumptions and limitations. These limitations inform the accurate interpretation of the presented results and assess the ability of the results to be extrapolated further. VTTI does not endorse any dash cam used or not used in this study. The results obtained are based on controlled testing in specific conditions using only one driver. VTTI only assessed how often alerts were generated during testing. Anytime a “successful” alert is referenced, it means that an alert was produced for the appropriate behavior being tested; it does not imply the alert changed a driver’s behavior. It is important to note that VTTI did not assess:

- 1) The ability of the systems to detect unsafe driving behaviors.
- 2) Any safety benefits of using a DSM or efficacy of any technology.
- 3) Whether any dash cam features are “better” than others.
- 4) In-depth false alerts.

The following conclusions highlight significant results from the two-phase study that pertain to the Motive system.

- Motive was more likely to issue a successful in-cab alert compared to Lytx and Samsara for phone call, lap texting, 45-degree texting, close following, and rolling stop trials across all lighting conditions.
- For Motive, *In-Cab Alert or Dashboard Alert* was the most frequent result for all tasks (phone calls, lap texting, 45-degree texting, close following, rolling stops, and no seatbelt).
- Motive had significantly shorter time to alert compared to Lytx for phone call tasks, 45-degree texting, and no seatbelt trials and significantly shorter time to alert for close following trials compared to Samsara.
- Motive was more likely to be ranked number one for daytime and nighttime video quality when compared to Lytx and Samsara.
- Motive was more likely to be ranked number one for nighttime image quality when compared to Samsara.
- Motive was more likely to be ranked number one for cell phone alert quality when compared to Lytx and more likely to be ranked first for close following alert quality when compared to both Lytx and Samsara.
- Motive had a significantly higher rating for hardware design when compared to Samsara.

Disclaimer

This research was funded by Motive. Although care has been taken to ensure complete and accurate study results, recipients of this document accept the possibility of unintended errors or omissions. The results expressed in this document are based on data collected from March 13th, 2023, through May 2nd, 2023, under specific testing conditions that may not apply to other circumstances. VTTI does not endorse any products included or not included in this study.

Table of Contents

EXECUTIVE SUMMARY	0
Project Objective	2
Methods	2
Analysis.....	2
Results.....	3
In-cab alerts	3
Dashboard Alerts	3
Time to Alert.....	3
Survey Likert Ratings	4
Survey Overall Rankings	4
Content Analysis.....	4
Conclusions	5
TABLE OF CONTENTS	7
LIST OF FIGURES	9
LIST OF TABLES	12
INTRODUCTION.....	18
Background	18
Introduction	18
METHODS	20
Phase One: Technical Performance Benchmarking.....	20
DMS Acquisition & Installation	20
Experimental Design.....	22
Analysis	30
Phase Two: User Performance Benchmarking	32
Survey Content	33
Analysis	36
RESULTS	39
Phase One: Technical Performance Benchmarking.....	39
In-cab Alerts: Outgoing Phone Call.....	39
In-cab Alert: Lap Texting	40
In-cab Alert: 45-Degree Texting.....	41

In-cab Alert: Close Following	42
In-cab Alert: Rolling Stop.....	43
In-cab Alert: No Seatbelt	44
Dashboard Alerts	45
Time to Alert.....	47
Phase Two: User Performance Benchmarking	49
Video Quality.....	49
Image Quality	52
Alert Quality	54
Hardware Design	56
LIMITATIONS.....	59
CONCLUSIONS	61
Opportunities	62
REFERENCES.....	64
APPENDIX.....	65
Additional Resources for Phase One: Technical Performance Benchmarking.....	65
Additional Resources for Phase Two: User Performance Benchmarking.....	67
Video Quality	67
Image Quality.....	78
Alert Quality.....	89
Hardware Design Quality.....	105

List of Figures

Figure 1 - The three DMSs used for benchmarking (Left to right: Motive, Samsara, and Lytx.) 18

Figure 2 - Example of ideal positioning for testing. 20

Figure 3 - Example of three positions used during testing. 21

Figure 4 - The three rotations used during testing: (a) Rotation 1: Motive, Samsara, Lytx; (b) Rotation 2: Lytx, Motive Samsara; (c) Rotation 3: Samsara, Lytx, Motive. 22

Figure 5 - Map of the Virginia Smart Roads. 23

Figure 6 - Example of driver during on-road testing for outgoing phone call..... 25

Figure 7 - Example of driver during on-road testing for texting condition (45-degrees). 26

Figure 8 - Example of driver during on-road testing for lap texting condition. 26

Figure 9 - Example of lead vehicle during on-road testing for close following. 28

Figure 10 - Example of driver during on-road testing for no seatbelt condition. 29

Figure 11 - Example of forward view during on-road testing for the rolling stop condition. 30

Figure 12 – Example of alert chart used by experimenter during testing. 31

Figure 13 - Example of video questions viewed from a mobile phone. 33

Figure 14 - Example of image questions viewed from a mobile phone. 34

Figure 15 - Example of audio alert questions viewed from a mobile device. 35

Figure 16 - Example of hardware questions viewed from a mobile device..... 36

Figure 17 - Successful phone call alert percentages during on-road testing across system, time of day, and rotation..... 40

Figure 18 - Successful lap texting alert percentages during on-road testing across system, time of day, and rotation..... 41

Figure 19 - Successful 45-degree texting alert percentages during on-road testing across system, time of day, and rotation. 42

Figure 20 - Successful close following alert percentages during on-road testing across system, time of day, and rotation. 43

Figure 21 - Successful rolling stop alert percentages during on-road testing across system, time of day, and rotation..... 44

Figure 22 - Successful no seatbelt alert percentages during on-road testing across system, time of day, and rotation..... 45

Figure 23. Likert scale ratings for daytime video quality comparing current Motive users’ opinions of all three systems..... 69

Figure 24 - Likert scale ratings for daytime video quality comparing current Motive users and non-users. 70

Figure 25. Likert scale ratings for nighttime video quality comparing current Motive users' opinions of all three systems.	72
Figure 26 - Likert scale ratings for nighttime video quality comparing current Motive users and non-users.	73
Figure 27 - Ranking for daytime video quality comparing current Motive users' opinions of all three systems.	75
Figure 28 - Ranking for daytime video quality comparing current Motive users and non-users.	75
Figure 29 - Ranking for nighttime video quality comparing current Motive users' opinions of all three systems.	77
Figure 30 - Ranking for nighttime video quality comparing current Motive users and non-users.	78
Figure 31 - Likert scale ratings for daytime image quality comparing current Motive users' opinions of all three systems.	80
Figure 32 - Likert scale ratings for daytime image quality comparing current Motive users and non-users.	81
Figure 33 - Likert scale ratings for nighttime image quality comparing current Motive users' opinions of all three systems.	83
Figure 34 - Likert scale ratings for nighttime image quality comparing current Motive users and non-users.	84
Figure 35 - Ranking for daytime image quality comparing current Motive users' opinions of all three systems.	85
Figure 36 - Ranking for daytime image quality comparing current Motive users and non-users.	86
Figure 37 - Ranking for nighttime image quality comparing current Motive users' opinions of all three systems.	88
Figure 38 - Ranking for nighttime image quality comparing current Motive users and non-users.	88
Figure 39 - Likert scale ratings for cell phone use alert quality comparing current Motive users' opinions of all three systems.	91
Figure 40 – Likert scale ratings for cell phone use alert quality comparing current Motive users and non-users.	92
Figure 41 - Likert scale ratings for seatbelt use alert quality comparing current Motive users' opinions of all three systems.	94
Figure 42 – Likert scale ratings for seatbelt use alert quality comparing current Motive users and non-users.	95
Figure 43 - Likert scale ratings for close following distance alert quality comparing current Motive users' opinions of all three systems.	97
Figure 44 – Likert scale ratings for close following distance alert quality comparing current Motive users and non-users.	98

Figure 45 - Ranking for cell phone use alert quality comparing current Motive users’ opinions of all three systems..... 99

Figure 46 - Ranking for cell phone use alert quality comparing current Motive users and non-users..... 100

Figure 47 - Ranking for seatbelt use alert quality comparing current Motive users’ opinions of all three systems..... 102

Figure 48 - Ranking for seatbelt use alert quality comparing current Motive users and non-users..... 102

Figure 49 - Ranking for close following distance alert quality comparing current Motive users’ opinions of all three systems..... 104

Figure 50 - Ranking for close following distance alert quality comparing current Motive users and non-users..... 105

Figure 51 – Likert scale ratings for hardware design quality comparing current Motive users’ opinions of all three systems..... 107

Figure 52 - Likert scale ratings for hardware design quality comparing current Motive users and non-users..... 108

Figure 53 - Ranking for hardware design quality comparing current Motive users’ opinions of all three systems..... 109

Figure 54 - Ranking for hardware design quality comparing current Motive users and non-users..... 110

List of Tables

Table 1 – Handheld Cell Phone Alerts for each DMS.....	24
Table 2 - Close Following Alerts for each DMS	27
Table 3 - No Seatbelt Alerts for each DMS.....	28
Table 4 - Rolling Stop Alerts for each DMS	29
Table 5 - Successful Phone Call Alerts by System Type and Time of Day	39
Table 6 - Successful Lap Texting Alerts by System Type and Time of Day	40
Table 7 - Successful 45-Degree Texting Alerts by System Type and Time of Day.....	41
Table 8 - Successful Close Following Alerts by System Type and Time of Day	42
Table 9 - Successful Rolling Stop Alerts by System Type and Time of Day.....	44
Table 10 - Successful No Seatbelt Alerts by System Type and Time of Day	45
Table 11 - Overall Dashboard Alert Status for Phone Call Task.....	46
Table 12 - Overall Dashboard Alert Status for Lap Texting Task.....	46
Table 13 - Overall Dashboard Alert Status for 45-Degree Texting Task.....	46
Table 14 - Overall Dashboard Alert Status for Close Following Task.....	46
Table 15 - Overall Dashboard Alert Status for Rolling Stop Task.....	47
Table 16 - Overall Dashboard Alert Status for No Seatbelt Task.....	47
Table 17 - Overall Time to Alert for Each System Across All Task Types	48
Table 18 – Overall Time to Alert Results of Post-hoc Dunn's Tests for Significance	49
Table 19 - Video Quality Likert Ratings by Time of Day	50
Table 20 - Video Quality Overall System Rankings by Time of Day	50
Table 21 - Driver Free Response Themes for Video Quality	51
Table 22 - Image Quality Likert Rankings by Time of Day.....	52
Table 23 - Image Quality Overall System Rankings by Time of Day.....	53
Table 24 - Driver Free Response Themes for Image Quality.....	53
Table 25 - Alert Quality Likert Ratings by Alert Type	54
Table 26 - Alert Quality Overall System Rankings by Alert Type.....	55
Table 27 - Driver Free Response Themes for Alert Quality.....	55
Table 28 - Hardware Quality Likert Ratings	57
Table 29 - Hardware Quality Overall System Rankings	57
Table 30 - Driver Free Response Themes for Hardware Quality	57
Table 31 - Type 3 Analysis of Effects for Phone Call Task Model.....	65

Table 32 - Analysis of Maximum Likelihood Estimates for Phone Call Task Model	65
Table 33 - Type 3 Analysis of Effects for Lap Texting Task Model.....	65
Table 34 - Analysis of Maximum Likelihood Estimates for Lap Texting Task.....	66
Table 35 - Type 3 Analysis of Effects for 45-Degree Texting Task Model.....	66
Table 36 - Analysis of Maximum Likelihood Estimates for 45-Degree Texting Task	66
Table 37 - Type 3 Analysis of Effects for Close Following Task Model.....	67
Table 38 - Analysis of Maximum Likelihood Estimates for Close Following Task	67
Table 39 – Frequency counts for Likert scale ratings of daytime video quality.....	67
Table 40 - Model results for Likert scale ratings for daytime video quality assessed as continuous variables, with parameter estimate values.	68
Table 41 - Model results for Likert scale ratings of daytime video quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.....	68
Table 42 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in daytime video quality.....	68
Table 43 – Frequency counts for Likert scale ratings of daytime video quality for current Motive users.	69
Table 44 – Frequency counts for Likert scale ratings of daytime video quality for current Motive users compared to non-users.....	69
Table 45 – Frequency counts for Likert scale ratings of nighttime video quality.	70
Table 46 - Model results for Likert scale ratings for nighttime video quality assessed as continuous variables, with parameter estimate values.	71
Table 47 - Model results for Likert scale ratings of nighttime video quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.	71
Table 48 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in nighttime video quality.....	71
Table 49 – Frequency counts for Likert scale ratings of nighttime video quality for current Motive users.	72
Table 50 – Frequency counts for Likert scale ratings of nighttime video quality for current Motive users compared to non-users.....	72
Table 51 – Frequency counts for overall order ranking of daytime video quality.	73
Table 52 - Model results for overall order ranking of daytime video quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.	73
Table 53 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in daytime video quality.....	74
Table 54 – Frequency counts for overall order ranking of daytime video quality for current Motive users.	74

Table 55 – Frequency counts for ranking of daytime video quality for current Motive users compared to non-users.	75
Table 56 – Frequency counts for overall order ranking of nighttime video quality.	76
Table 57 - Model results for overall order ranking of nighttime video quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.	76
Table 58 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in nighttime video quality.	76
Table 59 – Frequency counts for overall order ranking of nighttime video quality for current Motive users.	77
Table 60 – Frequency counts for ranking of nighttime video quality for current Motive users compared to non-users.	77
Table 61 – Frequency counts for Likert scale ratings of daytime image quality.	78
Table 62 - Model results for Likert scale ratings for daytime image quality assessed as continuous variables, with parameter estimate values.	79
Table 63 - Model results for Likert scale ratings of daytime image quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.	79
Table 64 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in daytime image quality.	79
Table 65 – Frequency counts for Likert scale ratings of daytime image quality for current Motive users.	80
Table 66 – Frequency counts for Likert scale ratings of daytime image quality for current Motive users compared to non-users.	80
Table 67 – Frequency counts for Likert scale ratings of nighttime image quality.	81
Table 68 - Model results for Likert scale ratings for nighttime image quality assessed as continuous variables, with parameter estimate values.	82
Table 69 - Model results for Likert scale ratings of nighttime image quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.	82
Table 70 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in nighttime image quality.	82
Table 71 – Frequency counts for Likert scale ratings of nighttime image quality for current Motive users.	83
Table 72 – Frequency counts for Likert scale ratings of nighttime image quality for current Motive users compared to non-users.	83
Table 73 – Frequency counts for overall order ranking of daytime image quality.	84
Table 74 - Model results for overall order ranking of daytime image quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.	84
Table 75 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in daytime image quality.	85

Table 76 – Frequency counts for overall order ranking of daytime image quality for current Motive users.	85
Table 77 – Frequency counts for ranking of daytime image quality for current Motive users compared to non-users.	86
Table 78 – Frequency counts for overall order ranking of nighttime image quality.	86
Table 79 - Model results for overall order ranking of nighttime image quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.	87
Table 80 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in nighttime image quality.	87
Table 81 – Frequency counts for overall order ranking of nighttime image quality for current Motive users.	87
Table 82 – Frequency counts for ranking of nighttime image quality for current Motive users compared to non-users.	88
Table 83 – Frequency counts for Likert scale ratings of cell phone use alert quality.	89
Table 84 - Model results for Likert scale ratings for cell phone use alert quality assessed as continuous variables, with parameter estimate values.	89
Table 85 - Model results for Likert scale ratings of cell phone use alert quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.	89
Table 86 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in cell phone use alert quality.	90
Table 87 – Frequency counts for Likert scale ratings of cell phone use alert quality for current Motive users.	90
Table 88 – Frequency counts for Likert scale ratings of cell phone use alert quality for current Motive users compared to non-users.	91
Table 89 – Frequency counts for Likert scale ratings of seatbelt use alert quality.	92
Table 90 - Model results for Likert scale ratings for seatbelt use alert quality assessed as continuous variables, with parameter estimate values.	93
Table 91 - Model results for Likert scale ratings of seatbelt use alert quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.	93
Table 92 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in seatbelt use alert quality.	93
Table 93 – Frequency counts for Likert scale ratings of seatbelt use alert quality for current Motive users.	94
Table 94 – Frequency counts for Likert scale ratings of seatbelt use alert quality for current Motive users compared to non-users.	94
Table 95 – Frequency counts for Likert scale ratings of close following distance alert quality. .	95
Table 96 - Model results for Likert scale ratings for close following distance alert quality assessed as continuous variables, with parameter estimate values.	96

Table 97 - Model results for Likert scale ratings of close following distance alert quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values. 96

Table 98 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in close following distance alert quality. 96

Table 99 – Frequency counts for Likert scale ratings of close following distance alert quality for current Motive users. 97

Table 100 – Frequency counts for Likert scale ratings of close following distance alert quality for current Motive users compared to non-users. 97

Table 101 – Frequency counts for overall order ranking of cell phone use alert quality. 98

Table 102 - Model results for overall order ranking of cell phone use alert quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values. 98

Table 103 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in cell phone use alert quality. 99

Table 104 – Frequency counts for overall order ranking of cell phone use alert quality for current Motive users. 99

Table 105 – Frequency counts for ranking of cell phone use alert quality for current Motive users compared to non-users. 100

Table 106 – Frequency counts for overall order ranking of seatbelt use alert quality. 100

Table 107 - Model results for overall order ranking of seatbelt use alert quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values. 101

Table 108 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in seatbelt use alert quality. 101

Table 109 – Frequency counts for overall order ranking of seatbelt use alert quality for current Motive users. 101

Table 110 – Frequency counts for ranking of seatbelt use alert quality for current Motive users compared to non-users. 102

Table 111 – Frequency counts for overall order ranking of close following distance alert quality. 103

Table 112 - Model results for overall order ranking of close following distance alert quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values. 103

Table 113 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in close following distance alert quality. 103

Table 114 – Frequency counts for overall order ranking of close following distance alert quality for current Motive users. 104

Table 115 – Frequency counts for ranking of close following distance alert quality for current Motive users compared to non-users. 104

Table 116 – Frequency counts for Likert scale ratings of hardware design quality. 105

Table 117 - Model results for Likert scale ratings for hardware design quality assessed as continuous variables, with parameter estimate values. 106

Table 118 - Model results for Likert scale ratings of hardware design quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values. 106

Table 119 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in hardware design quality. 106

Table 120 – Frequency counts for Likert scale ratings of hardware design quality for current Motive users. 107

Table 121 – Frequency counts for Likert scale ratings of hardware design quality for current Motive users compared to non-users. 107

Table 122 – Frequency counts for overall order ranking of daytime image quality. 108

Table 123 - Model results for overall order ranking of hardware design quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values. 108

Table 124 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in hardware design quality. 109

Table 125 – Frequency counts for overall order ranking of hardware design quality for current Motive users. 109

Table 126 – Frequency counts for ranking of hardware design quality for current Motive users compared to non-users. 110

Table 127 – Comparison of time to alert and average successful alerts across all systems. 110

Introduction

Background

The Federal Motor Carrier Association’s Large Truck Crash Causation Study (2006) determined that drivers were cited as the main precipitating factor in 55% of crashes involving heavy vehicles. Eighty-seven percent of the critical factors included errors related to decision, recognition, non-performance, and performance. Similarly, the National Highway Transportation Safety Administration found that 94% of the primary contributing factors to crashes were due to the driver (Singh, 2015). Forty-one percent of these crashes were due to some kind of driver inattention and 33% involved decision making such as speeding, illegal maneuvers, gap misjudgment, etc. This research stresses the importance of reducing unsafe driver behavior and improving driver performance and decision making.

Driver monitoring systems (DMSs) aim to reduce driver-related crashes by holding the driver accountable for risky driving behavior through in-cabin alerts and coaching from fleet safety managers. Boyle et al. (2016) found that the combination of in-cab alerts and coaching reduced the likelihood of safety-related events in two fleets by 59.8% and 50.1%. Additionally, in a similar study by Hickman and Hanowski (2011), the implementation of DMSs and coaching reduced the likelihood of safety-related events by 37% and 52.2% in two fleets. These independent research results illustrate how DMSs can address critical reasons associated with CMV-involved crashes.

Introduction

Motive contracted the Virginia Tech Transportation Institute (VTTI) to benchmark their technology against two other commercially available DMSs. VTTI benchmarked each of the following three DMSs (Figure 1).



Figure 1 - The three DMSs used for benchmarking (Left to right: Motive, Samsara, and Lytx.)

The purpose of the benchmarking effort was twofold. Firstly, VTTI assessed the frequency that each system successfully alerted the driver to risky behaviors and maneuvers while installed in a

heavy vehicle. A successful alert was operationally defined as the correct alert was generated when the behavior was performed. The six driving events used to benchmark the technology were:

- 1) making an outgoing phone call,
- 2) texting mid-air on a cellphone (45-degree angle),
- 3) texting in the lap on a cellphone,
- 4) close following,
- 5) not wearing a seatbelt, and
- 6) rolling through a stop sign.

To ensure a fair testing environment, VTTI followed strict experimental design protocols developed based on previous research. Although the task types were selected by the sponsor, the experimental design was developed and approved by two subject matter experts and a statistician.

Additionally, the team assessed CMV driver opinions of the three DMSs on various specifications. Specifically, the drivers provided ratings about each of the DMS's video/image quality, field of view, alert quality, hardware quality, and their overall preference. The videos and images included in the survey illustrated what a driver would see when reviewing footage with a manager or looking back on recorded video, while the alerts and hardware mimic what a driver would experience inside the cab of a truck.

Methods

The three DMSs were evaluated in a two-phase study: technical performance benchmarking and user performance benchmarking. The first assessment gauged how frequently each system issued an alert to the six driving behaviors using a controlled test-track experiment. The second phase involved recruiting 188 commercial vehicle drivers to provide ratings in an online survey about various qualities of the systems. These studies did not assess each DMS's performance with different driver characteristics, maintenance requirements, data storage or access, nor the interface drivers might use to review captured footage. The detailed procedures used for each phase of the benchmarking efforts are specified below.

Phase One: Technical Performance Benchmarking

Phase one included a controlled test-track experiment to determine the success rate of each system at providing in-cab alerts in response to specific driving behaviors. Data was collected between April 14th, 2023, through May 2nd, 2023. The below methods were deployed to maintain consistency throughout all testing.

DMS Acquisition & Installation

Each DMS was purchased new directly from Lytx, Samsara, and Motive, respectively, and shipped to the campus at VTTI. The systems arrived with packaging intact and no visible damage. The DMSs were installed simultaneously in a Class 8 tractor provided by VTTI. All systems were installed at once to minimize variability in sun position, cloud cover and driver performance that may occur if each system were tested independently. The systems were installed by a certified, professional third-party installer to ensure that the placement of the cameras met each technology provider's installation standards. The installer determined that ideal camera placement for any system was as high and centered on the windshield as possible. Therefore, the ideal position for the systems was determined to be the upper middle of the windshield as illustrated in Figure 2.



Figure 2 - Example of ideal positioning for testing.

Two other positions were selected to test the dashcams simultaneously. Each of the three positions is illustrated in Figure 3. The two other placements are still within the manufacturer’s guidelines, but slightly to the right and left of the ideal position.



Figure 3 - Example of three positions used during testing.

To prevent position bias, the three DMSs were tested in each location on the windshield. Each system experienced one full rotation of testing in each location, and all cameras were rotated twice after the initial installation. This ensured that all DMSs were tested in all locations. In addition to the DMSs, the VTTI team installed a data acquisition system (DAS) to collect data on the cameras and truck during testing. Figure 4 shows what the positions looked like from the DAS’s perspective and show a clearer image of each of the three DMSs to illustrate order.



Figure 4 - The three rotations used during testing: (a) Rotation 1: Motive, Samsara, Lytx; (b) Rotation 2: Lytx, Motive Samsara; (c) Rotation 3: Samsara, Lytx, Motive.

Prior to testing during each rotation, VTTI conducted a verification test of each DMS to confirm that each technology met their intended performance specifications and that there were no obstructions blocking the cameras.

Experimental Design

The three DMSs were tested using six risky driving behaviors, also referred to as “tasks,” and three lighting conditions. Each of the six tasks was performed five times during daylight with dry conditions and no external roadway lighting, five times during the night with dry conditions and roadway lighting consistent with current highway (i.e., a 0.75-mile section of 4000K HPS luminaries spaced 80 meters apart) and street light (i.e., one 150 HPS cobra head mounted at 30 feet) standards, and three trials in civil twilight with dry conditions and no roadway lighting. The daylight, nighttime, and civil twilight time constraints were determined using <https://suncalc.org/>.

This process was repeated with each of the three DMS positions, resulting in a total of 234 data points per DMS across the length of the study (15 for day, 15 for night, and nine for twilight across each of the six tasks). To minimize variability in the results, the following experimental controls were used:

- All testing occurred during dry road conditions.
- The same driver was used throughout the duration of the study, and the driver did not alter their appearance in any significant way (for example, the driver wears glasses, so the driver wore glasses for every trial). The driver wore a shirt with a high contrast to the seatbelt for seatbelt testing (i.e., a white shirt with a dark gray seatbelt).
- The driver used the same seat location across all conditions and DMSs.
- The exact same tractor, a 2014 International Lonestar, was used across all conditions and DMSs.
- The driver followed the same repeatable procedure for each driving task.
- Twilight testing occurred during civil twilight to ensure repeatable lighting conditions.

All testing was performed on the highway and surface street sections of the VTTI Smart Road, which is a closed test bed research facility (Figure 5).

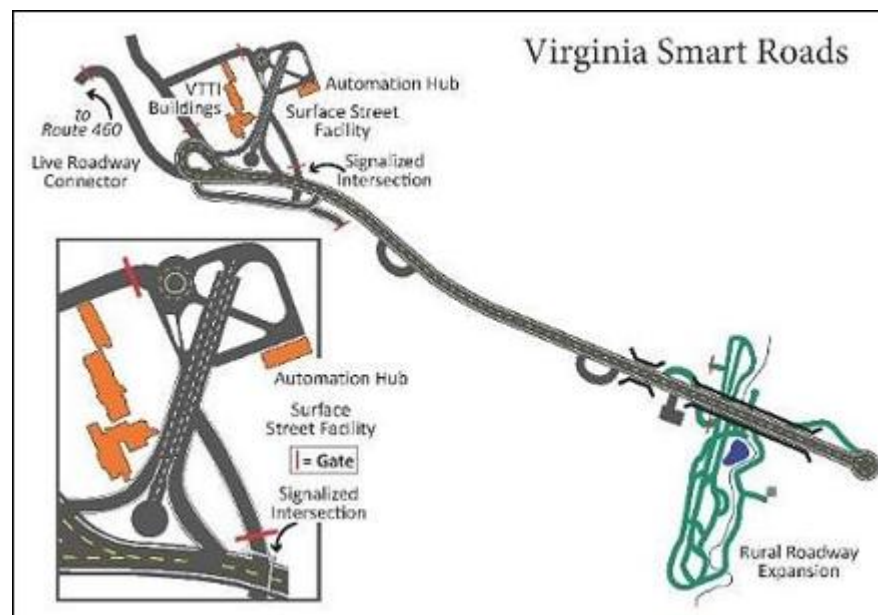


Figure 5 - Map of the Virginia Smart Roads.

On-road Protocol

Each task was standardized to ensure the test driver performed similarly across each trial. To improve consistency, the test driver practiced each behavior 15 times in the daytime, 15 times at nighttime, and nine times at twilight to ensure consistent performance across trials. To mimic realistic driving behavior for each of the six task types, the VTTI team used research-supported protocols as described in the following sections. All protocols were reviewed by two subject matter experts in the research field and one expert statistician.

Make an Outgoing Call on a Handheld Smartphone

Each of the DMSs has a form of cell phone calling alert. For Samsara¹, the alerts are called “mobile usage,” which “detects the phone itself” and “includes events when a driver holds a phone up to their ear without lowering their head.” Samsara also has an “inattentive driving” alert, which “detects abnormal changes in the driver’s head position.” Lytx² has “handheld device” alerts that “recognize objects and actions that indicate distracted driving behavior such as using a cell phone.” Lytx also mentions that their device can detect cell phone use “even if the object in question is hidden from view” because “the AI can detect behaviors such as looking down repeatedly to determine that a driver is texting.” The Motive³ DMS “detects behaviors like cell phone use” and “distraction.” These alerts are summarized in Table 1.

Table 1 – Handheld Cell Phone Alerts for each DMS

DMS	Alert Name(s)	Threshold Description
Motive DC-54	Cell Phone Usage/Distracted	The driver is using a phone for at least 5 s above 25 mph.
Samsara HW-CM3	Mobile Usage/Inattentive Driving	The driver is using a cell phone at speeds above 25 mph.
Lytx DriveCam SF400	Inattentive Driving/Handheld Device	The driver is holding/using a handheld cell phone that results in pulling driver visual attention away from roadway.

The following protocol was used to test each system’s alerts:

1. The driver looked down at the cup holder once.
2. The driver reached for a smartphone in the cup holder with their right hand.
3. The driver held the phone in their right hand at an elbow bend of 45 degrees within view of the camera.
4. Olson et al. (2009) found that drivers dialing their phone tend to look down for an average of 3.8 seconds over a 6-second period. Therefore, the driver looked up and down at the phone for 2 seconds with eyes off road and 1 second with eyes on road twice for a total of 6 seconds.
5. After 6 seconds, the driver held the phone up to their right ear for 30 seconds while looking at the road.

Figure 6 portrays what the outgoing phone call looked like when performed by the test driver.

¹ Samsara - <https://kb.samsara.com/hc/en-us/articles/360057221352-Distracted-Driving-Triggers>

² Lytx - <https://resources.lytx.com/blog/machine-vision-artificial-intelligence-triggers>

³ Motive - <https://gomotive.com/content-library/spec-sheet/ai-dashcam/>



Figure 6 - Example of driver during on-road testing for outgoing phone call.

Send an Outgoing Text Message

The alerts described in Table 1 also apply to the 45-degree texting task.

The following protocol was used to test each system's alerts:

1. The driver looked down at the cup holder once.
2. The driver reached for a smartphone in the cup holder with their right hand.
3. The driver held the phone in their right hand at an elbow bend of 45 degrees.
4. Olson et al. (2021) found that texting drivers tend to look down at their phone for an average of 5 seconds over a 6-second period. Therefore, the driver looked up and down at the phone for 2.5 seconds with eyes off road, 1 second with eyes on road, then 2.5 seconds with eyes off road, four times, for a total of 12 seconds.

Figure 7 portrays what the 45-degree text message behavior looked like when performed by the test driver.



Figure 7 - Example of driver during on-road testing for texting condition (45-degrees).

Discretely Use a Smartphone in Lap

The alerts described in Table 1 also apply to the lap texting task.

The following protocol was used to test each system's alerts:

1. The driver looked down at the cup holder once.
2. The driver reached for a smartphone in the cup holder with their right hand.
3. The driver held their phone in their hand while the back of their hand touched their lap.
4. Olson et al. (2021) found that drivers who are texting tend to look down at their phone for an average of 5 seconds over a 6-second period. Therefore, the driver looked up and down at the phone for 2.5 seconds with eyes off road, 1 second with eyes on road, then 2.5 seconds with eyes off road, four times, for a total of 12 seconds.

Figure 8 portrays what the lap texting behavior looked like when performed by the test driver.

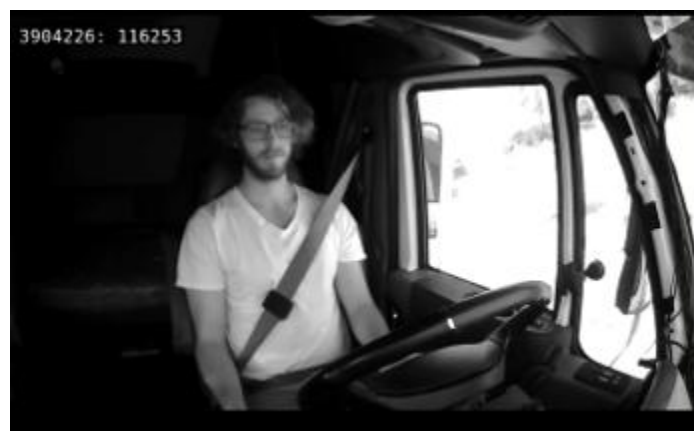


Figure 8 - Example of driver during on-road testing for lap texting condition.

Closely Following a Lead Vehicle.

Each of the DMS systems has a form of close following distance alert. For Samsara⁴, the alerts are called “following distance,” which “detects unsafe following distance, often referred to as tailgating.” Lytx⁵ has “Following Distance” alerts that use thresholds for the lead vehicle to determine when an unsafe following distance is maintained. The Motive⁶ DMS has “close following” alerts that detect when a lead vehicle is too close to the front of the truck. These claims are summarized in Table 2.

Table 2 - Close Following Alerts for each DMS

DMS	Alert Name(s)	Threshold Description
Motive DC-54	Close Following	Driver is travelling above 35 mph for 10 s at a minimum of 1.0 s following distance.
Samsara HW-CM3	Following Distance	Driver is travelling above 35 mph for 15 s at a minimum of 1.0 s following distance.
Lytx DriveCam SF400	Following Distance	Driver is travelling above 25 mph for 4 s at a minimum of 1.5 s following distance.

The following protocol was used to test each system’s alerts:

1. The truck followed the lead vehicle, a Ford Focus, accelerating to 50 mph from a stop.
2. Sarkar et al. (2022) found that truck drivers give an average of 2 seconds of headway with 0.62 seconds standard deviation. It was assumed that 1.5 standard deviations represent a risky driving population based on this data. This equals a 1.07-second following distance. Therefore, the test driver created an approximately 1-second following distance (representing the riskiest subset of drivers) and maintained that headway for 30 seconds.

Figure 9 portrays what the close following behavior looked like when performed by the test driver.

⁴ Samsara - <https://kb.samsara.com/hc/en-us/articles/360043193612-Following-Distance-Detection>

⁵ Lytx - <https://resources.lytx.com/blog/machine-vision-artificial-intelligence-triggers>

⁶ Motive - https://gomotive.com/wp-content/uploads/2023/05/ai_dashcam_spec_sheet.pdf



Figure 9 - Example of lead vehicle during on-road testing for close following.

Driver Not Wearing Seatbelt

Each of the DMS systems has a form of close following distance alert. However, VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled during the data collection period. For Samsara⁷, the alert is called “seatbelt unbuckled,” which “detects when the driver is not wearing a seatbelt at the beginning of a trip.” Lytx⁸ has a “driver unbelted” alert that detects if “a driver is not wearing a seatbelt or if it is worn improperly.” The Motive⁹ DMS has a “close following” alert that detects when a lead vehicle is too close to the front of the truck. These alerts are summarized in Table 3.

Table 3 - No Seatbelt Alerts for each DMS

DMS	Alert Name(s)	Threshold Description
Motive DC-54	Seatbelt Violation	Seatbelt is unbuckled for at least 15 s above 10 mph.
Samsara HW-CM3	Seatbelt Unbuckled	Driver does not have a seatbelt on at the beginning of the trip.
Lytx DriveCam SF400	Driver Unbelted	Driver is not wearing a seatbelt or is wearing it improperly at the beginning of a trip.

The following protocol was used to test each system’s alerts:

1. The driver removed his seatbelt when the truck was stopped.
2. The driver accelerated to above 45 mph and drove without a seatbelt for at least 30 seconds.

⁷ Samsara - <https://kb.samsara.com/hc/en-us/articles/360043619011-AI-Event-Detection>

⁸ Lytx - https://beldongroupofcompanies.files.wordpress.com/2020/06/guide_to_lytx_event_analysis.pdf

⁹ Motive - https://gomotive.com/wp-content/uploads/2023/05/ai_dashcam_spec_sheet.pdf

- To begin the next trial, the truck was stopped, and the driver put the seatbelt back on.

Figure 10 portrays what the no seatbelt behavior looked like when performed by the test driver.



Figure 10 - Example of driver during on-road testing for no seatbelt condition.

Rolling Stop Through a Stop Sign

Each of the DMSs has a form of rolling stop detection. However, it is important to note that Lytx reported a known issue with their rolling stop detection. The company disabled rolling stop alerts due to a high number of false alerts. Therefore, this system was not included in the analysis for rolling stops. Samsara¹⁰ uses an alert called “rolling stop detection” that “determines when a driver does not stop completely for a posted stop sign.” Motive¹¹ uses an alert called “stop sign violation” that detects when a driver has rolled through a stop sign without stopping. These alerts are summarized in Table 4.

Table 4 - Rolling Stop Alerts for each DMS

DMS	Alert Name(s)	Threshold Description
Motive DC-54	Stop Sign Violation	Driver rolls through a stop sign and does not go below 6 mph within 7 s of the stop sign being invisible to the camera.
Samsara HW-CM3	Rolling Stop	Driver rolls through a stop sign at a minimum of 10 mph.
Lytx DriveCam SF400	N/A	N/A

¹⁰ Samsara - <https://kb.samsara.com/hc/en-us/articles/360029629972-Rolling-Stop-Detection>

¹¹ Motive - https://gomotive.com/wp-content/uploads/2023/05/ai_dashcam_spec_sheet.pdf

The following protocol was used to test each system's alerts:

1. The driver travelled straight on a roadway at 25 mph in the right lane.
2. The driver slowed to 15 mph, a minimum of 20 feet before the stop sign.
3. The driver rolled through the intersection at 15 mph.

Figure 11 portrays what the rolling stop behavior looked like when performed by the test driver.



Figure 11 - Example of forward view during on-road testing for the rolling stop condition.

Analysis

DMS alert data was collected through a trimodal approach. First, an experimenter in the passenger seat recorded in-cabin alerts in real time during experimentation. Second, an in-house VTTI DAS recorded footage of visual in-cabin alerts and truck kinematics for post-experimental review. Third, the DMSs recorded dashboard alerts to their respective vendor's platforms and these dashboards were used to collect audio data for all auditory in-cabin alerts. The illumination in lux during each trial was also collected.

In-cab Alerts

The main goal of the analysis was to determine which DMSs successfully detected each of the risky driving behaviors performed by the test driver. During on-road testing, an experimenter in the passenger seat manually recorded all in-cabin alerts via visual and auditory detection. If the system provided an alert during the duration of a driving task, this was counted as a "success" and the experimenter indicated the success on a printed sheet (Figure 12).

	Rotation 1 - Motive, Samsara, Lytx																	
	Rolling Stop			Phone Call			45-Degree Texting			Lap Texting			Close Following			No Seatbelt		
	Motive	Samsara	Lytx	Motive	Samsara	Lytx	Motive	Samsara	Lytx	Motive	Samsara	Lytx	Motive	Samsara	Lytx	Motive	Samsara	Lytx
Day Trial 1																		
Day Trial 2																		
Day Trial 3																		
Day Trial 4																		
Day Trial 5																		
Twilight Trial 1																		
Twilight Trial 2																		
Twilight Trial 3																		
Night Trial 1																		
Night Trial 2																		
Night Trial 3																		
Night Trial 4																		
Night Trial 5																		

Figure 12 – Example of alert chart used by experimenter during testing.

False alarms, or alerts that occurred without a warranted behavior, were not automatically counted as a “failure” (unless a true alert did not occur). However, the total number of false alarms was noted. In the event there were multiple successful alerts sent for the same behavior, only the first alert was included in the analysis. In other words, there were no benefits or penalties for a device to send multiple alerts. For example, during the no seatbelt condition, one of the systems would produce an alert multiple times since the driver did not correct the behavior until the end of the trial.

The success of in-cab alerts naturally follows a binary response (i.e., yes/no, 1/0). Therefore, a logistic regression model was built for each task type where the response variable equals the binary variable representing in-cab alert status for the task. The response variable is shown below:

$$Y = \begin{cases} 1 & \text{success, in - cab alert issued} \\ 0 & \text{failure, in - cab alert not issued} \end{cases}$$

The probability of a successful in-cab alert is measured as p . The variables used to explain the success or failure of an alert were system type (Lytx, Motive, Samsara), system rotation (1, 2, 3), and time of day (day, night, twilight). The model structure is shown below:

$$\text{logit or } \ln\left(\frac{p}{1-p}\right) = X'\beta + \varepsilon$$

The model intercept represents one level of each explanatory variable (i.e., system, rotation, and time of day). The model coefficients for each explanatory level explain how the level compares to the intercept in predicting successful in-cab alerts for the task. Relative risk ratios were calculated for each task type to compare the likelihood of receiving an alert for each system.

Dashboard Alerts Analysis

With DMS in-cab alerts, discussed previously, drivers hear an alert in real time when they perform risky behaviors. In some instances, these alerts are documented to a dashboard so the events can be reviewed later by the driver or a safety manager. The second analysis reviewed the number of alerts sent to each DMS's dashboard, respectively. The dashboard refers to the system's virtual platform. It is important to note that each company has a different threshold for sending alerts to the dashboard for review. For example, one company may only document alerts after the second time the driver performs the behavior, whereas another may document each time an in-cab alert is triggered. The dashboard alerts were not used to assess the accuracy of the DMSs and were instead used to identify if there were alerts that appeared on a dashboard but that were not given to a driver in-cab. Dashboard alerts were counted if the alert happened within the time frame of testing and accurately identified the risky behavior performed.

For each task performed, researchers noted whether the system issued an alert in-cab or on the dashboard. These results were summarized by in-cab or dashboard alert, in-cab alert only, and no in-cab or dashboard alert. These were then summarized using counts in tabular and graphical form.

Time to Alert Analysis

In addition to successful dashboard and in-cab alerts, the video collected through the DAS was reviewed to determine the time between the task start and the successful DMS alert given. The "time to alert" metric was calculated for each of the DMSs and for each task type. Due to individual technology differences in thresholds for each of the DMS providers, "time to alert" was not included in the performance score but was calculated as an independent metric.

Time to alert is a continuous variable measured as the time in seconds between task begin and in-cab alert begin time. Therefore, a typical analysis used to test a continuous dependent variable for significant differences between groups is the analysis of variance (ANOVA) test. However, after reviewing the in-cab alert success analysis, it was clear that the in-cab alert performance varied widely across system types and between task types. To account for the unbalanced data and the small sample sizes, a non-parametric approach was used to assess data for significant differences in time to alert. For each task type a Kruskal-Wallis test was used to test for significant differences overall. To perform a Kruskal-Wallis test on a single task, all data from all systems is grouped and then ranked in order and testing against the following hypotheses:

The null hypothesis H_o = all groups have the same distribution with the same median

The alternative H_a = at least one group follows a different distribution and median

If the test was significant, post-hoc Dunn tests with Bonferroni corrections were used to compare two systems at a time for significant differences.

Phase Two: User Performance Benchmarking

To determine how users viewed each DMS's video quality, image quality, alert quality, field of view, and hardware aesthetics, an online questionnaire was developed. Survey participants included 188 commercial vehicle drivers who drive trucks on a regular basis or provide safety

coaching to commercial vehicle drivers. Participants answered at least one question about the content via an electronic survey platform called QuestionPro. All questions were formatted to be viewed appropriately on a mobile device or desktop.

Survey Content

The three DMSs were de-identified to avoid bias. Code names such as “Oak,” “Pine,” and “Maple” were used in place of the DMS names. Firstly, participants indicated what level of experience they had with DMSs. Then, participants rated each quality of the three DMSs on a Likert scale from 1 to 5 where one indicates they strongly *disagree* with the statement and five indicates they strongly *agree*. Participants also made an overall rating of preference by ordering the DMSs 1, 2, and 3 for each quality. For the video and image questions, participants rated both nighttime media and daytime media. All images and videos were the same size, approximately the same lighting, and similar task types to avoid confounds. Additionally, the presentation of the questions within each category was randomized to avoid anchoring bias.

Video Quality Questions

The video questions included in the survey were downloaded from each DMS company’s dashboard. The interior and exterior views from the dashboards were stitched together and the video quality was not altered in any way. The videos did not include audio. The videos were rated based on whether the participant perceived the video to have “good video quality,” which comprised “clear visuals and details, appropriate view of the cabin and front of truck, and continuous video flow.” Figure 13 gives an example of the individual and overall video questions.

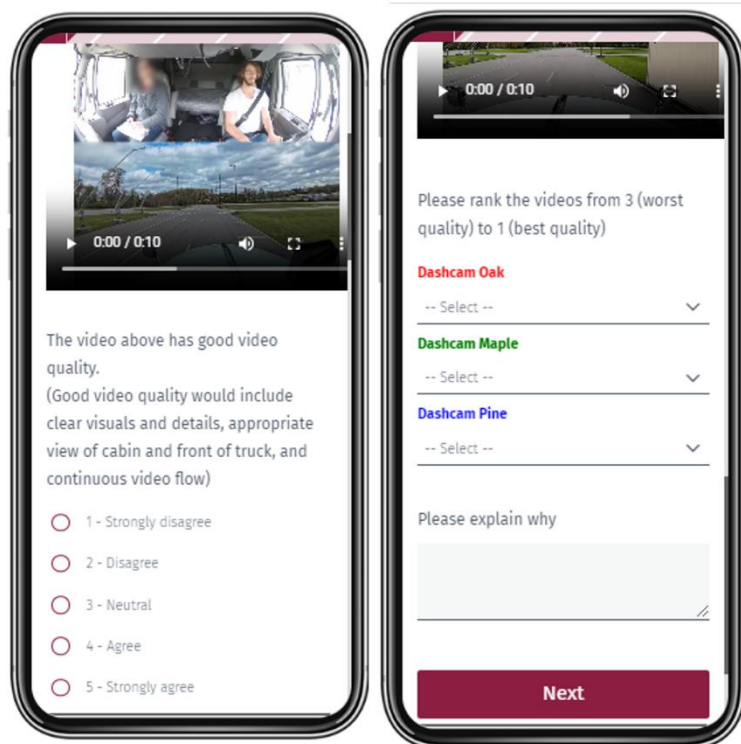


Figure 13 - Example of video questions viewed from a mobile phone.

Image Quality Questions

The image questions included in the survey were screen captures of downloaded video from each company’s dashboard. The images were rated based on whether the participant perceived the image to have “good image quality,” which comprised “clear details and appropriate view of the cabin and front of the truck.” Figure 14 gives an example of the individual and overall image questions.

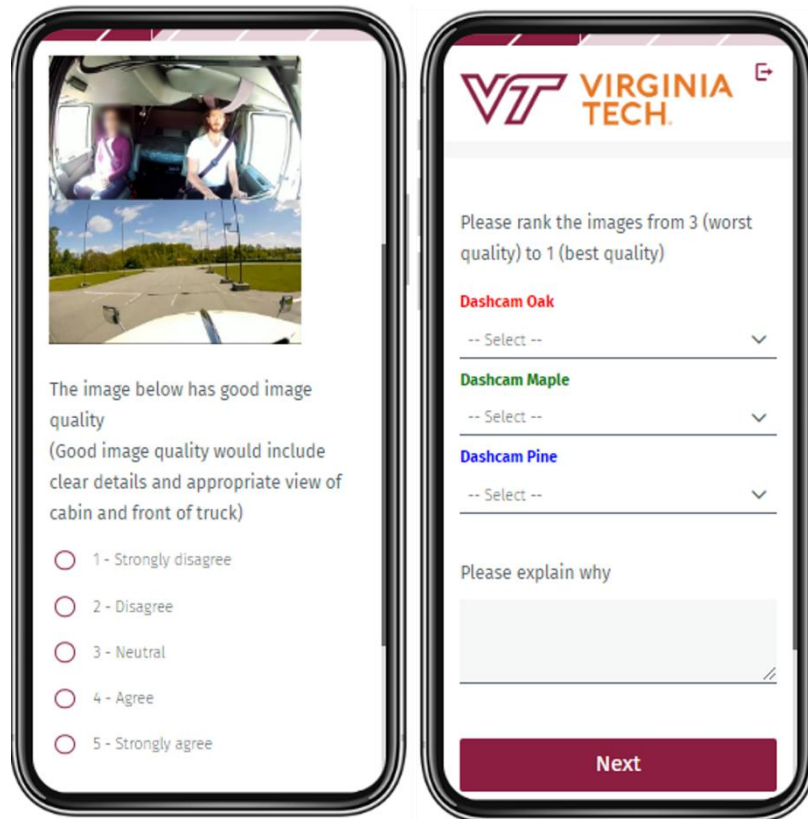


Figure 14 - Example of image questions viewed from a mobile phone.

Audio Alert Questions

The audio alert questions included in the survey were downloaded audio from each of the DMS company’s websites. These audio clips were not recorded from inside the truck to avoid bias in perceived loudness. Participants rated the audio clips in terms of “good quality,” which comprised “clear audio, clear communication of alert intent, and pleasing audio.” The participants rated alerts for distracted driving, seat belt violations, and close following. Figure 15 gives a visual example of the individual and overall audio alert questions.

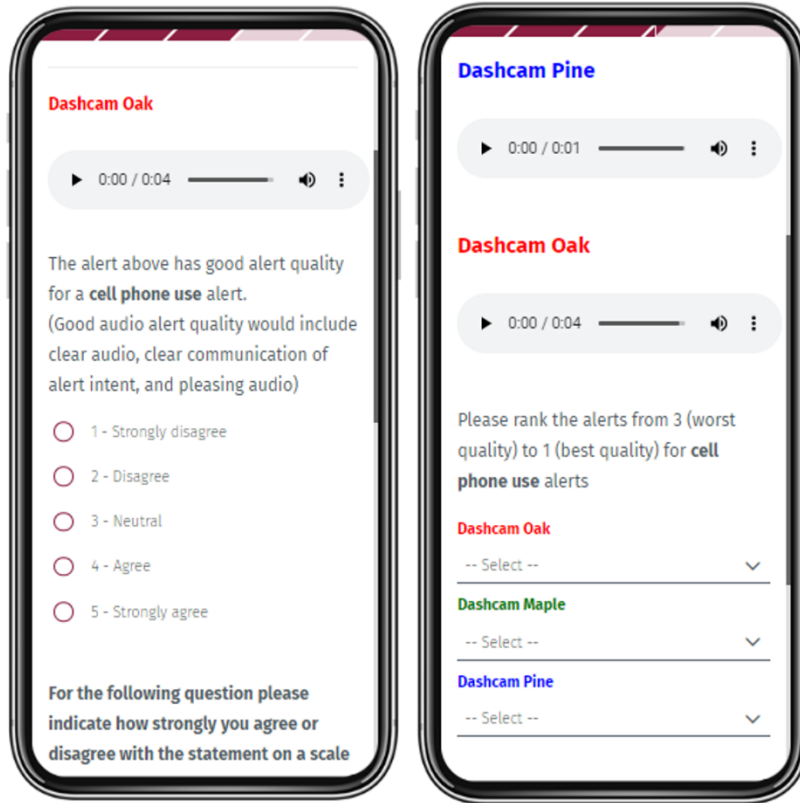


Figure 15 - Example of audio alert questions viewed from a mobile device.

Hardware Quality Questions

The hardware questions included in the survey were photographed by a professional photographer with the same lighting and angle. The overall hardware questions were photographed on a plain table with a quarter as a reference for size. Participants rated the individual hardware based on how well the “shape, color, form, etc.” of the hardware is “fitting for its use, does not block line of sight out of the windshield, and is aesthetically pleasing.” Participants rated the overall hardware comparison based on “size, shape, and aesthetics.” Figure 16 gives an example of the individual and overall hardware questions.

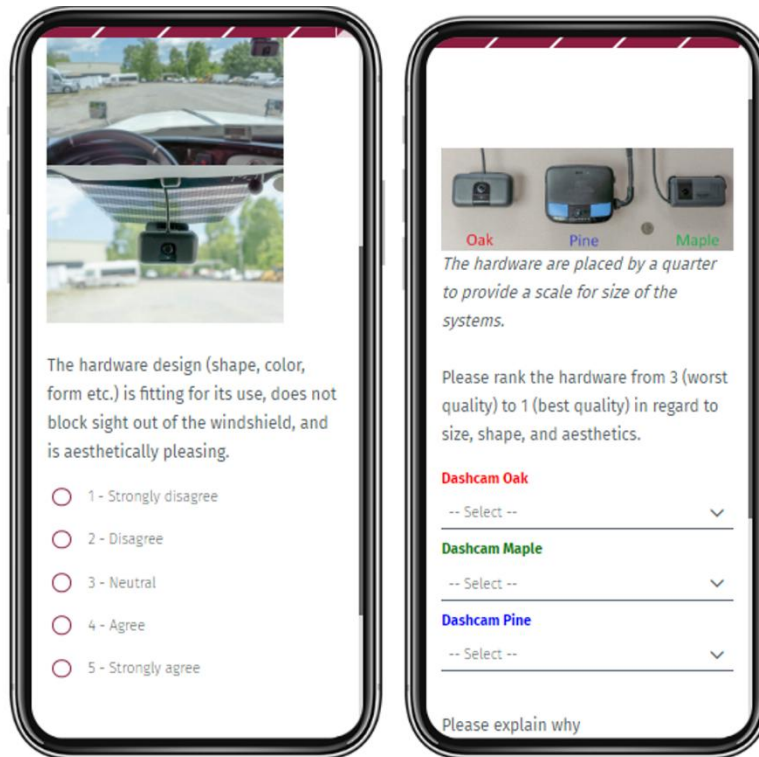


Figure 16 - Example of hardware questions viewed from a mobile device.

Analysis

A quantitative data analysis was conducted on the survey responses using the statistical software SAS version 9.4. Survey responses were included from 188 commercial vehicle drivers who completed at least one question in the survey. Rating distributions for the four system characteristics were developed. The system characteristics include video quality, image quality, alert quality, and hardware design. These distributions included tables with summary statistics and plots of the full range of rating responses. The survey consisted of three main question types: Likert scale ratings, ordering, and free response questions. For the Likert scale ratings and overall order questions, descriptive statistics were used to assess the frequency of positive ratings across all DMS qualities. In addition, distributions of Likert scale ratings and overall order questions were calculated for current Motive users, to understand current users' assessments of all three systems. Current Motive users were compared to non-users for their distribution of Likert scale ratings and overall order questions.

Ratings

The Likert scale ratings were assessed in two ways. In the first approach, the ratings were treated as continuous values and assessed using a generalized linear model built per survey question. The model predicted survey rating by two explanatory variables: the system evaluated (Lytix, Motive, Samsara) and the Motive user status of the respondent (current Motive user or non-Motive user). The model structure is as follows:

$$Y_{System\ Facet} = \text{continuous value ranging between 1 and 5}$$

$$y_{System\ Facet} = X'\beta + \varepsilon$$

where y is the rating score for one system facet, β is the model coefficient vector for system evaluated and Motive user status, and ε is the error term. The model intercept represents one level of each explanatory variable (i.e., system and Motive user status). The model coefficients for each explanatory level explain how the level compares to the intercept in predicting rating value for the system characteristic. A model was built for each system characteristic. For video quality and image quality, an individual model was built to evaluate ratings of day videos and images and night videos and images. For alert quality, an individual model was built for each alert type assessed (cell phone use, seatbelt use, and close following distance).

In the second approach, the ratings were assessed as a binary response variable using a logistic regression model. The model predicted the likelihood of a system characteristic to be rated 4 or higher. The response variable is shown below:

$$Y = \begin{cases} 1 & \text{positive response, system facet rated 4 +} \\ 0 & \text{negative or neutral response, system facet rated 3 or below} \end{cases}$$

The probability of a positive response is measured as p . The variables used to explain a positive response were the system evaluated (Lytx, Motive, Samsara) and the Motive user status of the respondent (current Motive user or non-Motive user). The model structure is shown below:

$$\text{logit or } \ln\left(\frac{p}{1-p}\right) = X'\beta + \varepsilon$$

The model intercept represents one level of each explanatory variable (i.e., system evaluated and Motive user status). The model coefficients for each explanatory level explain how the level compares to the intercept in predicting a positive response for the system characteristic. Odds ratios were calculated for each system characteristic to compare the odds of a positive response by the system evaluated. As in the previous analysis, a model was built for each system characteristic, time of day, and alert type included in the survey.

Overall Order

The overall order data was analyzed to understand the likelihood of being ranked first. The overall order data was transformed to a binary variable, with rankings of first and rankings of second and third combined. The ranking data was then analyzed using a logistic regression model. The model predicted the likelihood of a system characteristic to be ranked first. The response variable is shown below:

$$Y = \begin{cases} 1 & \text{system ranked first} \\ 0 & \text{system ranked second or third} \end{cases}$$

The probability of ranking first is measured as p . The variables used to explain a ranking of first were the system evaluated (Lytx, Motive, Samsara) and Motive user status of the respondent (current Motive user or non-Motive user). The model structure is shown below:

$$\text{logit or } \ln\left(\frac{p}{1-p}\right) = X'\beta + \varepsilon$$

The model intercept represents one level of each explanatory variable (i.e., system evaluated and Motive user status). The model coefficients for each explanatory level explains how the level compares to the intercept in predicting a ranking of first for the system characteristic. Odds ratios were calculated for each system characteristic to compare the odds of a first ranking by system evaluated. As in the previous analyses, models were built for each system characteristic, time of day, and alert type included in the survey.

Content Analysis

For open-ended questions, a content analysis was performed to identify recurring themes and patterns. The written responses for each open-ended question were compiled into a spreadsheet. Each set of answers for all corresponding questions was read over once then upon a second pass each answer was bucketed into reoccurring themes for each question (Ex. For audio alerts, volume and tone were two of the primary reoccurring themes mentioned in free response). Day and night free response answers for video and image quality were combined to yield four content analysis groups: (1) video quality, (2) image quality, (3) alert quality, and (4) hardware quality. Mentions of each theme from all questions per analysis group were counted to yield four tables that visualize driver areas of focus regarding quality for the four analysis groups. The reoccurring themes expressed in the results are a compilation of the explanations as to why each participant gave the ranking they did. Because of the randomization of the questions, for a vast majority of the free response answers, there is no clear means to discern which system is being discussed in the free response. Therefore, the themes are high level counts determined by key words in the responses.

Results

Phase One: Technical Performance Benchmarking

The following section presents the results obtained from the analyses conducted in this study, which aimed to investigate the prevalence of successful alerts when compared across three DMSs and the opinions of users on different qualities of each system.

In-cab Alerts: Outgoing Phone Call

The percentage of successful alerts during the outgoing phone call trials was calculated for each of the three systems across the three lighting conditions. The percentages are included in Table 5.

Table 5 - Successful Phone Call Alerts by System Type and Time of Day

Time of Day	Total Runs	Lytix	Motive	Samsara
Day	15	73%	100%	53%
Night	15	0%	100%	33%
Twilight	9	0%	78%	22%
Overall	39	28%	95%	38%

For the 39 phone call trials, the Motive system produced a successful alert 95% of the time, while Samsara and Lytx produced a successful in-cab alert 38% and 28% of the time, respectively. Figure 17 displays the successful alert percentages for the lighting conditions and illustrates the difference in alerts for each rotation. The logistic regression model for the phone call task assessed whether the different explanatory study variables had a statistically significant effect on predicting a successful alert. System ($\chi^2(2) = 25.39, p < 0.0001$) and time of day ($\chi^2(2) = 25.39, p < 0.0001$) were significant explanatory variables. Specifically, Lytx and Samsara were significantly less likely to produce a successful alert for the phone call task than the Motive system across all lighting types and rotations ($p < 0.0001$ for both the comparison of Lytx to Motive and Samsara to Motive). Additional model results are included in the Appendix.

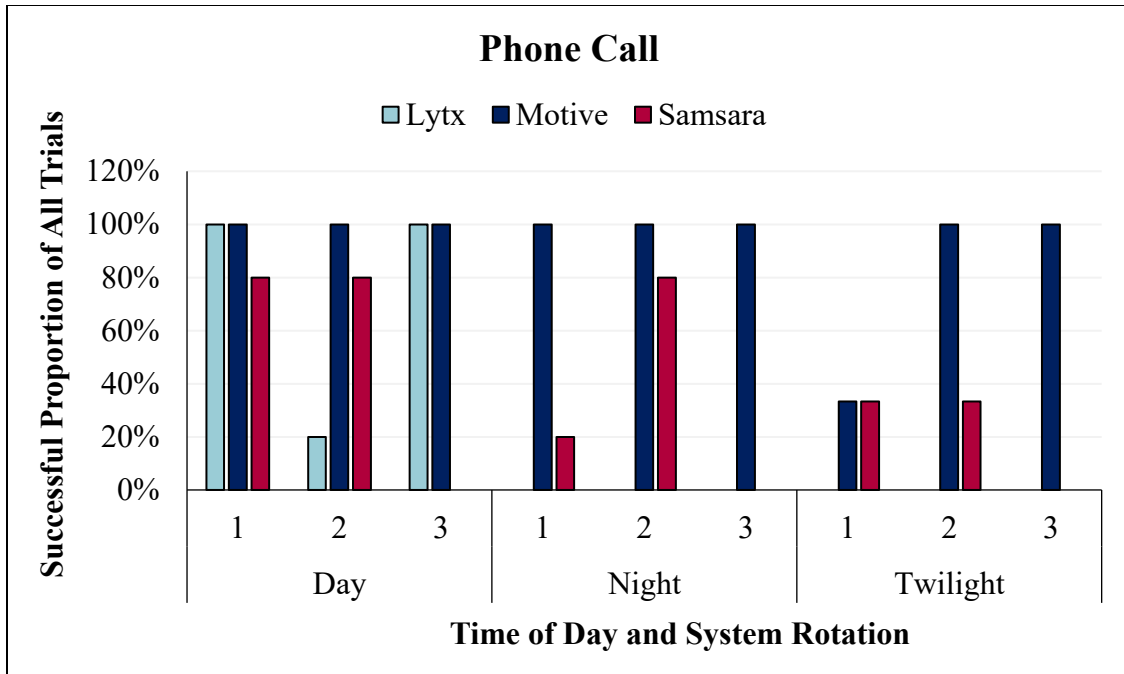


Figure 17 - Successful phone call alert percentages during on-road testing across system, time of day, and rotation.

Additionally, the relative risk of in-cab alerts for the cell phone task was calculated to determine the likelihood of a successful alert. The Motive system was 3.36 times more likely to produce an alert for the phone call task than Lytx (95% CI = 2.01, 5.58) and 2.47 times more likely than Samsara (95% CI = 1.65, 3.69).

In-cab Alert: Lap Texting

The percentage of successful alerts during the lap texting task was calculated for each of the three systems across the three lighting conditions. The percentages are included in Table 6.

Table 6 - Successful Lap Texting Alerts by System Type and Time of Day

Time of Day	Total Runs	Lytx	Motive	Samsara
Day	15	20%	67%	40%
Night	15	0%	40%	0%
Twilight	9	0%	50%	0%
Overall	39	8%	53%	15%

For the 39 lap texting trials, the Motive system produced a successful alert 53% of the time while Samsara and Lytx produced a successful in-cab alert 15% and 8% of the time, respectively. Figure 18 displays successful alert percentages for the lighting conditions and illustrates the difference in alerts for each rotation during lap texting. The logistic regression model for the lap texting assessed whether the different explanatory study variables had a statistically significant effect on predicting a successful alert. System ($\chi^2(2) = 21.34, p < 0.0001$), rotation ($\chi^2(2) = 8.91, p = 0.0116$), and time of day ($\chi^2(2) = 13.10, p = 0.0014$) were significant explanatory variables. Specifically, Lytx and Samsara were significantly less likely to produce a successful alert for the

lap texting task than the Motive system across all lighting types and rotations ($p < 0.0001$ for the comparison of Lytx to Motive, and $p = 0.0003$ for the comparison of Samsara to Motive). Additional model results are included in the appendix.

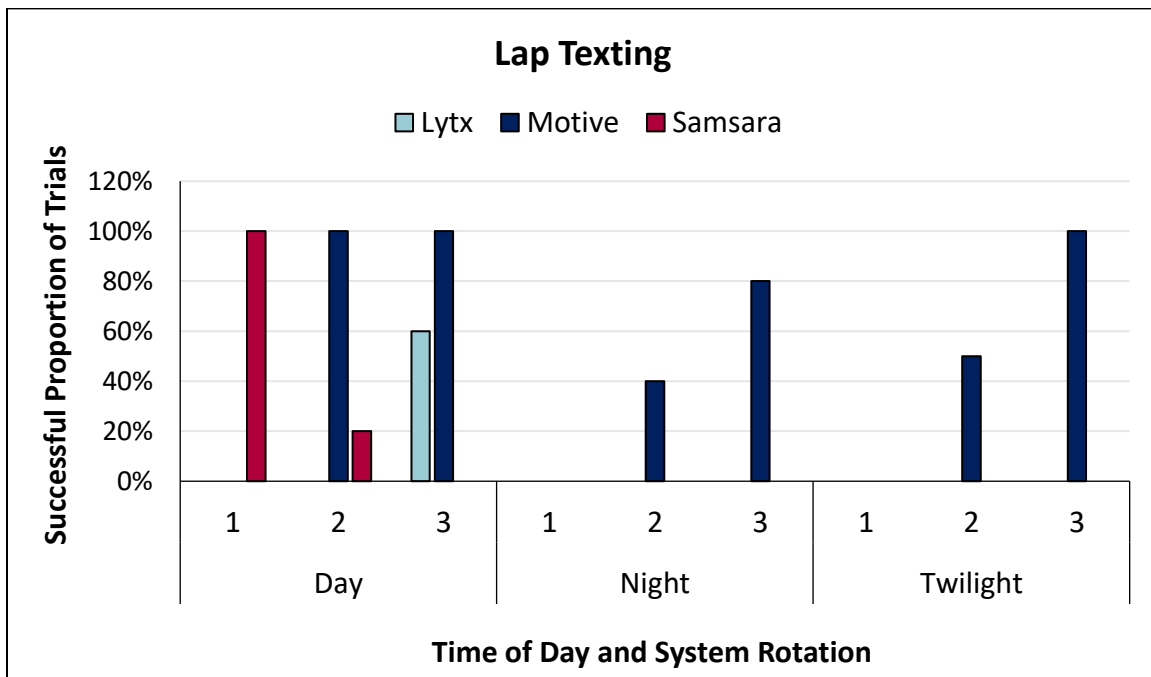


Figure 18 - Successful lap texting alert percentages during on-road testing across system, time of day, and rotation.

Additionally, the relative risk of in-cab alerts for the lap texting task was calculated to understand how likely an alert would be if the Motive system was being used as opposed to another system type. The Motive system was 7 times more likely to produce an alert for the lap texting task than Lytx (95% CI = 2.27, 21.62) and 3.50 times more likely than Samsara (95% CI = 1.58, 7.75).

In-cab Alert: 45-Degree Texting

The percentage of successful alerts during the 45-degree texting task was calculated for each of the three systems across the three lighting conditions. The percentages are included in Table 7.

Table 7 - Successful 45-Degree Texting Alerts by System Type and Time of Day

Time of Day	Total Runs	Lytx	Motive	Samsara
Day	15	47%	100%	53%
Night	15	0%	93%	53%
Twilight	9	0%	75%	25%
Overall	39	18%	92%	47%

For the 39 45-degree texting trials, the Motive system produced a successful alert 92% of the time, while Samsara and Lytx produced a successful in-cab alert 47% and 18% of the time, respectively. Figure 19 displays successful alert percentages for the lighting conditions and

illustrates the difference in alerts for each rotation during 45-degree texting. The logistic regression model for the 45-degree texting task found system ($\chi^2(2) = 28.26, p < 0.0001$), rotation ($\chi^2(2) = 7.57, p = 0.0227$), and time of day ($\chi^2(2) = 10.71, p = 0.0047$) statistically significant in predicting a successful alert. Both Lytx and Samsara were significantly less likely to produce a successful alert for the 45-degree texting task than the Motive system across all lighting types and rotations ($p < 0.0001$ for both the comparison of Lytx to Motive and Samsara to Motive). Additional model results are included in the Appendix.

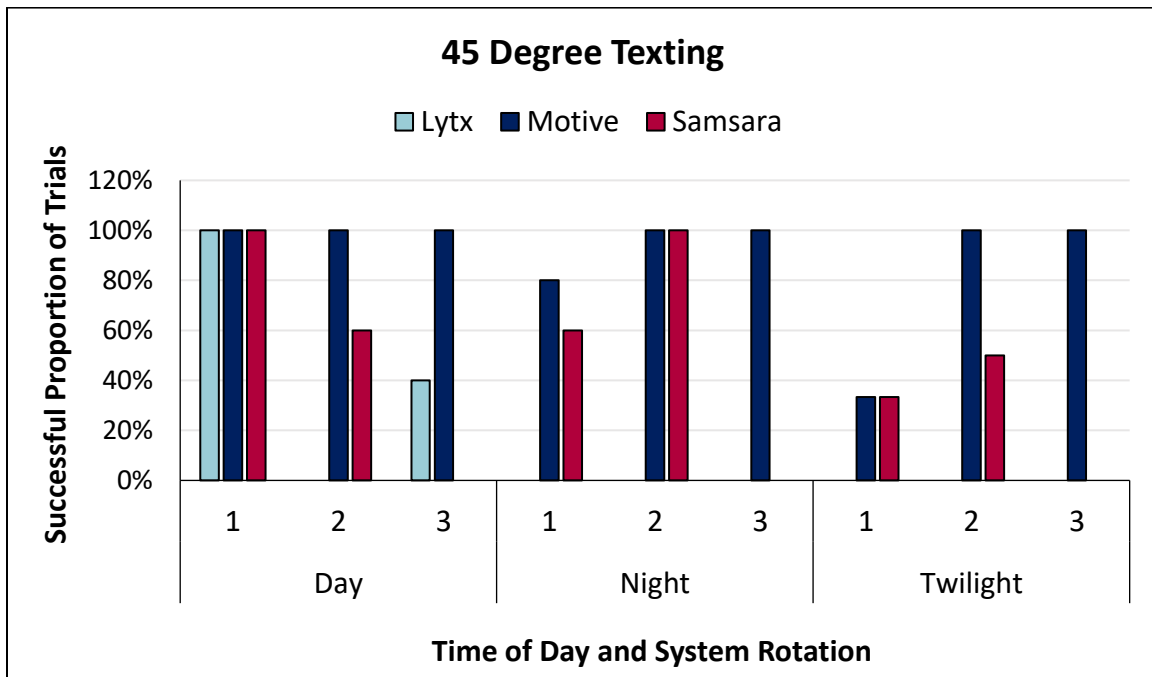


Figure 19 - Successful 45-degree texting alert percentages during on-road testing across system, time of day, and rotation.

Additionally, the relative risk of in-cab alerts for the 45-degree texting task was calculated to determine the likelihood of a successful alert. The Motive system was 5.14 times more likely to produce an alert for the 45-degree texting task than Lytx (95% CI = 2.61, 10.12) and two times more likely than Samsara (95% CI = 1.41, 2.84).

In-cab Alert: Close Following

The percentage of successful alerts during the close following task were calculated for each of the three systems across the three lighting conditions. The percentages are included in Table 8.

Table 8 - Successful Close Following Alerts by System Type and Time of Day

Time of Day	Total Runs	Lytx	Motive	Samsara
Day	15	33%	67%	20%
Night	15	20%	67%	13%
Twilight	9	33%	67%	22%
Overall	39	28%	67%	18%

For the 39 close following trials, the Motive system produced a successful alert 67% of the time, while Samsara and Lytx produced a successful in-cab alert 18% and 28% of the time, respectively. Figure 20 displays successful alert percentages for the lighting conditions and illustrates the difference in alerts for each rotation during close following. The logistic regression model for the close following task found system [$\chi^2(2) = 20.08, p < 0.0001$] to be statistically significant in predicting a successful alert. Both Lytx and Samsara were significantly less likely to produce a successful alert for the close following task than the Motive system across all lighting types and rotations ($p < 0.0008$ and $p < 0.0001$, respectively). Additional model results are included in the Appendix.

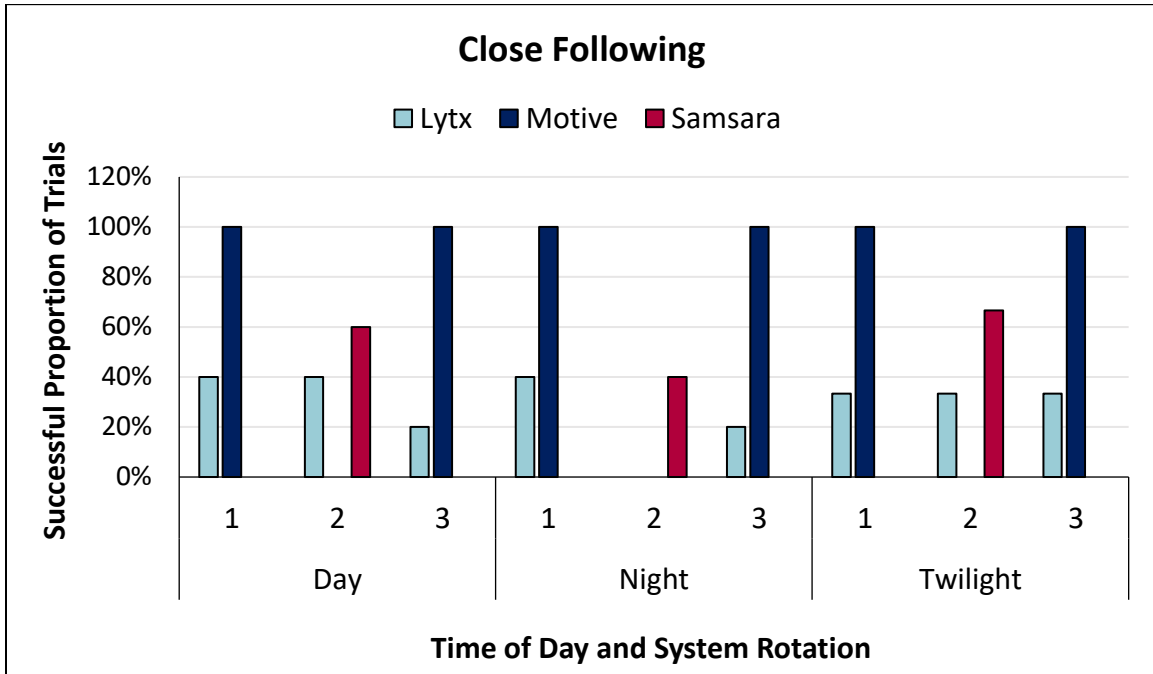


Figure 20 - Successful close following alert percentages during on-road testing across system, time of day, and rotation.

Additionally, the relative risk of in-cab alerts for the close following task was calculated to understand the likelihood of a successful alert. The Motive system was 2.89 times more likely to produce an alert for the close following task than Lytx (95% CI = 1.56, 5.34) and 3.71 times more likely than Samsara (95% CI = 1.83, 7.53).

In-cab Alert: Rolling Stop

The percentage of successful alerts during the rolling stop task was calculated for each of the three systems across the three lighting conditions. The percentages are included in Table 9.

Table 9 - Successful Rolling Stop Alerts by System Type and Time of Day

Time of Day	Total Runs	Lytx*	Motive	Samsara
Day	15	.	87%	0%
Night	15	.	60%	0%
Twilight	9	.	89%	0%
Overall	39	.	77%	0%

*Lytx system not evaluated for rolling stop alerts. The team was informed that the alert was deactivated by the system for all customers for the duration of the study.

For the 39 rolling stop trials, the Motive system produced a successful alert 77% of the time, while Samsara did not produce a successful in-cab alert. Figure 21 displays successful alert percentages for the lighting conditions and illustrates the difference in alerts for each rotation during the rolling stop trials. Due to the lack of successful alert data between all three systems for rolling stops, a chi-square test was run to compare only Motive and Samsara over all conditions. For the rolling stop task, a significant difference in successful alert distribution exists between Motive and Samsara [$\chi^2(2) = 48.75, p\text{-value} < 0.0001$].

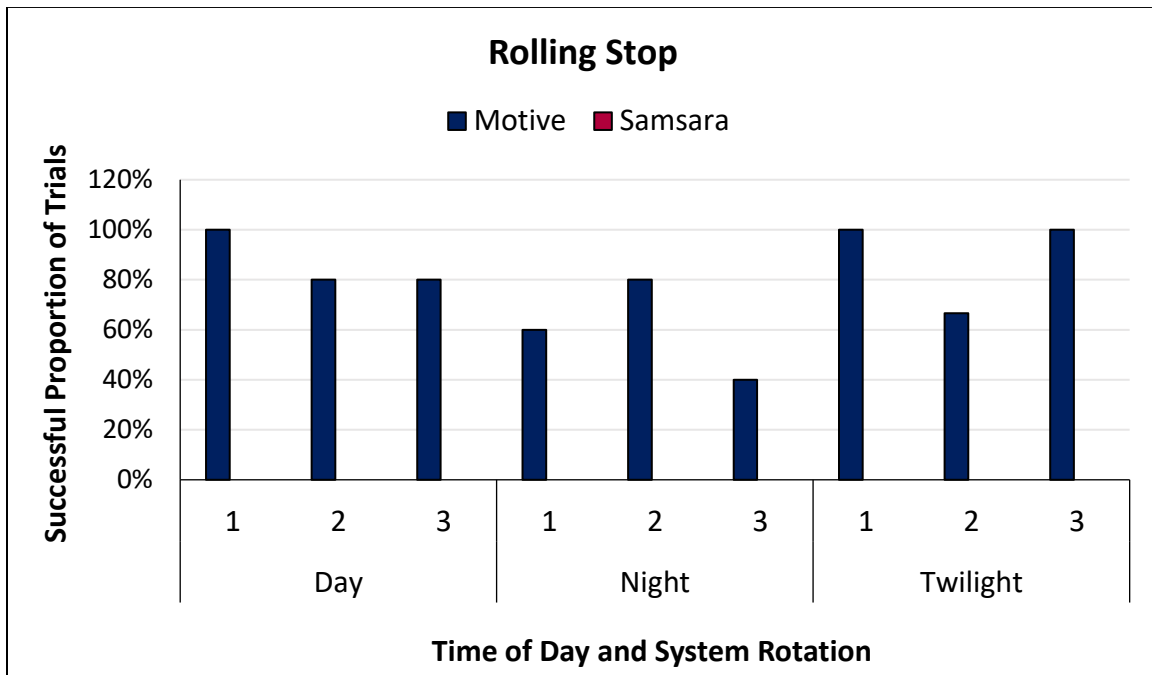


Figure 21 - Successful rolling stop alert percentages during on-road testing across system, time of day, and rotation.

In-cab Alert: No Seatbelt

The percentage of successful alerts during the no seatbelt task was calculated for each of the three systems across the three lighting conditions. The percentages are included in Table 10.

Table 10 - Successful No Seatbelt Alerts by System Type and Time of Day

Time of Day	Total Runs	Lytx	Motive	Samsara*
Day	15	100%	100%	0%
Night	15	100%	100%	0%
Twilight	9	100%	100%	0%
Overall	39	100%	100%	0%

*VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled.

For the 39 no seatbelt trials, the Motive system and Lytx system produced a successful alert 100% of the time, while Samsara did not produce a single successful in-cab alert. Figure 22 displays successful alert percentages for the lighting conditions and illustrates the difference in alerts for each rotation during no seatbelt tasks. Due to the drastic differences in numbers for no seatbelt trials, a chi-square test was run to compare the three systems over all conditions. For the no seatbelt task, a significant difference in successful alert distribution exists between Motive and Samsara and Lytx and Samsara [$\chi^2(2) = 117.00, p\text{-value} < 0.0001$]. However, VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled.

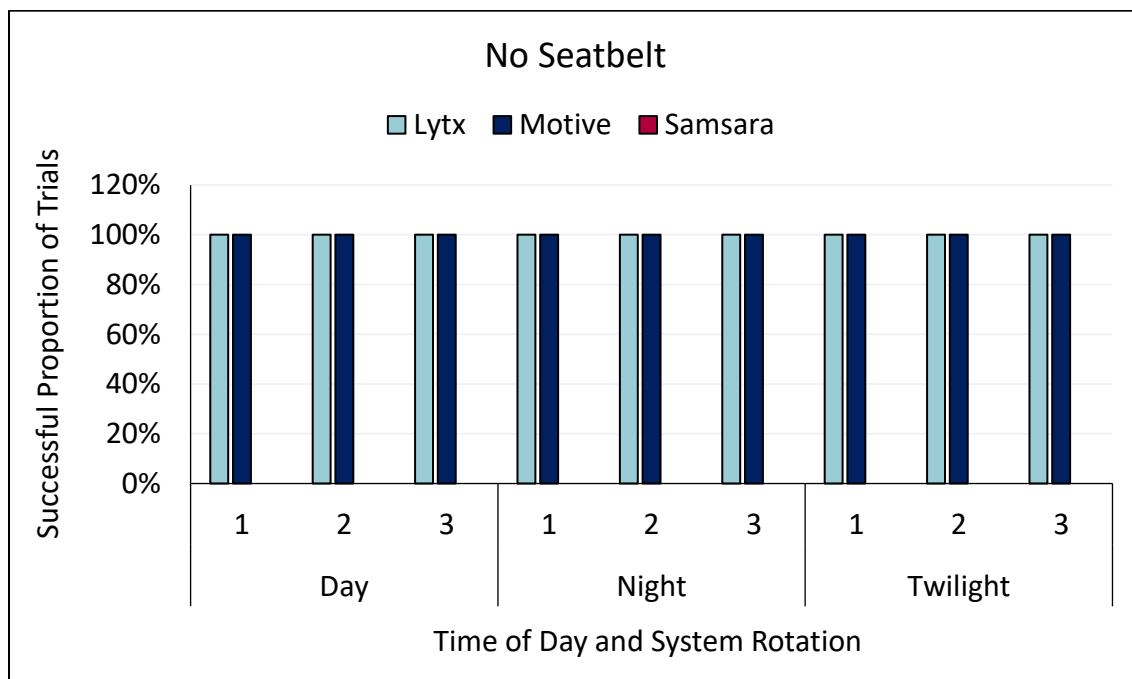


Figure 22 - Successful no seatbelt alert percentages during on-road testing across system, time of day, and rotation.

Dashboard Alerts

For each system and across all task types, the percentage of successful and unsuccessful alerts was categorized based on dashboard or in-cab alerts. If a DMS received a dashboard alert or an in-cab alert, it was counted as an “In-Cab Alert or Dashboard Alert.” If the DMS only triggered

an in-cab alert and there was no associated dashboard alert, then it was counted as “In-Cab Alert Only.” If the DMS neither triggered a successful in-cab nor a dashboard alert, then it was categorized as “Neither In-Cab Alert or Dashboard Alert.”

For Motive, “In-Cab Alert or Dashboard Alert” was the most frequent categorization across all task types. For Samsara, the most frequent option was “No In-Cab or Dashboard Alert” for all task types. For Lytx, the most frequent option was “No In-Cab or Dashboard Alert” for all task types except the no seatbelt task (in-cab alert only was the most frequent). It is also important to note that the Lytx system was not evaluated for rolling stop tasks as the alert was deactivated by the company for all customers during the duration of the study, and VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled.

Table 11–Table 16 provide exact percentages for each of the categorizations.

Table 11 - Overall Dashboard Alert Status for Phone Call Task

Phone Call Alert Status	Lytx	Motive	Samsara
In-cab alert OR dashboard alert	28.21%	94.87%	38.46%
In-cab alert only	23.08%	0.00%	20.51%
Neither in-cab alert or dashboard alert	71.79%	5.13%	61.54%

Table 12 - Overall Dashboard Alert Status for Lap Texting Task

Lap Texting Alert Status	Lytx	Motive	Samsara
In-cab alert OR dashboard alert	7.50%	52.50%	15.00%
In-cab alert only	7.50%	0.00%	12.50%
Neither in-cab alert or dashboard alert	92.50%	47.50%	85.00%

Table 13 - Overall Dashboard Alert Status for 45-Degree Texting Task

45-degree Texting Alert Status	Lytx	Motive	Samsara
In-cab alert OR dashboard alert	18.42%	92.11%	47.37%
In-cab alert only	15.79%	2.63%	31.58%
Neither in-cab alert or dashboard alert	81.58%	7.89%	52.63%

Table 14 - Overall Dashboard Alert Status for Close Following Task

Close Following Alert Status	Lytx	Motive	Samsara
In-cab alert OR dashboard alert	35.90%	100.00%	27.50%
In-cab alert only	23.08%	0.00%	10.00%
Neither in-cab alert or dashboard alert	64.10%	0.00%	72.50%

Table 15 - Overall Dashboard Alert Status for Rolling Stop Task

Rolling Stop Alert Status	Lytx	Motive	Samsara
In-cab alert OR dashboard alert	n/a	76.92%	0.00%
In-cab alert only	n/a	0.00%	0.00%
Neither in-cab alert or dashboard alert	n/a	23.08%	100.00%

Table 16 - Overall Dashboard Alert Status for No Seatbelt Task

No Seatbelt Alert Status	Lytx	Motive	Samsara
In-cab alert OR dashboard alert	100.00%	100.00%	0.00%
In-cab alert only	82.05%	82.05%	0.00%
Neither in-cab alert or dashboard alert	0.00%	0.00%	100.00%

Time to Alert

The time to alert results indicate the time between task start time and a successful alert. There are some instances when the time to alert could not be calculated due to a limited number of successful alerts. Table 17 lists the time to alert for each of the six task types across each of the systems. Due to a known issue with Lytx’s rolling stop alerts, there is no data for time to alert. All other missing data is due to a lack of successful alerts produced by the system. Additionally, VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled.

For time to alert, a Kruskal-Wallis test was used to test for significant differences in time to alert across all rotations and lighting conditions for each task type. In some instances, there was a lack of alert data from the systems, which resulted in the inability to use this statistical testing. Therefore, in Table 18, tasks marked with “n/a” were not tested for significance due to a lack of alert data. Additionally, due to a low number of valid in-cab alert counts, the analyses were not stratified by day.

Table 17 - Overall Time to Alert for Each System Across All Task Types

Task Type	System	N Obs with Time to Alert Data	Time to Alert Mean (s)	Std Dev (s)
Phone Call	Lytx	7	14.50	2.15
	Motive	35	6.89	3.24
	Samsara	17	7.48	4.48
Lap Texting	Lytx	3	16.22	0.15
	Motive	21	7.03	3.93
	Samsara	6	8.65	8.27
45-Degree Texting	Lytx	7	14.50	2.15
	Motive	35	6.89	3.24
	Samsara	17	7.48	4.48
Close Following	Lytx	11	13.12	4.07
	Motive	26	15.46	10.13
	Samsara	7	26.14	8.99
Rolling Stop	Lytx	n/a	n/a	n/a
	Motive	30	3.78	0.51
	Samsara	0	.	.
No Seatbelt	Lytx	39	20.05	4.47
	Motive	39	15.54	7.51
	Samsara	0	.	.

Post-hoc Dunn’s tests for significance were used to follow-up significant Kruskal-Wallis tests and identify which systems differed in time to alert. Results of these tests are included in Table 18. Motive had a significantly shorter time to alert compared to Lytx for 45-degree texting, phone call, and no seatbelt tasks. Motive had a significantly shorter time to alert compared to Samsara for the close following trials. The lap texting and rolling stop conditions had too few observations to test further. As mentioned previously, Lytx has a known issue with rolling stop alerts and was, therefore, removed from the comparison. All other comparisons between Motive and the other systems were not statistically significant.

Table 18 – Overall Time to Alert Results of Post-hoc Dunn's Tests for Significance

Task Type	System Comparison	Dunn's z-test Statistic	p-value	Significance
Phone Call	Motive v. Lytx	3.9480	0.0001	*
	Motive v. Samsara	2.3378	0.0291	
	Lytx v. Samsara	5.2182	<0.0001	*
Lap Texting	n/a	n/a	n/a	n/a
45-Degree Texting	Motive v. Lytx	3.5959	0.0005	*
	Motive v. Samsara	0.5114	0.9136	
	Lytx v. Samsara	3.6519	0.0004	*
Close Following	Motive v. Lytx	-0.1120	>0.9999	
	Motive v. Samsara	-2.8449	0.0067	*
	Lytx v. Samsara	-2.5889	0.0144	*
Rolling Stop	n/a	n/a	n/a	n/a
No Seatbelt	Motive v. Lytx	5.0616	<0.0001	*

**Indicates significance in time to alert between systems.*

Phase Two: User Performance Benchmarking

The following section presents the results obtained from the analyses of survey data focused on acquiring driver feedback on video quality, image quality, alert quality, and hardware quality. Survey responses were included from 188 commercial vehicle drivers who completed at least one question. The drivers answered Likert ratings, rankings, and free response questions. The free response questions provided a place for drivers to elaborate on their ranking of each of the three systems for the criteria being analyzed.

Video Quality

For ratings of daytime video quality, the average ratings were similar across the three systems. Table 19 displays the average participant Likert ratings for all systems by time of day. A generalized linear model was used to assess significant differences between scores treated as continuous variables. The model indicated no statistically significant difference in daytime video quality between Motive and both Lytx ($p=0.5465$) and Samsara ($p=0.1343$). The systems were also tested for the likelihood of being rated 4 or higher on the Likert scale, using a logistic regression model. The model indicated no significant difference in likelihood to be rated 4 or higher between Motive and both Lytx ($p=0.8345$) and Samsara ($p=0.1467$).

For the ratings of nighttime video quality, the average ratings were significant across the three systems. Again, a generalized linear model was used to test for significant differences between scores treated as continuous variables. The model showed a statistically significant difference in nighttime video quality between Motive and both Lytx ($p=0.0022$) and Samsara ($p<0.0001$). The systems were also tested using a logistic regression model for the likelihood of being rated 4 or higher on the Likert scale. The model showed a significant difference between Motive and both

Lytix ($p=0.0113$) and Samsara ($p<0.0001$). Motive was 1.72 times more likely to be rated a 4 or higher compared to Lytx for nighttime video quality [95% CI = (1.13, 2.62)]. Motive was 2.96 times more likely to be rated 4 or higher compared to Samsara for nighttime video quality [95% CI = (1.89, 4.65)].

Table 19 - Video Quality Likert Ratings by Time of Day

Time of Day	System	Total Responses	Average Score	SD of Score	Median	% of Scores 4+ (Agree or Strongly Agree)
Daylight	Motive	152	3.74	1.26	4	71.05%
	Lytix	152	3.66	1.28	4	69.74%
	Samsara	153	3.53	1.23	4	61.44%
Night	Motive	133	3.68	1.14	4	63.91%
	Lytix	132	3.25	1.20	3	46.21%
	Samsara	133	3.00	1.07	3	30.83%

Table 20 displays the overall participant rankings across the three systems by time of day. For the overall rankings, the logistic regression model assessing likelihood to be ranked first in daytime video quality showed a significant difference between Motive and both Lytx ($p=0.0083$) and Samsara ($p=0.0086$). Motive was 1.75 times more likely to be ranked first compared to Lytx for daytime video quality [95% CI = (1.16, 2.66)] and 1.75 times more likely to be ranked first compared to Samsara for daytime video quality [95% CI = (1.15, 2.66)]. The test of likelihood to be ranked first in nighttime video quality showed a significant difference between Motive and both Lytx ($p=0.0209$) and Samsara ($p=0.0047$). Motive was 1.64 times more likely to be ranked first compared to Lytx for nighttime video quality [95% CI = (1.08, 2.49)]. Motive was 1.83 times more likely to be ranked first compared to Samsara for nighttime video quality [95% CI = (1.20, 2.78)].

Table 20 - Video Quality Overall System Rankings by Time of Day

Time of Day	System	Total Responses	Average Score	SD of Score	% Respondents Ranking System #1
Daylight	Motive	136	1.79	0.82	45.59%
	Lytix	134	1.99	0.73	26.87%
	Samsara	134	2.22	0.82	26.87%
Night	Motive	124	1.81	0.85	46.77%
	Lytix	122	1.95	0.73	28.69%
	Samsara	122	2.22	0.82	24.59%

There were 79 responses for daylight video rankings and 64 responses for night video rankings, providing a total of 143 entries for the video quality content analysis. In the following sections participant quotes were altered to reflect the name of the systems (Motive, Samsara, and Lytx) rather than the code names used (Oak, Maple, and Pine). Table 21 displays the number of participants who commented about each theme.

Table 21 - Driver Free Response Themes for Video Quality

Video Quality	Discernable Details	Video Continuity	Brightness	Field of View	Color	Other or N/A
66	11	11	21	18	5	49

Video Clarity

Clarity of the video was the number one area of focus for study participants when rating video quality. Participants remarked heavily about blurriness or graininess visible in the video. From the responses, it appears that participants favored video that was clear and void of any graininess or blur. For example, one participant said, “Lytx is blurry [and the] other two are clear.” This participant ranked Lytx in third place.

Discernible Details

The ability to discern specific details (such as, but not limited to, text on road signs or the hand position of the driver) was commonly mentioned by participants. When mentioning specific details in the video, participants most frequently focused on the ability to read the stop sign that is visible in the footage. Participants favored video where detail was easy to discern. One participant responded with, “Motive [has] clearer resolution [and] you can read the stop sign. Samsara is blurrier, but you can see through the windows on the doors and slightly read the sign. Lytx you can see there’s a sign and assume it’s a stop sign but cannot read it...”

Video Continuity

Many participants commented on the frame rate and continuity of the videos. Participants did not like footage that had a low framerate or “choppy” continuity. Most remarks on video continuity were made when a participant was commenting on a system that had a poor frame rate or continuity. One participant remarked, “Motive has the best quality and smoothest flow. Samsara seems smooth, but the quality is lacking. Lytx has poor quality and terrible flow.”

Brightness

The second most important area of focus for the participants was the brightness of the video. The participants appeared to favor video that had a subjectively “appropriate” brightness for day and greater brightness for night. There were many comments regarding nighttime quality being unsatisfactory, and superior night video quality was called out specifically in multiple responses. For example, “The night clarity seems slightly better with Motive than Samsara.”

Field of View

The third most remarked theme by the participants was field of view from the cameras. Participants mentioned the quality of the width of the views from the cameras, preferring wider angles of view. Participants also mentioned wanting to have side views of the truck. Some examples are “All are too dark, and none have side view[s]” and “Dashcam Lytx has no hood mirrors to see the blind spots from [the] dashcam.”

Color

There were a handful of participants that remarked specifically on the color in the videos. Participants favored videos that had more vibrant color. Following are some remarks from

participants: “The color and sharpness are the best on [the] Motive dashcam” and “Not enough color.”

Image Quality

For ratings of daytime image quality, the average ratings were similar across the three systems. A generalized linear model was built to test for significant differences between scores treated as continuous variables. The model showed no statistically significant difference in daytime image quality between Motive and both Lytx ($p=0.1256$) and Samsara ($p=0.5319$). The systems were also tested for the likelihood of being rated 4 or higher on the Likert scale, using a logistic regression model. The model found a significant difference between Motive and Lytx ($p=0.0376$), where Lytx was 1.55 times more likely to be rated 4 or higher compared to Motive for daytime image quality [95% CI = (1.03, 2.33)]. There was no significant difference between Motive and Samsara ($p=0.6724$).

For the ratings of nighttime image quality, results of a generalized linear model testing scores as continuous variables showed a statistically significant difference in nighttime image quality between Motive and Samsara ($p=0.0001$). There was no significant difference between Motive and Lytx ($p=0.2337$). The systems were also tested for the likelihood of being rated 4 or higher on the Likert scale. The test of likelihood showed a significant difference between Motive and Samsara ($p=0.0003$). Motive was 2.55 times more likely to be rated 4 or higher compared to Samsara for nighttime image quality [95% CI = (1.54, 4.24)]. Motive and Lytx did not show a significant difference in likelihood to be rated 4 or higher for nighttime image quality ($p=0.4263$). Table 22 displays the average participant Likert ratings for all systems by time of day.

Table 22 - Image Quality Likert Rankings by Time of Day

Time of Day	System	Total Responses	Average Score	SD of Score	Median	% of Scores 4+ (Agree or Strongly Agree)
Daylight	Motive	124	3.69	1.11	4	58.06%
	Lytx	125	3.89	0.99	4	73.60%
	Samsara	126	3.60	1.04	4	60.32%
Night	Motive	113	3.40	1.19	4	51.33%
	Lytx	117	3.22	1.19	3	43.59%
	Samsara	114	2.83	0.98	3	24.56%

Table 23 displays the overall participant rankings across the three systems by time of day. For the overall rankings, the logistic regression model assessing the likelihood to be ranked first in daytime image quality showed no significant difference between Motive and both Lytx ($p=0.5696$) and Samsara ($p=0.9655$). The logistic regression model assessing the likelihood to be ranked first in nighttime image quality showed a significant difference between Motive and Samsara ($p=0.0068$). Motive was 1.80 times more likely to be ranked first compared to Samsara for nighttime image quality [95% CI = (1.18, 2.75)]. There was no significant difference between Motive and Lytx ($p=0.0638$) for nighttime image quality.

Table 23 - Image Quality Overall System Rankings by Time of Day

Time of Day	System	Total Responses	Average Score	SD of Score	% Respondents Ranking System #1
Daylight	Motive	117	2.10	0.85	31.62%
	Lytix	118	1.81	0.71	36.44%
	Samsara	118	2.07	0.84	32.20%
Night	Motive	111	1.79	0.83	46.85%
	Lytix	108	1.98	0.77	30.56%
	Samsara	110	2.19	0.80	23.64%

There were 52 responses for daylight image rankings and 50 responses for night image rankings, providing a total of 104 entries for the image quality content analysis. In the following sections participant quotes were altered to reflect the name of the systems (Motive, Samsara, and Lytx) rather than the code names used (Oak, Maple, and Pine). Table 24 displays the number of participants who commented about each theme.

Table 24 - Driver Free Response Themes for Image Quality

Image Clarity	Brightness	Field of View	Color	Other or N/A
40	20	19	8	34

Image Clarity

Image clarity is by far the most important detail to participants in determining image quality. Participants most often mentioned the words “clear” and “clarity” in their free responses for this theme. This is often in reference to one system, specifically, being clear or having clarity over the others. Occasionally, participants made note of one system being blurrier or having more glare than the rest. Following are some example quotes: “Lytx is blurry” and “Motive is very clear.”

Brightness

Most responses that mention brightness for image quality are regarding the level of light for the nighttime images. Some participants believed the images were too dark and others, like the video quality answers, noted which systems had better nighttime imaging. Some participants specifically made mention of “brightness,” but this was often in shorthand or a one-word response. Here are some quotes from participants: “You can see clearer at night with Motive” and “Samsara’s night vision in the cab is the clearest and brightest.”

Field of View

Most participants remark on the systems having a “good” or “better” view. Using context from other responses, it is assumed that this means having a wider field of view and greater visibility for the inside and outside images. One good example from a participant is as follows: “The views were different. Lytx was okay, but [the] view is limited. Motive was a little better, but [the] view was still limited. Samsara has the widest view, and you can see more of your surroundings.”

Color

Like the video quality responses, participants for image quality preferred images with vibrant colors. Following are quotes from participants: “I ranked in order of closeness to lifelike color. Samsara lost its vibrancy” and “It has a more crisp picture. The colors are more bright.”

Alert Quality

The alert ratings are separated by alert type. For the cell phone alert ratings, the average ratings were similar across the three systems. Table 25 displays the average participant Likert ratings for all systems by alert type. A generalized linear model was used to assess significant differences between scores treated as continuous variables. The model showed no statistically significant difference in cell phone alert quality between Motive and both Lytx ($p=0.7110$) and Samsara ($p=0.3873$). The systems were also tested for the likelihood of being rated 4 or higher on the Likert scale, using a logistic regression model. The model showed no significant difference between Motive and both Lytx ($p=0.6004$) and Samsara ($p=0.4007$) in likelihood to be rated 4 or higher.

For the seatbelt alert ratings, a generalized linear model found a significant difference in seatbelt alert quality ratings between Motive and Samsara ($p=0.0002$). There was no significant difference in seatbelt alert quality between Motive and Lytx ($p=0.0717$). The logistic regression model assessing likelihood to be rated 4 or higher showed a significant difference between Motive and both Lytx and Samsara, where Lytx was 1.65 times more likely to be rated 4 or higher compared to Motive, [95% CI = (1.07, 2.54)] and Samsara was 2.05 times more likely to be rated 4 or higher compared to Motive [95% CI = (1.34, 3.15)].

For close following alerts, there was no significant difference in close following alert quality between Motive and both Lytx ($p=0.1110$) and Samsara ($p=0.6463$). The logistic regression model testing likelihood to be rated a 4 or higher also showed no significant difference between Motive and both Lytx ($p=0.2731$) and Samsara ($p=0.8216$) for close following alert quality.

Table 25 - Alert Quality Likert Ratings by Alert Type

Alert Type	System	Total Responses	Average Score	SD of Score	Median	% of Scores 4+ (Agree or Strongly Agree)
Cell Phone Use	Motive	111	3.89	1.06	4	72.97%
	Lytx	111	3.84	1.12	4	68.47%
	Samsara	111	3.77	1.11	4	65.77%
Seatbelt	Motive	105	3.50	1.08	4	51.43%
	Lytx	106	3.75	0.98	4	70.75%
	Samsara	106	4.02	1.00	4	80.19%
Close Following	Motive	105	3.50	1.08	4	51.43%
	Lytx	106	3.75	0.98	4	70.75%
	Samsara	106	4.02	1.00	4	80.19%

Table 26 displays the overall participant rankings across the three systems for each alert type. For the overall cell phone use rankings, the logistic regression model assessing likelihood to be ranked first in cell phone alert quality showed a significant difference between Motive and Lytx

($p=0.0070$). Motive was 1.79 times more likely to be ranked first when compared to Lytx [95% CI = (1.17, 2.73)]. There was no significant difference between Motive and Samsara for cell phone alerts ($p=0.2587$). For the overall seatbelt alert rankings, the likelihood of being ranked first in seatbelt alert quality was not significantly different between Motive and both Lytx ($p=0.1702$) and Samsara ($p=0.1405$). For the overall close following alerts, the likelihood of being ranked first in close following alert quality was significantly different between Motive and both Lytx ($p=0.0475$) and Samsara ($p=0.0132$). Motive was 1.56 times more likely to be ranked first compared to Lytx [95% CI = (1.01, 2.41)] and 1.73 times more likely to be ranked first compared to Samsara for close following alert quality [95% CI = (1.12, 2.68)].

Table 26 - Alert Quality Overall System Rankings by Alert Type

Task Type	System	Total Responses	Average Score	SD of Score	% Respondents Ranking System #1
Cell Phone Use	Motive	107	1.90	0.89	44.86%
	Lytx	106	2.09	0.71	20.75%
	Samsara	106	2.00	0.84	34.91%
Seatbelt	Motive	102	2.08	0.86	33.33%
	Lytx	102	2.10	0.71	20.59%
	Samsara	100	1.81	0.84	46.00%
Close Following	Motive	99	1.81	0.85	47.47%
	Lytx	99	1.92	0.71	29.29%
	Samsara	97	2.27	0.82	23.71%

There were 51 responses for cell phone alerts, 46 responses for no seatbelt alerts, and 41 responses for close following alerts, providing a total of 138 entries for the alert quality content analysis. Table 27 displays the number of participants who commented about each theme.

Table 27 - Driver Free Response Themes for Alert Quality

Volume	Clarity	Tone	Clear Intention	Catches Attention	Politeness	Other
35	20	24	17	9	10	43

In the following sections participant quotes were altered to reflect the name of the systems (Motive, Samsara, and Lytx) rather than the code names used (Oak, Maple, and Pine).

Volume

Volume was the key area of focus for participants regarding DMS alerts. The drivers ranked systems that sounded quiet or muffled lower than the systems deemed “loud” and “clear.” The responses were split between participants noting that alerts were too quiet versus noting that one alert was louder and clearer than the others. Some participants mentioned that an alert or alerts were too loud, but to note, this could be due to the device the participant used to take the survey. In summary, the participants had a clear disposition towards alerts with a higher volume. The following are quotes from the participants: “Samsara is a lot louder, and you can hear it clearly” and “Volume was the reason for my pickings. It has to be loud to go over natural cab noise.”

Clarity

Clarity of alerts is the ability for the driver to clearly hear and interpret the voice for an alert. Participants used language such as “clear” and “detailed” when speaking of alert clarity. Roughly a third of the responses mentioning clarity also mentioned volume. Participants preferred alerts that were clear, most importantly the voice being easily discernible. Following are some examples of responses: “You can understand the voice” and “Samsara is a lot louder, and you can hear it clearly.”

Tone

Tone was the second most important theme for the participants. Tone is in reference to how soothing or agitating the voice and beeps are for the alerts. The participant responses were split roughly into three subcategories: those that ranked based on the least “annoying” or “abrasive” alert, those that ranked based on the most “soothing” or “easy on the ears” alert, and those that simply mentioned “tone” or “voice.” The following are some examples from the participants: “Some are annoying and piercing” and “Motive [has] the softer tone and [is] more enjoyable to hear...”

Clear Intention

The communication of the alert was an important theme for the drivers that took the survey. Most drivers that mentioned the intention of the alert focused on the importance of having a simple and concise message to the driver. Wordier alerts were deemed unpreferable and potentially confusing to the driver. The following are good examples from the participants: “If I heard ‘put down phone’ I’d listen. The other ones seem less direct” and “Motive was direct though quiet, but Samsara was too long of a message...”

Catches Attention

The ability of an alert to catch the attention of the driver was a common theme referenced by the participants. The participants prefer an alert that has a distinct sound and encourages the driver to focus on the message to be conveyed. The following are some examples: “Dash cam Motive gives a more distinct sound and alert. It gives more detail for you to listen to [and] the sound catches you more so than the other two” and “I give Motive [the rank of one] because it does give a more distinct sound that alerts you to pay attention.”

Politeness

Several of participants ranked the alerts based on the politeness of the language used for the respective alerts. Drivers appreciate the use of language such as “please” in an alert as well as not using accusatory language. Drivers also mentioned they prefer their alerts to be more like reminders of desired behavior rather than commands to refrain from a specific behavior. Here are some examples of driver opinion: “Lytx’s message would be better received by the driver as it is not assuming...,” “Samsara gives orders, not reminders. No good,” and “Again, I like being reminded, and I love the ‘please’!”

Hardware Design

For the ratings of hardware design, the average scores were similar across systems. Table 28 displays the average participant Likert ratings for all systems. A generalized linear model was used to assess differences in hardware design rating, as treated as a continuous variable. The model indicated there was no significant difference in hardware design ratings between Motive

and Lytx ($p=0.4703$). There was a significant difference in hardware design ratings between Motive and Samsara ($p=0.0115$). The systems were also compared for differences in likelihood of being rated 4 or higher on the Likert scale. The logistic regression model results showed no significant difference between Motive and both Lytx ($p=0.7493$) and Samsara ($p=0.0617$).

Table 28 - Hardware Quality Likert Ratings

System	Total Responses	Average Score	SD of Score	Median	% of Scores 4+ (Agree or Strongly Agree)
Motive	101	3.76	0.95	4	69.31%
Lytx	101	3.86	0.93	4	72.28%
Samsara	101	3.42	1.07	4	52.48%

Table 29 displays the overall participant rankings for hardware design across the three systems. For the overall rankings, the logistic regression model assessing likelihood to be ranked first in hardware quality showed no significant difference between Motive and both Lytx ($p=0.0899$) and Samsara ($p=0.0604$).

Table 29 - Hardware Quality Overall System Rankings

System	Total Responses	Average Score	SD of Score	% Respondents Ranking System #1
Motive	98	1.73	0.67	38.78%
Lytx	97	1.79	0.75	40.21%
Samsara	96	2.45	0.83	21.88%

There were a total of 42 responses for the hardware quality content analysis. Table 30 displays the number of participants who commented about each theme.

Table 30 - Driver Free Response Themes for Hardware Quality

Size	Aesthetics	Other
26	6	12

In the following sections participant quotes were altered to reflect the name of the systems (Motive, Samsara, and Lytx) rather than the code names used (Oak, Maple, and Pine).

Size

The size of the hardware was the key theme for discussion for participant responses regarding hardware quality. Drivers had concerns with DMSs being too large or bulky and inhibiting visibility through the front windshield. The consensus was that the smaller the DMS the more desirable it was. Following are some quotes from the participants: “The smaller the better,” “Smaller and doesn’t hinder vision,” and “As a driver you need less on the windshield blocking [your] view...”

Aesthetics

Aesthetics were mentioned by a handful of participants in the content analysis; therefore, it was included as a theme for hardware quality. The language used in determining what the participants, specifically, found aesthetically pleasing was vague or inconsistent for determining subthemes for aesthetic preferences. The following are some examples of the language used: “Slim design looks better” and “I like that both Motive and Lytx have a centered camera. All camera body designs look great!”

Limitations

While the study provided a scientific benchmarking effort between the three DMSs, it is important to note the study assumptions and limitations. These limitations inform the accurate interpretation of the presented results and assess the ability of the results to be extrapolated further. This section contextualizes the results of the benchmarking effort by identifying key limitations in analysis methods, experimental methods, and comparisons.

For the survey analysis, there was a question regarding drivers' experience with DMSs to understand if there were any biases in users who currently use the Motive system when completing the survey. The planned sampling strategy included collecting near-equal response counts from users of each of the three systems plus non-DMS users. However, the response rate was higher for certain system users; therefore, the final sample does not include near-equal response counts from users of all three systems and non-system users. To counter this limitation, multiple attempts were made to collect data from all system users, the data was analyzed with system user as a covariate when possible, and summary statistics were presented by current Motive users and current non-users. Additionally, the survey was anonymous and administered virtually, which impacted the participants' ability to ask clarifying questions or troubleshoot technical difficulties. All survey responses were included, even if a driver chose not to complete the entire survey.

During experimentation, the DMSs were tested in three different locations on the windshield. These locations were determined by a certified, third-party installer who followed the installation requirements provided by each vendor. Although these three locations met vendor requirements, the results indicated that location and angle of the device played a significant role in the variance in performance across various tasks (for some tasks/DMSs). The significance of positioning highlights the importance of clear guidelines and recommendations to customers regarding optimal DMS placement. Customer results may vary depending on the location and angle of the device installed in their specific vehicle.

To increase the validity of the results, experimental controls were put in place to mitigate the influence of weather factors and driver characteristics during testing. The DMSs were exclusively tested in dry weather conditions during one time of the year. Camera-based DMSs rely on computer vision to detect objects in the environment. Rain, snow, fog, or other obstructions on the windshield may impact the ability of the exterior camera to detect rolling stops or close following. The impacts of these conditions were not tested; therefore, the results of these functions may only apply to DMSs used in dry weather conditions. The operability of the interior camera is not impacted by weather factors; however, it is important to consider the ability of the DMS for varying driver characteristics. The same test driver was used across all trials during experimentation. The test driver was a Caucasian male who wore glasses. Additionally, during the seatbelt trials, the driver wore a white shirt to increase contrast. Various races, genders, and clothing types were not tested. Therefore, the results of the interior camera may differ under different driver characteristics.

Maintaining consistency in the abilities and functionalities of the DMSs under comparison is essential for drawing meaningful conclusions. This ensures that any observed differences or similarities in performance can be more confidently attributed to their inherent characteristics, rather than disparities in their capabilities. For the rolling stop alerts, it is important to note that

Lytx intentionally disabled their rolling stop in-cab alerts due to a known issue with false alarms. With this in mind, a fair comparison between the three devices could not be made, and Lytx was removed from the rolling stop alert benchmark. This information was not published on their website and was gathered by calling the company directly. Although the VTTI team verified with each vendor, respectively, that all alerts relevant to the study were turned on and functioning properly, it is possible that other DMSs were experiencing unknown errors that negatively impacted their performance results. For example, VTTI was informed after the study was completed that the seatbelt alert for the Samsara device was not properly enabled. Furthermore, when interpreting the results for the “time to alert” section, it is important to note that companies intentionally have different thresholds for alerts, and many can be adjusted by the customer. For example, if a result indicates that one company was “faster” to issue an alert for a behavior, this does not imply that a faster alert is better. Alert thresholds are adjustable by the customer due to complaints of too frequent alerts being bothersome to drivers. Therefore, this metric should be interpreted by the customer in the context of each system’s thresholds.

Conclusions

The following conclusions highlight significant results from the two-phases of the study that pertain to the Motive system. Conclusions describing statistically insignificant results are not included.

Conclusion #1: Motive was more likely to issue a successful in-cab alert than Samsara and Lytx for phone calls, lap texting, 45-degree texting, close following, and rolling stops.

From the relative risk equation used to calculate the success of in-cab alerts for all task types and across all lighting conditions, it was found that Motive was 3.36 times more likely to produce an alert for the phone call task than Lytx and 2.47 times more likely than Samsara. Additionally, Motive was 7 times more likely to produce an alert for the lap texting task than Lytx and 3.50 times more likely than Samsara. Motive was also 5.14 times more likely to produce an alert for the 45-degree texting task than Lytx and 2 times more likely than Samsara. For the close following task, Motive was 2.89 times more likely to produce an alert than Lytx and 3.71 times more likely than Samsara. Lastly, due to the lack of successful alert data between all three systems for rolling stops, a chi-square test was run to compare only Motive and Samsara over all conditions. For the rolling stop task, a significant difference in successful alert distribution exists between Motive and Samsara.

Conclusion #2: For Motive, “in-cab alert or dashboard alert” was the most frequent result for all tasks (phone calls, lap texting, 45-degree texting, close following, rolling stops, and no seatbelt)

For phone call alerts, Motive gave a successful in-cab or dashboard alert 94.87% of the time, whereas Samsara and Lytx more frequently gave neither an in-cab or dashboard alert. For lap texting alerts, Motive gave a successful in-cab or dashboard alert 52.50% of the time while Samsara and Lytx more frequently gave neither an in-cab or dashboard alert.

Conclusion #3: Motive had a significantly shorter time to alert compared to Lytx for phone call tasks, 45-degree texting, and no seatbelt and a significantly shorter time to alert compared to Samsara for close following.

When compared to Lytx, Motive had a 7.61 second shorter time to alert for phone call, a 7.19 second shorter time to alert for 45-degree texting, and a 4.51-second shorter time to alert for no seatbelt. When compared to Samsara, Motive had a 10.68-second shorter time to alert for close following.

Conclusion #4: Motive was most likely to be ranked number one for daytime and nighttime video when compared to Lytx and Samsara.

Motive was 1.75 times more likely to be ranked first when compared to both Lytx and Samsara for daytime video quality. Motive was 1.64 times more likely to be ranked first compared to Lytx and 1.83 times more likely to be ranked first compared to Samara for nighttime video quality.

Conclusion #5: Motive was more likely to be ranked number one for nighttime images when compared to Samsara.

Motive was 1.8 times more likely to be ranked first when compared to Samsara for nighttime image quality.

Conclusion #6: Motive was more likely to be ranked number one for cell phone alerts when compared to Lytx and close following alerts when compared to Lytx and Samsara.

Motive was 1.79 times more likely to be ranked first when compared to Lytx for cell phone alert quality. Motive was 1.56 times more likely to be ranked first compared to Lytx and 1.73 times more likely to be ranked first compared to Samsara for close following alert quality.

Conclusion #7: Motive had a significantly higher rating for hardware design when compared to Samsara.

For the rating of hardware design and quality, Motive had an average Likert rating of 3.76 (1 being strongly disagree and 5 being strongly agree) with a standard deviation of 0.95 whereas Samsara held an average Likert rating of 3.42 with a standard deviation of 1.07.

Opportunities

The previous conclusions are specific to the Motive system. The following points detail areas where other systems performed better or equally well to the Motive system. These may be used as opportunities for improvement for the Motive DMS.

Opportunity #1: Time of day and system rotation was a significant factor in successful in-cab alerts for certain tasks.

System rotation and time of day were noted as significant predictors for successful in-cab alerts. This emphasizes the need for DMS providers to ensure their installation guidelines provide specific locations for placement that best fits their system's device. Additionally, customers should be aware that lighting impacts the success of in-cab alerts, specifically at nighttime.

Opportunity #2: Lytx and Samsara were more likely to be rated 4 or higher on seat belt use alert quality compared to Motive.

The logistic regression model used to assess the likelihood of one system to be ranked 4 or higher compared to the other two found that Lytx was 1.65 times more likely to be rated 4 or higher compared to Motive [95% CI = (1.07, 2.54)] and Samsara was 2.05 times more likely to be rated 4 or higher compared to Motive [95% CI = (1.34, 3.15)]. From the content analysis some participants noted that Motive's seatbelt alert was "soothing," "soft," and "calm." It was found in the content analysis that participants preferred audio alerts that grabbed the driver's attention. Therefore, Motive should be mindful of the tonality and vocal delivery of their system's voice alert.

Opportunity #3: Lytx was 1.55 times more likely to be rated 4 or higher on daytime image quality compared to Motive.

The generalized linear model found a significant difference between Motive and Lytx ($p=0.0376$) where Lytx was 1.55 times more likely to be rated 4 or higher compared to Motive for daytime image quality [95% CI = (1.03, 2.33)]. This means that participants were more likely to express strong positive emotions towards the daytime image quality from Lytx's system than Motive's system. Many factors could impact participants' ratings of images, so an explicit recommendation is not given; however, Motive may consider reviewing differences in image quality such as clarity, brightness, and field of view.

Opportunity #4: Motive was equally likely to be rated 4 or higher on daytime video quality compared to Lytx and Samsara.

The linear regression model used to determine significant difference for the likelihood of being rated 4 or higher between Motive and both Lytx and Samsara found no significant difference. As mentioned above, many factors impact the ratings of visual media. Motive may consider possible differences in video quality compared to competitors such as visual clarity, brightness, and field of view, all of which were expressed as key themes in the video quality content analysis.

Opportunity #5: Motive was equally likely to be ranked #1 in daytime image quality compared to Lytx and Samsara.

The linear model identified no significant difference between rankings of daytime image quality between Motive and both Lytx and Samsara. Therefore, Motive may consider possible factors that influence participants' ratings of image quality such as clarity, brightness, and field of view.

References

- Boyle, L., Guo, F., Hammond, R., Hanowski, R., & Soccolich, S. (2016). *Performance assessment of an onboard monitoring system for commercial motor vehicle drivers: A field operational test* (Report No: FMCSA-RRR-15-019). Federal Motor Carrier Safety Administration.
- Doerzaph, Z. R. (2007). *Development of a threat assessment algorithm for intersection collision avoidance systems* (Doctoral dissertation, Virginia Tech).
- Federal Motor Carrier Safety Administration. (2006). *Report to Congress on the Large Truck Crash Causation Study* (Report No. MC-R/MC-RRA). Federal Motor Carrier Safety Administration.
- Hammond, R. L., Soccolich, S. A., Guo, F., Glenn, T. L., Hanowski, R. J., & Han, S. (2021). *Analysis of naturalistic driving data to assess distraction and drowsiness in drivers of commercial motor vehicles* (No. FMCSA-RRR-20-003). Federal Motor Carrier Safety Administration.
- Hickman, J. S., & Hanowski, R. J. (2011). Use of a video monitoring approach to reduce at-risk driving behaviors in commercial vehicle operations. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(3), 189-198.
- Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). *Driver distraction in commercial vehicle operations* (No. FMCSA-RRT-09-042). Federal Motor Carrier Safety Administration.
- Sarkar, A., Engstrom, J., & Hanowski, R. (2022). *Analysis of car cut-ins between trucks based on existing naturalistic driving data*. National Surface Transportation Safety Center for Excellence. <http://hdl.handle.net/10919/109365>
- Singh, S. (2015). *Critical reasons for crashes investigated in the national motor vehicle crash causation survey* (No. DOT HS 812 115). National Highway Traffic Safety Administration. <http://www-nrd.nhtsa.dot.gov/Pubs/812115.pdf>

Appendix

Additional Resources for Phase One: Technical Performance Benchmarking

The following section presents additional analysis results and model output for Phase One: Technical Performance Benchmarking. The information is provided for reference purposes.

Phone Call Task

For the phone call task, a logistic regression model assessed whether the different explanatory study variables had a statistically significant effect on predicting a successful alert. Table 31 presents the model Type III Analysis of Effects results for the phone call task. Table 32 presents the model estimate values for each of the variables in the phone call task.

Table 31 - Type 3 Analysis of Effects for Phone Call Task Model

Effect	DF	Wald Chi-Square	p-value
System*	2	25.39	<.0001
Rotation	2	2.85	0.2410
Time of Day*	2	18.52	<.0001

*Indicates significance at $\alpha = 0.05$.

Table 32 - Analysis of Maximum Likelihood Estimates for Phone Call Task Model

Model Parameter	Parameter Level	DF	Estimate	Standard Error	p-value
Intercept		1	3.45	0.80	<.0001
System	Lytx*	1	-5.02	1.01	<.0001
System	Samsara*	1	-4.35	0.96	<.0001
Rotation	1	1	0.22	0.38	0.5633
Rotation	2	1	0.44	0.39	0.2579
Time of Day	Day*	1	1.83	0.43	<.0001
Time of Day	Night	1	-0.32	0.40	0.4332

*Indicates significance at $\alpha = 0.05$.

Lap Texting

For the lap texting task, a logistic regression model assessed whether the different explanatory study variables had a statistically significant effect on predicting a successful alert. Table 33 presents the model Type III Analysis of Effects results for the lap texting task. Table 34 presents the model estimate values for each of the variables in the lap texting task.

Table 33 - Type 3 Analysis of Effects for Lap Texting Task Model

Effect	DF	Wald Chi-Square	p-value
System*	2	21.34	<.0001
Rotation*	2	8.91	0.0116
Time of Day*	2	13.10	0.0014

*Indicates significance at $\alpha = 0.05$.

Table 34 - Analysis of Maximum Likelihood Estimates for Lap Texting Task

Model Parameter	Parameter Level	DF	Estimate	Standard Error	p-value
Intercept		1	0.10	0.38	0.7877
System	Lytx*	1	-3.50	0.84	<.0001
System	Samsara*	1	-2.55	0.70	0.0003
Rotation	1	1	-1.17	0.47	0.0121
Rotation	2	1	0.05	0.39	0.8940
Time of Day	Day*	1	1.48	0.42	0.0004
Time of Day	Night	1	-0.93	0.43	0.0315

*Indicates significance at $\alpha = 0.05$.

45-Degree Texting

For the 45-degree texting task, a logistic regression model assessed whether the different explanatory study variables had a statistically significant effect on predicting a successful alert. Table 35 presents the model Type III Analysis of Effects results for the 45-degree texting task. Table 36 presents the model estimate values for each of the variables in the 45-degree texting task.

Table 35 - Type 3 Analysis of Effects for 45-Degree Texting Task Model

Effect	DF	Wald Chi-Square	p-value
System*	2	28.26	<.0001
Rotation*	2	7.57	0.0227
Time of Day*	2	10.71	0.0047

*Indicates significance at $\alpha = 0.05$.

Table 36 - Analysis of Maximum Likelihood Estimates for 45-Degree Texting Task

Model Parameter	Parameter Level	DF	Estimate	Standard Error	p-value
Intercept		1	2.91	0.70	<.0001
System	Lytx*	1	-5.10	0.96	<.0001
System	Samsara*	1	-3.32	0.84	<.0001
Rotation	1	1	0.75	0.39	0.0517
Rotation	2	1	0.38	0.38	0.3243
Time of Day	Day*	1	1.34	0.42	0.0013
Time of Day	Night	1	-0.01	0.38	0.9743

*Indicates significance at $\alpha = 0.05$.

Close Following

For the close following task, a logistic regression model assessed whether the different explanatory study variables had a statistically significant effect on predicting a successful alert. Table 37 presents the model Type III Analysis of Effects results for the close following task. Table 38 presents the model estimate values for each of the variables in the close following.

Table 37 - Type 3 Analysis of Effects for Close Following Task Model

Effect	DF	Wald Chi-Square	p-value
System*	2	20.08	<.0001
Rotation	2	4.52	0.1044
Time of Day	2	0.73	0.6935

*Indicates significance at $\alpha = 0.05$.

Table 38 - Analysis of Maximum Likelihood Estimates for Close Following Task

Model Parameter	Parameter Level	DF	Estimate	Standard Error	p-value
Intercept		1	0.76	0.36	0.0339
System	Lytx*	1	-1.73	0.51	0.0008
System	Samsara*	1	-2.34	0.56	<.0001
Rotation	1	1	0.47	0.30	0.1240
Rotation	2*	1	-0.67	0.33	0.0391
Time of Day	Day	1	0.11	0.30	0.7125
Time of Day	Night	1	-0.26	0.30	0.3934

*Indicates significance at $\alpha = 0.05$.

Additional Resources for Phase Two: User Performance Benchmarking

The following section presents additional analysis results and model output for Phase Two: Technical Performance Benchmarking. The information is provided for reference purposes.

Video Quality

The following section presents tables relevant to the analysis of video quality, for daytime and nighttime videos, and by rating and overall order ranking questions.

Daytime Video Quality Ratings

Table 39 – Frequency counts for Likert scale ratings of daytime video quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	16	10	18	61	47	152
Lytx	17	14	15	64	42	152
Samsara	15	17	27	60	34	153

Table 40 - Model results for Likert scale ratings for daytime video quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.6588	0.1215	3.4207	3.8969	907.07	<.0001
System Evaluated	LytX	1	-0.0865	0.1435	-0.3677	0.1947	0.36	0.5465
System Evaluated	Samsara	1	-0.2144	0.1432	-0.4952	0.0663	2.24	0.1343
System Evaluated	Motive	0	0.0000	0.0000	0.0000	0.0000	.	.
Current Motive User	No	1	0.1496	0.1181	-0.0819	0.3812	1.60	0.2053
Current Motive User	Yes	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		1	1.2507	0.0414	1.1721	1.3344		

Table 41 - Model results for Likert scale ratings of daytime video quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.2809	0.1483	3.5884	0.0582	1.324
System Evaluated	LytX	1	-0.0437	0.2090	0.0437	0.8345	0.957
System Evaluated	Samsara	1	-0.3021	0.2082	2.1067	0.1467	0.739
Current Motive User	No	1	0.1657	0.0856	3.7478	0.0529	1.180

Table 42 – Odds ratio calculations comparing Motive to LytX and Samsara systems for ratings of 4 or higher vs. 3 or lower in daytime video quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	1.0446	0.2183	0.05	0.6935	1.5735	0.0437	0.8345
Samsara vs Motive	1.3527	0.2816	0.05	0.8996	2.0342	2.1067	0.1467

Ratings of daytime video quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 43 and Table 44 present the frequency counts of daytime video quality ratings for the stratified assessment. Figure 23 illustrates the distribution of daytime video quality ratings for all three systems, as rated by current Motive users. Figure 24 illustrates the distribution of ratings for current Motive users and non-Motive users for daytime video quality.

Table 43 – Frequency counts for Likert scale ratings of daytime video quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	7	5	10	23	21	66
LytX	10	7	6	21	21	65
Samsara	9	6	14	22	15	66

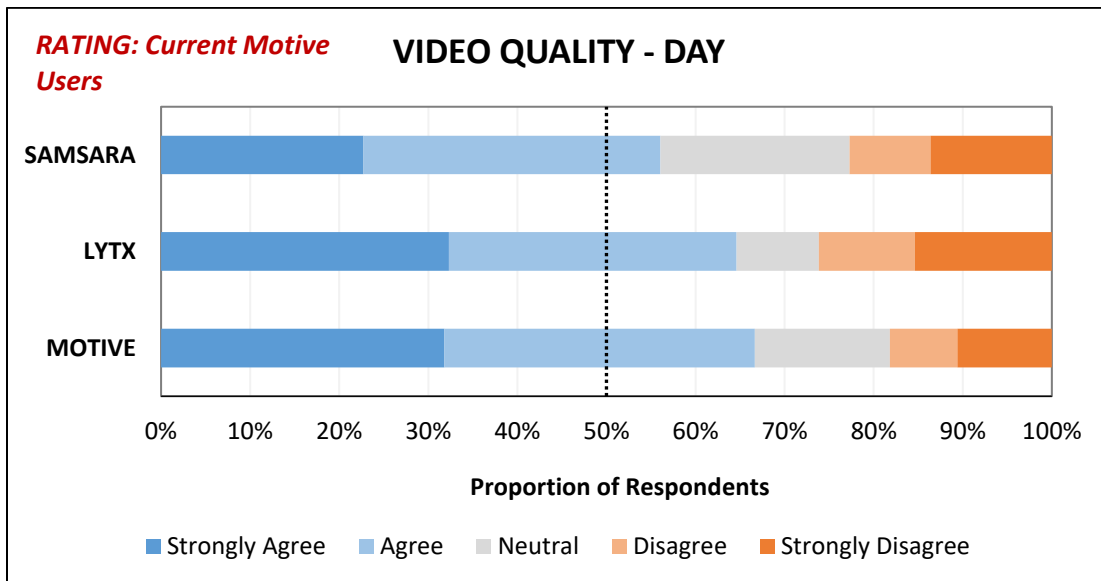


Figure 23. Likert scale ratings for daytime video quality comparing current Motive users' opinions of all three systems.

Table 44 – Frequency counts for Likert scale ratings of daytime video quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	9	5	8	38	26	86
Yes	7	5	10	23	21	66

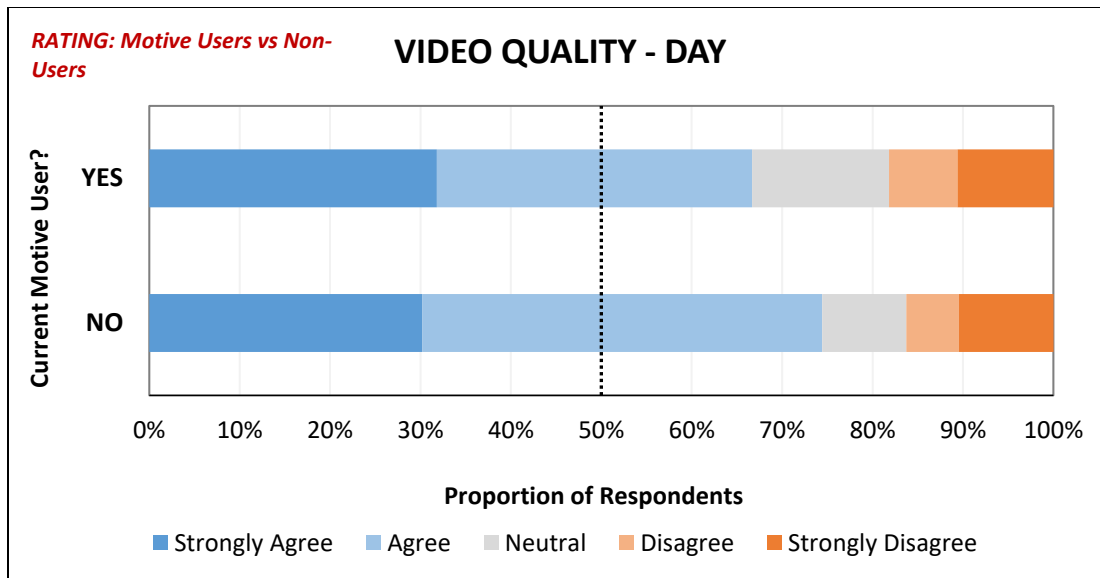


Figure 24 - Likert scale ratings for daytime video quality comparing current Motive users and non-users.

Nighttime Video Quality Ratings

Table 45 – Frequency counts for Likert scale ratings of nighttime video quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	8	13	27	51	34	133
LytX	10	30	31	39	22	132
Samsara	11	31	50	29	12	133

Table 46 - Model results for Likert scale ratings for nighttime video quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.6890	0.1194	3.4551	3.9230	955.18	<.0001
System Evaluated	Lytx	1	-0.4268	0.1393	-0.6998	-0.1537	9.38	0.0022
System Evaluated	Samsara	1	-0.6767	0.1391	-0.9492	-0.4041	23.68	<.0001
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	-0.0210	0.1154	-0.2472	0.2051	0.03	0.8553
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	1.1340	0.0402	1.0579	1.2156		

Table 47 - Model results for Likert scale ratings of nighttime video quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.2023	0.1472	1.8892	0.1693	0.817
System Evaluated	Lytx	1	-0.5420	0.2140	6.4129	0.0113	0.582
System Evaluated	Samsara	1	-1.0862	0.2297	22.3692	<.0001	0.338
Current Motive User	No	1	0.0777	0.0925	0.7062	0.4007	1.081

Table 48 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in nighttime video quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytx vs Motive	1.7194	0.3680	0.05	1.1303	2.6155	6.4129	0.0113
Samsara vs Motive	2.9629	0.6804	0.05	1.8890	4.6472	22.3692	<.0001

Ratings of nighttime video quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 49 and Table 50 present the frequency counts of nighttime video quality ratings for the stratified assessment. Figure 25 illustrates the distribution of nighttime video quality ratings

for all three systems, as rated by current Motive users. Figure 26 illustrates the distribution of ratings for current Motive users and non-Motive users for nighttime video quality.

Table 49 – Frequency counts for Likert scale ratings of nighttime video quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	2	6	12	17	18	55
Lytix	4	14	11	20	6	55
Samsara	6	10	23	10	6	55

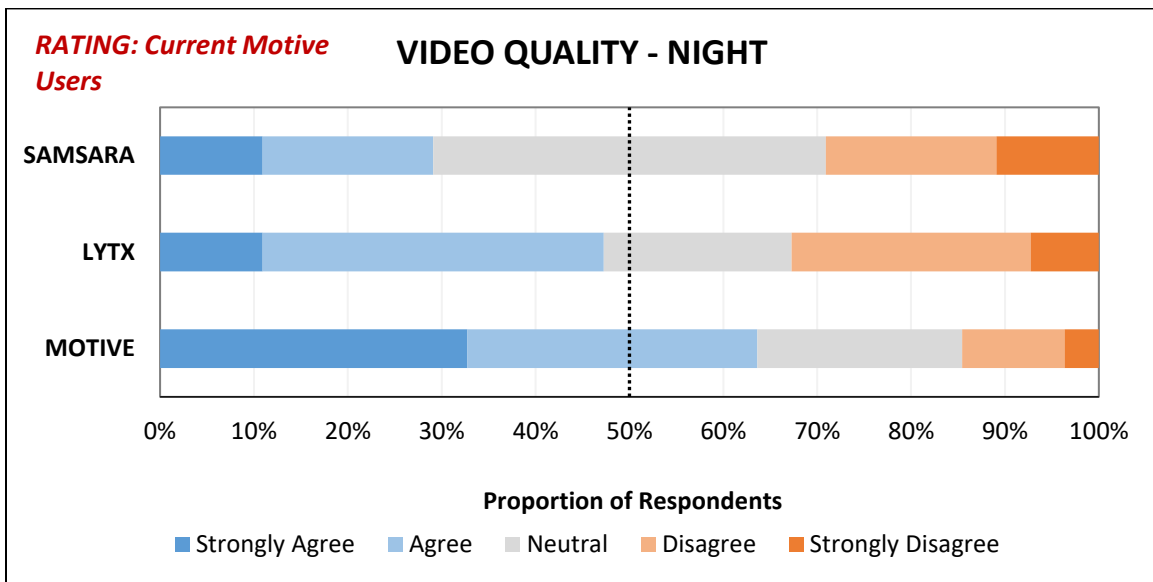


Figure 25. Likert scale ratings for nighttime video quality comparing current Motive users' opinions of all three systems.

Table 50 – Frequency counts for Likert scale ratings of nighttime video quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	6	7	15	34	16	78
Yes	2	6	12	17	18	55

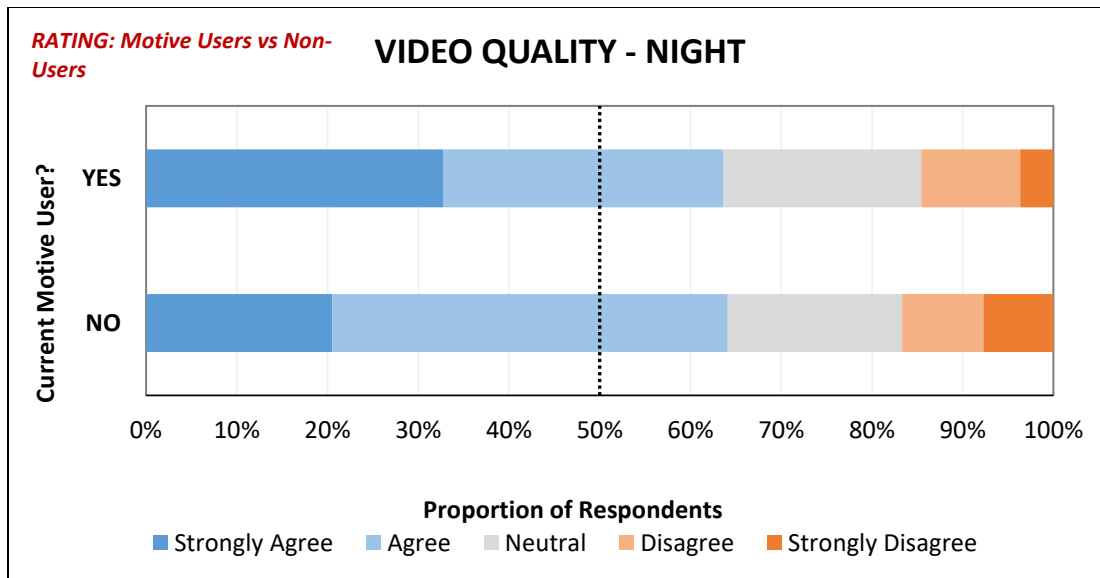


Figure 26 - Likert scale ratings for nighttime video quality comparing current Motive users and non-users.

Daytime Video Quality Overall Order Rankings

Table 51 – Frequency counts for overall order ranking of daytime video quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	62	40	34	136
LytX	36	63	35	134
Samsara	36	33	65	134

Table 52 - Model results for overall order ranking of daytime video quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.3916	0.1516	6.6697	0.0098	1.479
System Evaluated	LytX	1	-0.5617	0.2129	6.9636	0.0083	0.570
System Evaluated	Samsara	1	-0.5596	0.2128	6.9136	0.0086	0.571
Current Motive User	No	1	-0.0962	0.0873	1.214	0.2705	0.908

Table 53 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in daytime video quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	<i>p</i> -value
LytX vs Motive	1.7537	0.3733	0.05	1.1555	2.6615	6.9636	0.0083
Samsara vs Motive	1.7499	0.3724	0.05	1.1531	2.6556	6.9136	0.0086

Rankings of daytime video quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 54 and Table 55 present the frequency counts of daytime video quality rankings for the stratified assessment. Figure 27 illustrates the distribution of daytime video quality ranking for all three systems, as ranked by current Motive users. Figure 28 illustrates the distribution of ranking for current Motive users and non-Motive users for daytime video quality.

Table 54 – Frequency counts for overall order ranking of daytime video quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	25	17	14	56
LytX	17	25	17	59
Samsara	16	16	25	57

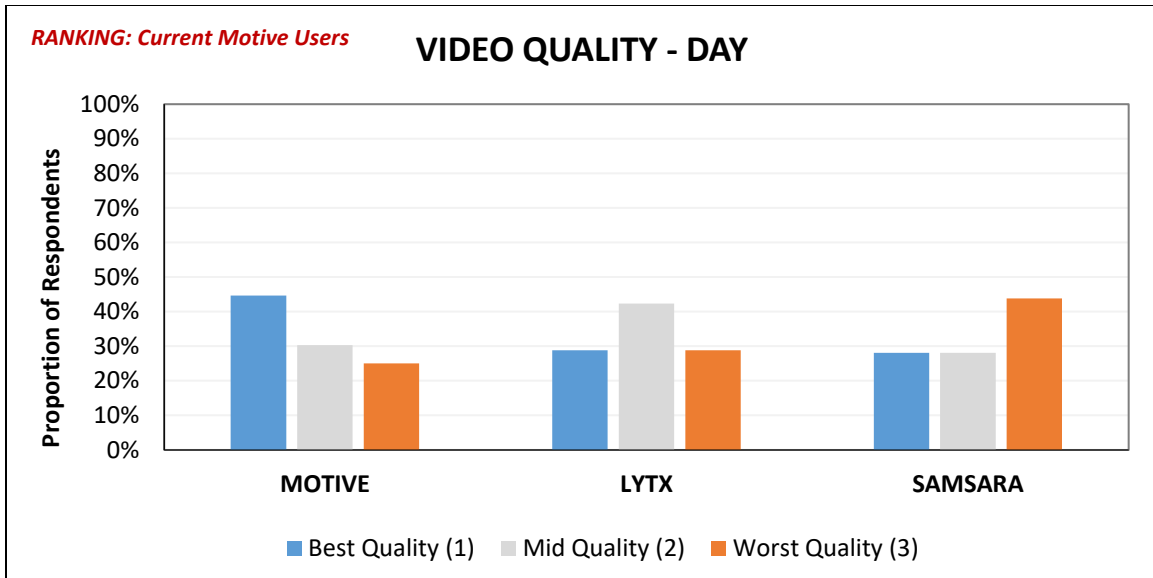


Figure 27 - Ranking for daytime video quality comparing current Motive users' opinions of all three systems.

Table 55 – Frequency counts for ranking of daytime video quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	37	23	20	80
Yes	25	17	14	56

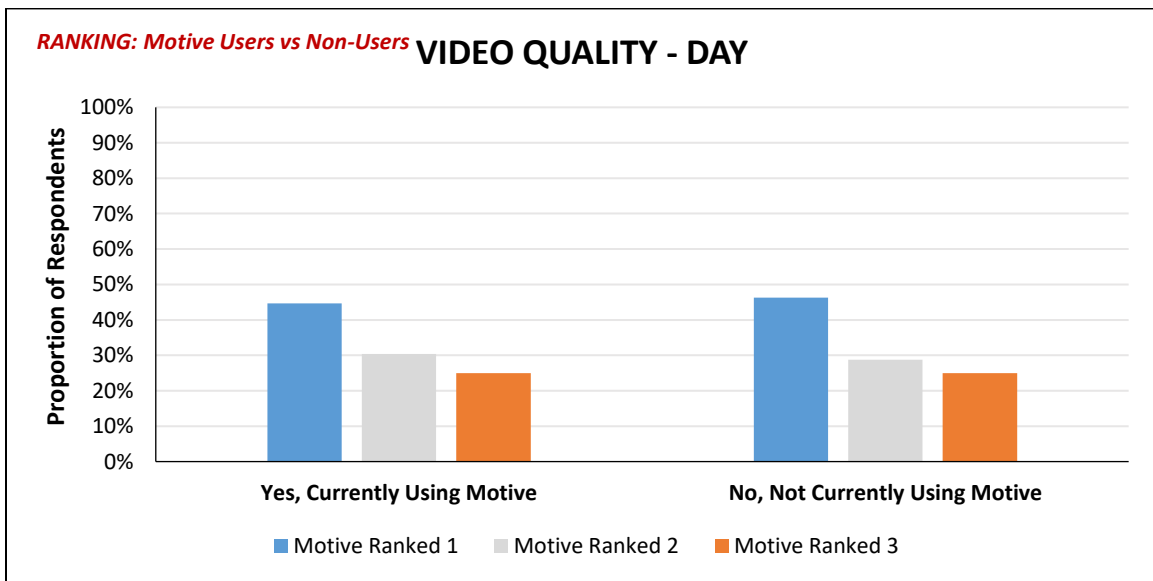


Figure 28 - Ranking for daytime video quality comparing current Motive users and non-users.

Nighttime Video Quality Overall Order Rankings

Table 56 – Frequency counts for overall order ranking of nighttime video quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	58	31	35	124
Lytix	35	58	29	122
Samsara	30	35	57	122

Table 57 - Model results for overall order ranking of nighttime video quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.5862	0.1542	14.4588	0.0001	1.797
System Evaluated	Lytix	1	-0.4936	0.2137	5.3360	0.0209	0.610
System Evaluated	Samsara	1	-0.6036	0.2136	7.9858	0.0047	0.547
Current Motive User	No	1	-0.0414	0.0873	0.2248	0.6354	0.959

Table 58 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in nighttime video quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytix vs Motive	1.6383	0.3501	0.05	1.0777	2.4905	5.3360	0.0209
Samsara vs Motive	1.8288	0.3906	0.05	1.2032	2.7796	7.9858	0.0047

Rankings of nighttime video quality were further stratified by DMS experience, with assessment of Motive users' opinions of all three systems and a comparison of Motive users to non-Motive users. Table 59 and Table 60 present the frequency counts of nighttime video quality rankings for the stratified assessment. Figure 29 illustrates the distribution of nighttime video quality ranking for all three systems, as ranked by current Motive users. Figure 30 illustrates the distribution of ranking for current Motive users and non-Motive users for daytime video quality.

Table 59 – Frequency counts for overall order ranking of nighttime video quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	27	14	12	53
Lytix	16	25	12	53
Samsara	10	15	28	53

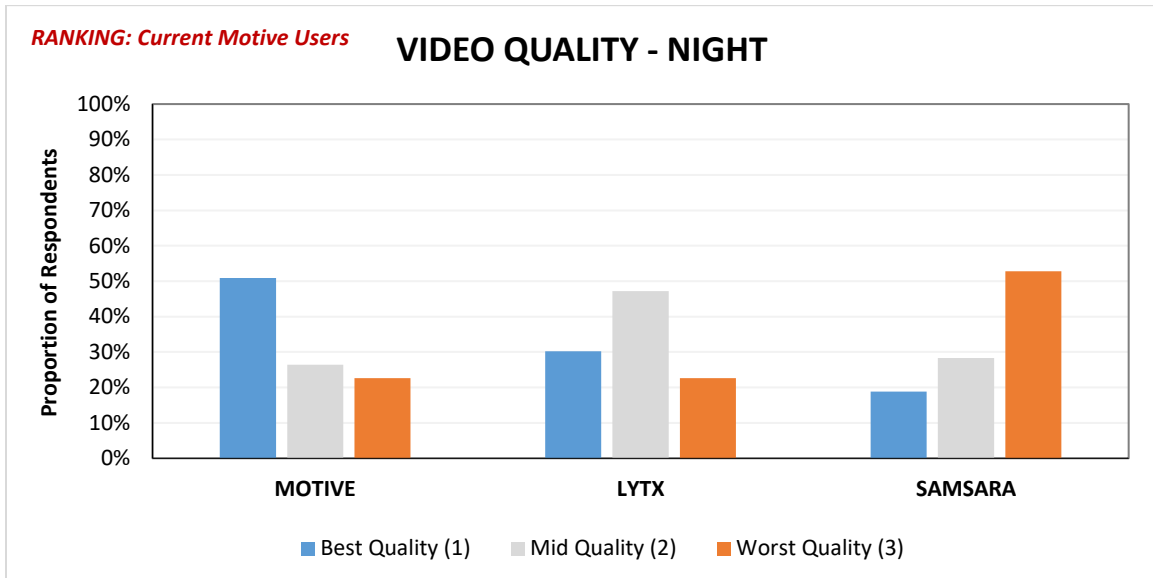


Figure 29 - Ranking for nighttime video quality comparing current Motive users' opinions of all three systems.

Table 60 – Frequency counts for ranking of nighttime video quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	31	17	23	71
Yes	27	14	12	53

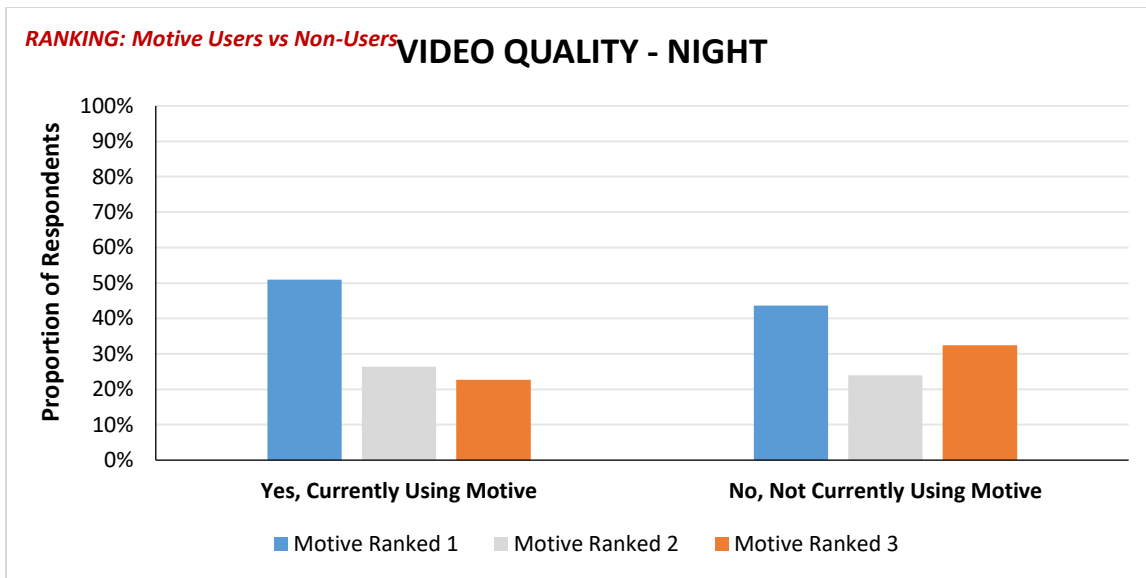


Figure 30 - Ranking for nighttime video quality comparing current Motive users and non-users.

Image Quality

The following section presents tables relevant to the analysis of image quality, for daytime and nighttime image stills, and by rating and overall order ranking questions.

Daytime Image Quality Ratings

Table 61 – Frequency counts for Likert scale ratings of daytime image quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	5	12	35	37	35	124
Lytix	4	8	21	57	35	125
Samsara	4	16	30	52	24	126

Table 62 - Model results for Likert scale ratings for daytime image quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.6762	0.1125	3.4558	3.8967	1068.01	<.0001
System Evaluated	Lytx	1	0.2023	0.1321	-0.0566	0.4612	2.35	0.1256
System Evaluated	Samsara	1	-0.0824	0.1318	-0.3408	0.1760	0.39	0.5319
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	0.0161	0.109	-0.1975	0.2298	0.02	0.8823
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	1.0422	0.0381	0.9702	1.1195		

Table 63 - Model results for Likert scale ratings of daytime image quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.4931	0.1509	10.6841	0.0011	0.611
System Evaluated	Lytx	1	0.4357	0.2096	4.3213	0.0376	1.546
System Evaluated	Samsara	1	0.0894	0.2115	0.1788	0.6724	1.094
Current Motive User	No	1	0.1146	0.0866	1.7529	0.1855	1.121

Table 64 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in daytime image quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytx vs Motive	0.6468	0.1356	0.05	0.4289	0.9754	4.3213	0.0376
Samsara vs Motive	0.9145	0.1934	0.05	0.6041	1.3842	0.1788	0.6724

Ratings of daytime image quality were further stratified by DMS experience, with assessment of Motive users' opinions of all three systems and a comparison of Motive users to non-Motive users. Table 65 and Table 66 present the frequency counts of daytime image quality ratings for the stratified assessment. Figure 31 illustrates the distribution of daytime image quality ratings for all three systems, as rated by current Motive users. Figure 32 illustrates the distribution of ratings for current Motive users and non-Motive users for daytime image quality.

Table 65 – Frequency counts for Likert scale ratings of daytime image quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	2	9	13	13	16	53
LytX	1	4	11	22	14	52
Samsara	1	5	15	20	12	53

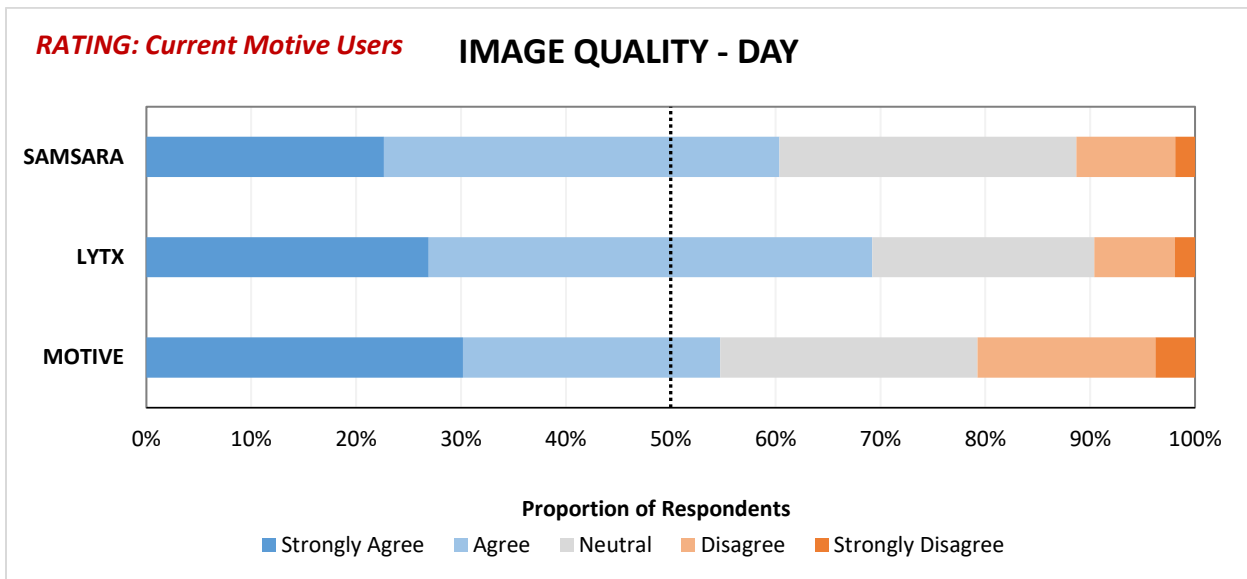


Figure 31 - Likert scale ratings for daytime image quality comparing current Motive users' opinions of all three systems.

Table 66 – Frequency counts for Likert scale ratings of daytime image quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	3	3	22	24	19	71
Yes	2	9	13	13	16	53

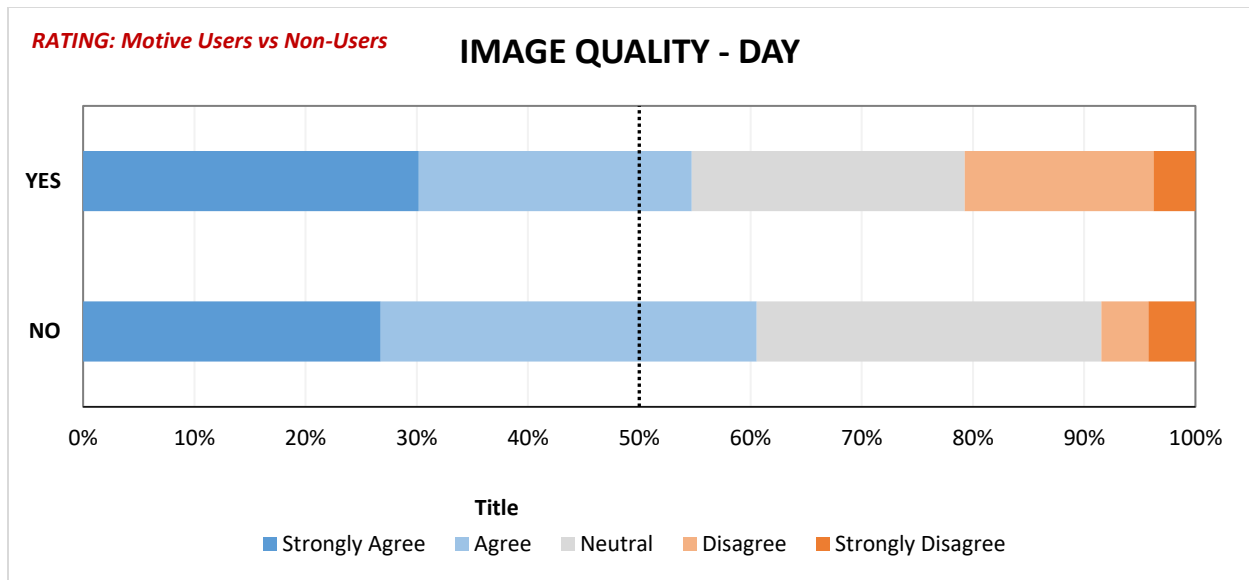


Figure 32 - Likert scale ratings for daytime image quality comparing current Motive users and non-users.

Nighttime Image Quality Ratings

Table 67 – Frequency counts for Likert scale ratings of nighttime image quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	7	22	26	35	23	113
LytX	9	26	31	32	19	117
Samsara	10	31	45	24	4	114

Table 68 - Model results for Likert scale ratings for nighttime image quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.4153	0.1259	3.1685	3.6621	735.64	<.0001
System Evaluated	LytX	1	-0.1758	0.1476	-0.4652	0.1136	1.42	0.2337
System Evaluated	Samsara	1	-0.5648	0.1486	-0.8560	-0.2735	14.45	0.0001
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	-0.0301	0.1219	-0.2691	0.2088	0.06	0.8047
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	1.1194	0.0427	1.0388	1.2063		

Table 69 - Model results for Likert scale ratings of nighttime image quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.8148	0.1587	26.3709	<.0001	0.443
System Evaluated	LytX	1	-0.1812	0.2278	0.6328	0.4263	0.834
System Evaluated	Samsara	1	-0.9364	0.2587	13.0991	0.0003	0.392
Current Motive User	No	1	0.0556	0.1007	0.3052	0.5806	1.057

Table 70 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in nighttime image quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	1.1986	0.2730	0.05	0.7670	1.8731	0.6328	0.4263
Samsara vs Motive	2.5507	0.6599	0.05	1.5362	4.2352	13.0991	0.0003

Ratings of nighttime image quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 71 and Table 72 present the frequency counts of nighttime image quality ratings for the stratified assessment. Figure 33 illustrates the distribution of nighttime image quality ratings

for all three systems, as rated by current Motive users. Figure 34 illustrates the distribution of ratings for current Motive users and non-Motive users for nighttime image quality.

Table 71 – Frequency counts for Likert scale ratings of nighttime image quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	2	11	14	14	8	49
Lytix	2	13	15	14	6	50
Samsara	2	13	19	12	3	49

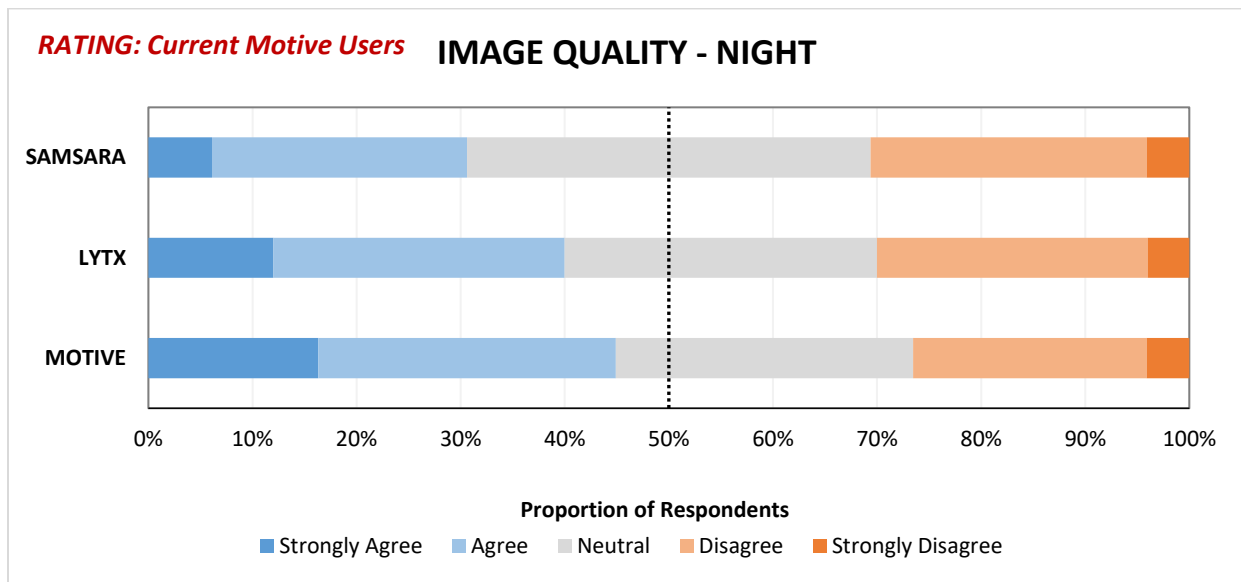


Figure 33 - Likert scale ratings for nighttime image quality comparing current Motive users’ opinions of all three systems.

Table 72 – Frequency counts for Likert scale ratings of nighttime image quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	5	11	12	21	15	64
Yes	2	11	14	14	8	49

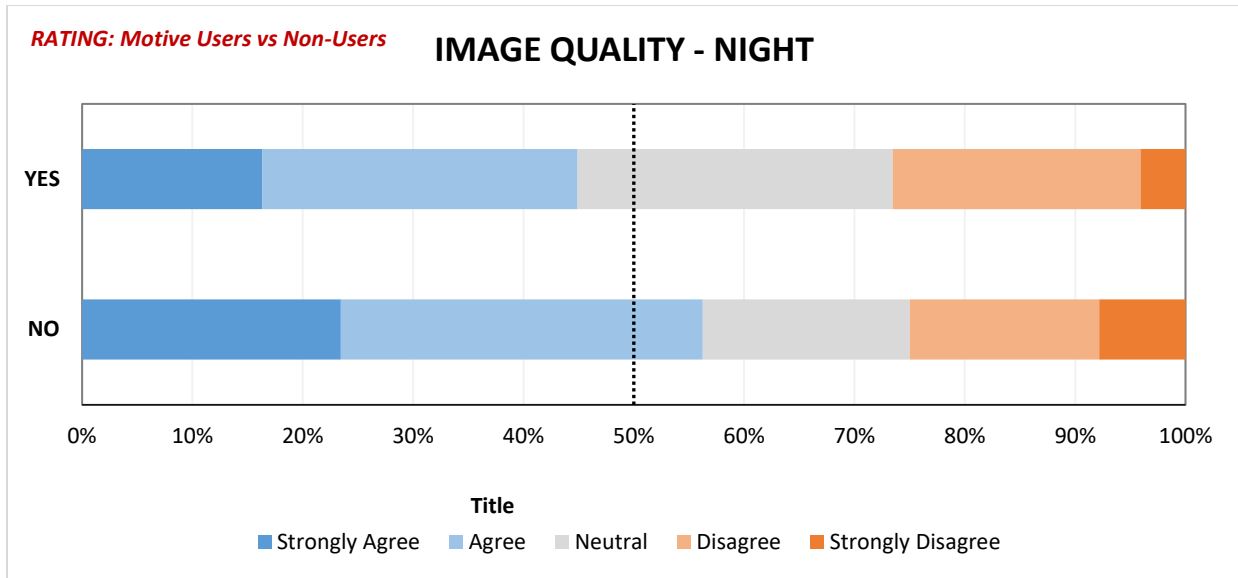


Figure 34 - Likert scale ratings for nighttime image quality comparing current Motive users and non-users.

Daytime Image Quality Overall Order Rankings

Table 73 – Frequency counts for overall order ranking of daytime image quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	37	31	49	117
Lytx	43	55	20	118
Samsara	38	34	46	118

Table 74 - Model results for overall order ranking of daytime image quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.2573	0.1495	2.9592	0.0854	1.293
System Evaluated	Lytx	1	0.1203	0.2115	0.3234	0.5696	1.128
System Evaluated	Samsara	1	0.0091	0.2106	0.0019	0.9655	1.009
Current Motive User	No	1	-0.0328	0.0869	0.1429	0.7055	0.968

Table 75 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in daytime image quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	0.8867	0.1875	0.05	0.5858	1.3421	0.3234	0.5696
Samsara vs Motive	0.9909	0.2087	0.05	0.6559	1.4972	0.0019	0.9655

Rankings of daytime image quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 76 and Table 77 present the frequency counts of daytime image quality rankings for the stratified assessment. Figure 35 illustrates the distribution of daytime image quality ranking for all three systems, as ranked by current Motive users. Figure 36 illustrates the distribution of ranking for current Motive users and non-Motive users for daytime image quality.

Table 76 – Frequency counts for overall order ranking of daytime image quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	19	10	20	49
LytX	15	26	11	52
Samsara	17	16	18	51

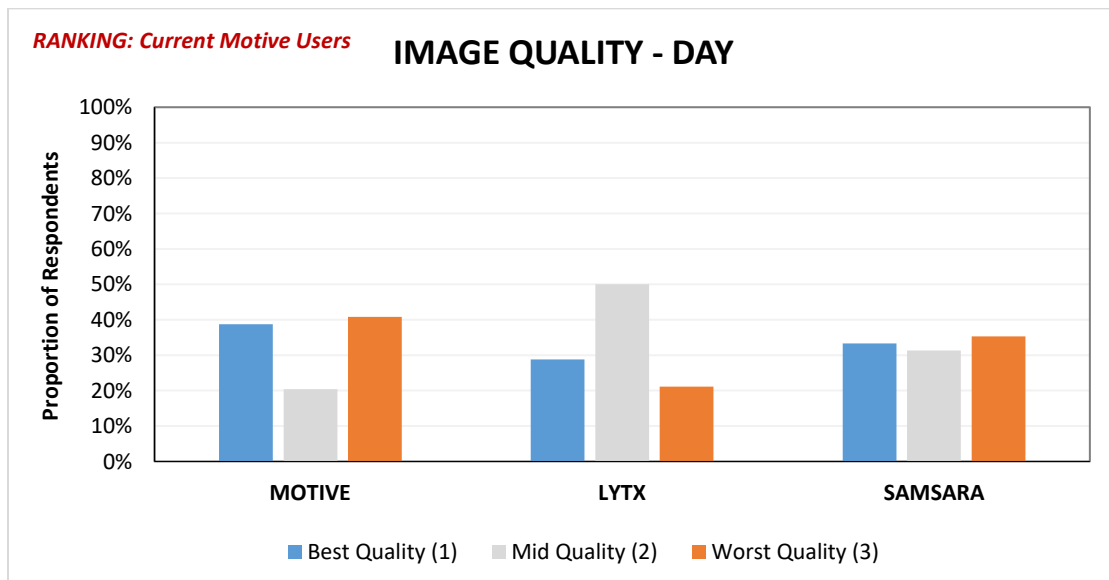


Figure 35 - Ranking for daytime image quality comparing current Motive users’ opinions of all three systems.

Table 77 – Frequency counts for ranking of daytime image quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	18	21	29	68
Yes	19	10	20	49

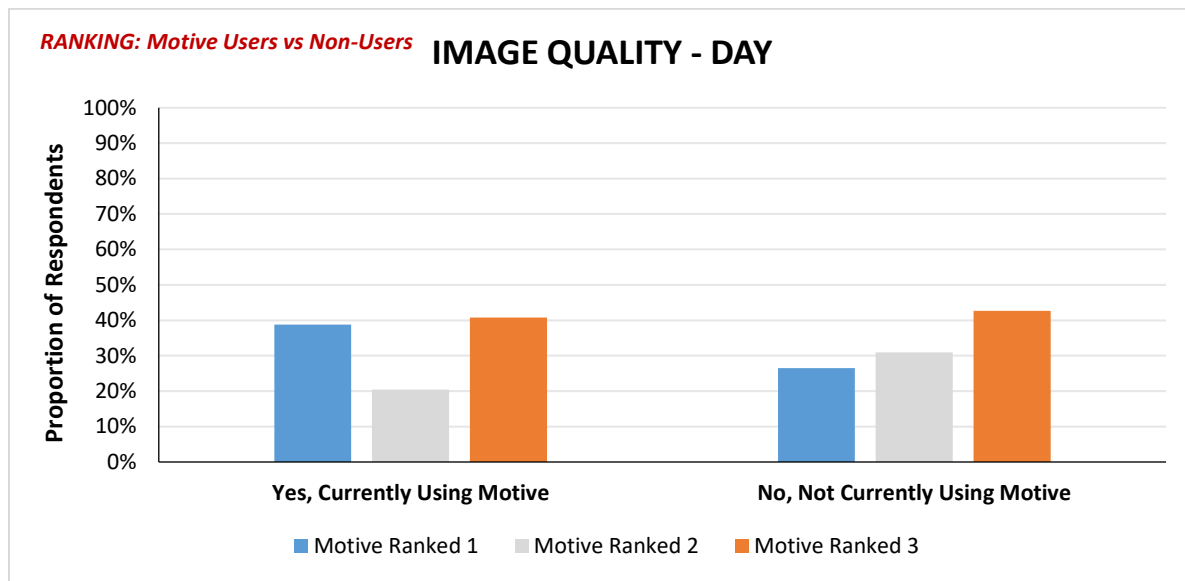


Figure 36 - Ranking for daytime image quality comparing current Motive users and non-users.

Nighttime Video Image Overall Order Rankings

Table 78 – Frequency counts for overall order ranking of nighttime image quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	52	30	29	111
Lytix	33	44	31	108
Samsara	26	37	47	110

Table 79 - Model results for overall order ranking of nighttime image quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.7521	0.1583	22.5642	<.0001	2.121
System Evaluated	Lytx	1	-0.4050	0.2184	3.4367	0.0638	0.667
System Evaluated	Samsara	1	-0.5865	0.2167	7.3258	0.0068	0.556
Current Motive User	No	1	-0.0112	0.0885	0.0161	0.8991	0.989

Table 80 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in nighttime image quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytx vs Motive	1.4993	0.3275	0.05	0.9771	2.3005	3.4367	0.0638
Samsara vs Motive	1.7977	0.3896	0.05	1.1756	2.7490	7.3258	0.0068

Rankings of nighttime image quality were further stratified by DMS experience, with assessment of Motive users' opinions of all three systems and a comparison of Motive users to non-Motive users. Table 81 and Table 82 present the frequency counts of nighttime image quality rankings for the stratified assessment. Figure 37 illustrates the distribution of nighttime image quality ranking for all three systems, as ranked by current Motive users. Figure 38 illustrates the distribution of ranking for current Motive users and non-Motive users for daytime image quality.

Table 81 – Frequency counts for overall order ranking of nighttime image quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	21	12	15	48
Lytx	18	16	14	48
Samsara	9	21	18	48

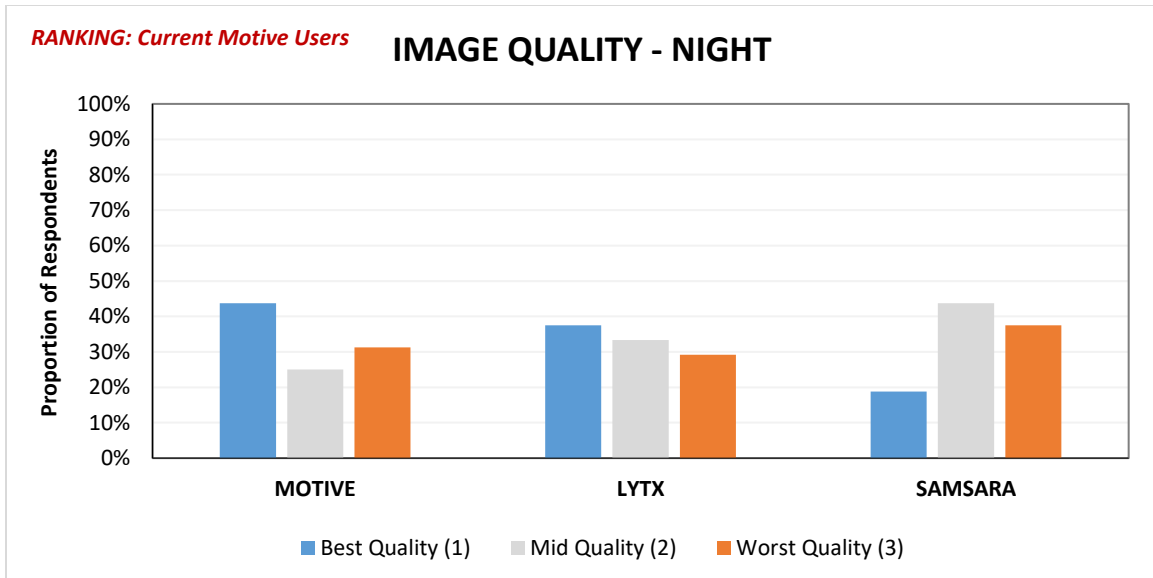


Figure 37 - Ranking for nighttime image quality comparing current Motive users' opinions of all three systems.

Table 82 – Frequency counts for ranking of nighttime image quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	31	18	14	63
Yes	21	12	15	48

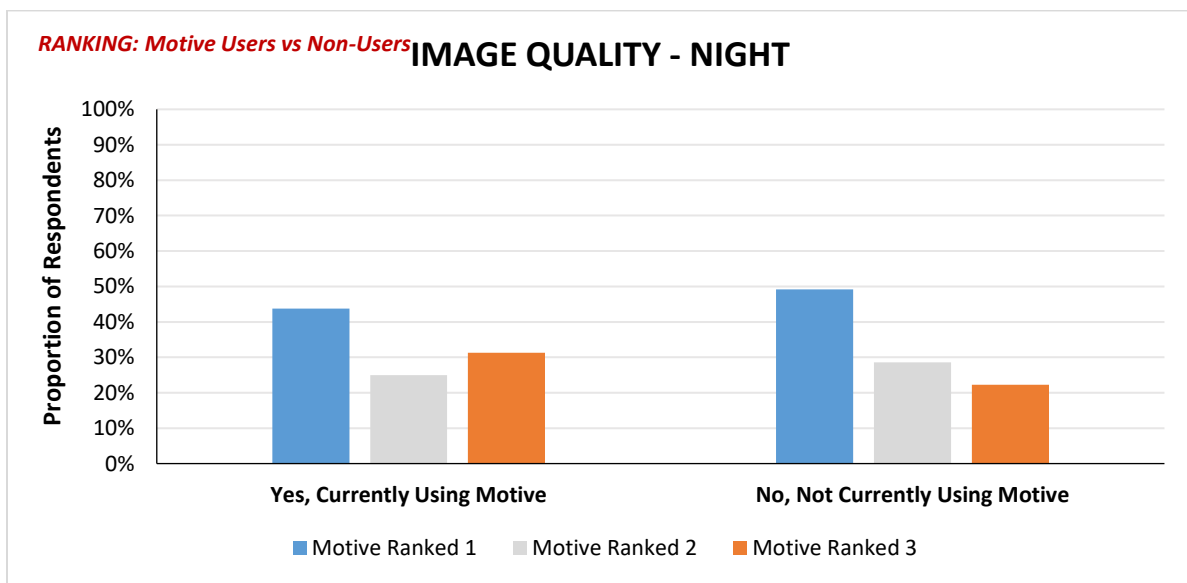


Figure 38 - Ranking for nighttime image quality comparing current Motive users and non-users.

Alert Quality

The following section presents tables relevant to the analysis of alert quality, for cell phone use, seatbelt use, and close following distance alert types and by rating and overall order ranking questions.

Cell Phone Use Alert Quality Ratings

Table 83 – Frequency counts for Likert scale ratings of cell phone use alert quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	4	9	17	46	35	111
Lytx	5	10	20	39	37	111
Samsara	6	8	24	41	32	111

Table 84 - Model results for Likert scale ratings for cell phone use alert quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	4.0192	0.1230	3.7781	4.2604	1067.49	<.0001
System Evaluated	Lytx	1	-0.0541	0.1459	-0.3400	0.2319	0.14	0.7110
System Evaluated	Samsara	1	-0.1261	0.1459	-0.4121	0.1598	0.75	0.3873
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	-0.2280	0.1200	-0.4631	0.0071	3.61	0.0573
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	1.0869	0.0421	1.0075	1.1727		

Table 85 - Model results for Likert scale ratings of cell phone use alert quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.2629	0.1479	3.1611	0.0754	0.769
System Evaluated	Lytx	1	-0.1098	0.2097	0.2744	0.6004	0.896
System Evaluated	Samsara	1	-0.1768	0.2104	0.7062	0.4007	0.838
Current Motive User	No	1	-0.1296	0.0864	2.2490	0.1337	0.878

Table 86 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in cell phone use alert quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytx vs Motive	1.1161	0.2340	0.05	0.7400	1.6833	0.2744	0.6004
Samsara vs Motive	1.1934	0.2511	0.05	0.7901	1.8025	0.7062	0.4007

Ratings of cell phone use alert quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 87 and Table 88 present the frequency counts of cell phone use alert quality ratings for the stratified assessment.

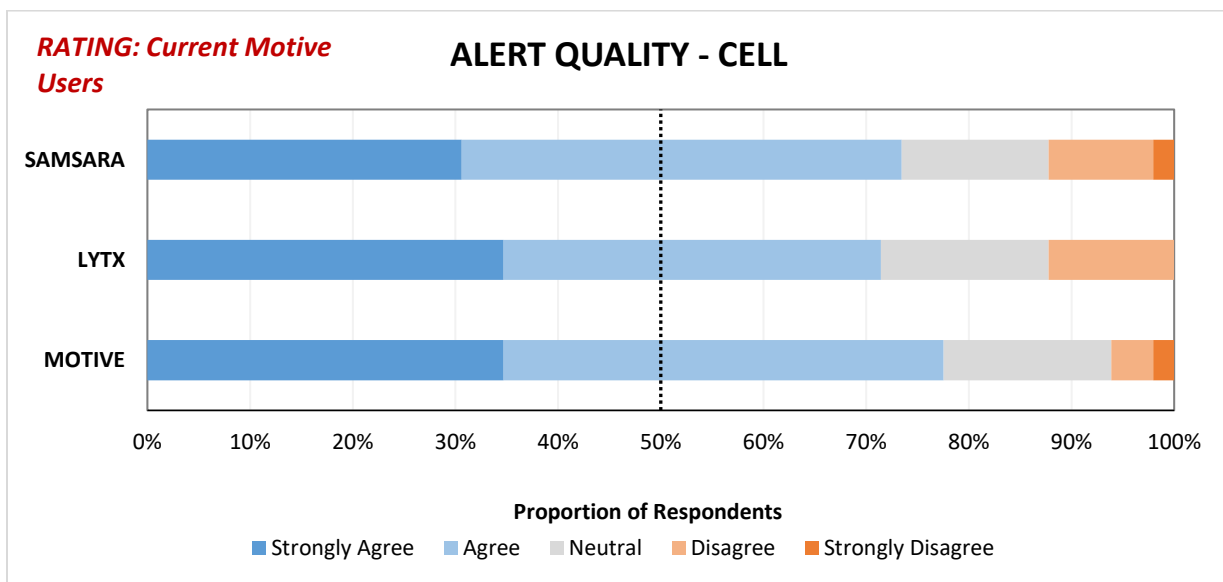


Figure 39 illustrates the distribution of cell phone use alert quality ratings for all three systems, as rated by current Motive users. Figure 40 illustrates the distribution of ratings for current Motive users and non-Motive users for cell phone use alert quality.

Table 87 – Frequency counts for Likert scale ratings of cell phone use alert quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	1	2	8	21	17	49
Lytx	0	6	8	18	17	49
Samsara	1	5	7	21	15	49

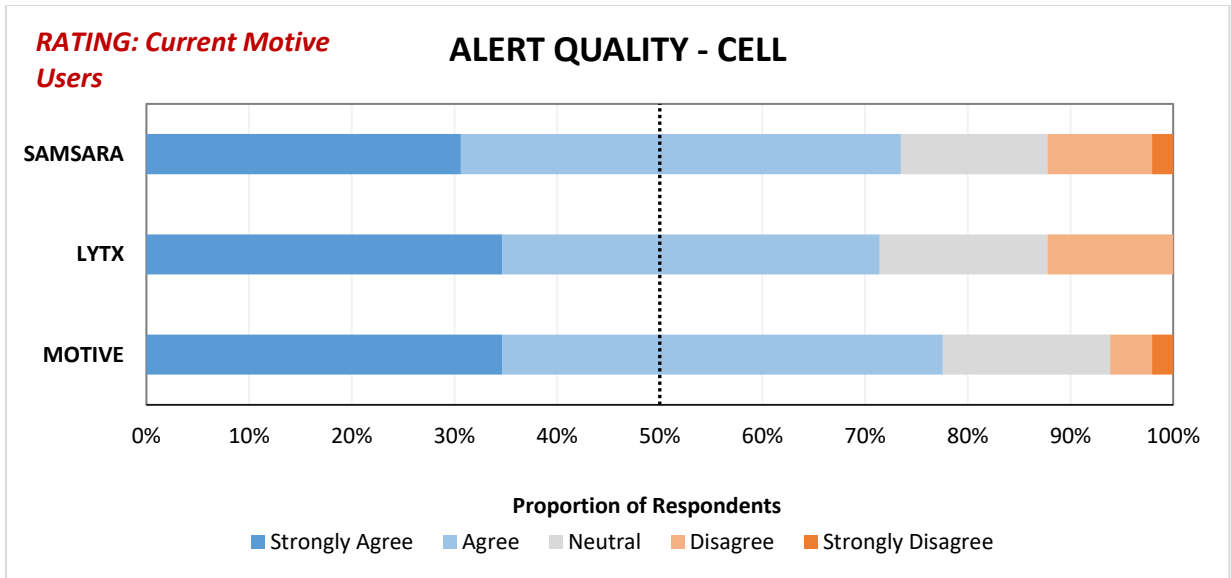


Figure 39 - Likert scale ratings for cell phone use alert quality comparing current Motive users' opinions of all three systems.

Table 88 – Frequency counts for Likert scale ratings of cell phone use alert quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	3	7	9	25	18	62
Yes	1	2	8	21	17	49

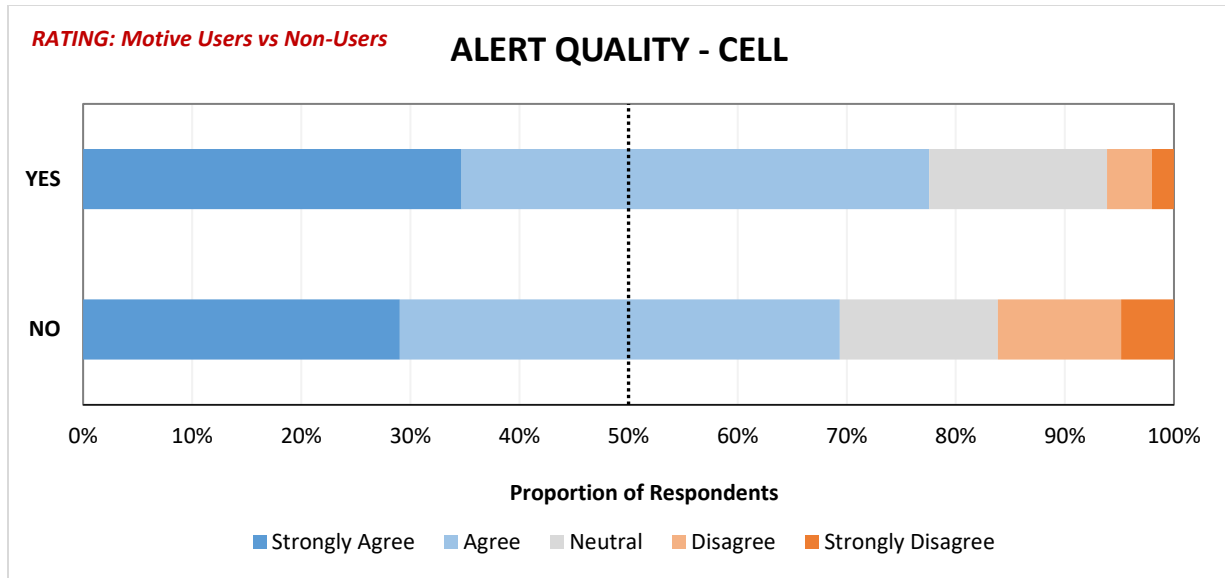


Figure 40 – Likert scale ratings for cell phone use alert quality comparing current Motive users and non-users.

Seatbelt Use Alert Quality Ratings

Table 89 – Frequency counts for Likert scale ratings of seatbelt use alert quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	3	17	31	33	21	105
Lytix	4	8	19	55	20	106
Samsara	4	5	12	49	36	106

Table 90 - Model results for Likert scale ratings for seatbelt use alert quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.4406	0.1184	3.2086	3.6726	845.07	<.0001
System Evaluated	LytX	1	0.2506	0.1391	-0.0221	0.5232	3.24	0.0717
System Evaluated	Samsara	1	0.5241	0.1391	0.2515	0.7968	14.19	0.0002
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	0.0956	0.1146	-0.1289	0.3201	0.70	0.404
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	1.0104	0.0401	0.9347	1.0922		

Table 91 - Model results for Likert scale ratings of seatbelt use alert quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.9235	0.1620	32.4785	<.0001	0.397
System Evaluated	LytX	1	0.4999	0.2197	5.1768	0.0229	1.649
System Evaluated	Samsara	1	0.7182	0.2181	10.8452	0.0010	2.051
Current Motive User	No	1	0.0994	0.0888	1.2544	0.2627	1.105

Table 92 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in seatbelt use alert quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	0.6066	0.1333	0.05	0.3944	0.9331	5.1768	0.0229
Samsara vs Motive	0.4876	0.1063	0.05	0.3180	0.7477	10.8452	0.0010

Ratings of seatbelt use alert quality were further stratified by DMS experience, with assessment of Motive users' opinions of all three systems and a comparison of Motive users to non-Motive users. Table 93 and Table 94 present the frequency counts of seatbelt use alert quality ratings for the stratified assessment. Figure 41 illustrates the distribution of seatbelt use alert quality ratings for all three systems, as rated by current Motive users. Figure 42 illustrates the distribution of ratings for current Motive users and non-Motive users for seatbelt use alert quality.

Table 93 – Frequency counts for Likert scale ratings of seatbelt use alert quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	1	8	14	13	9	45
Lytix	1	4	9	22	10	46
Samsara	3	2	8	19	14	46

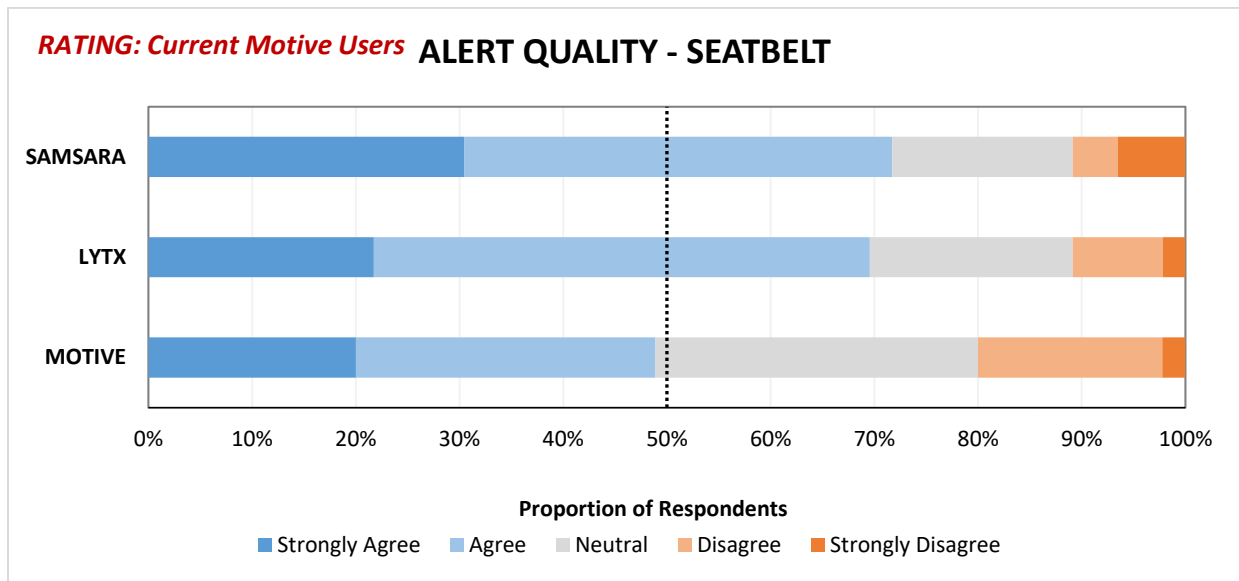


Figure 41 - Likert scale ratings for seatbelt use alert quality comparing current Motive users' opinions of all three systems.

Table 94 – Frequency counts for Likert scale ratings of seatbelt use alert quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	2	9	17	20	12	60
Yes	1	8	14	13	9	45

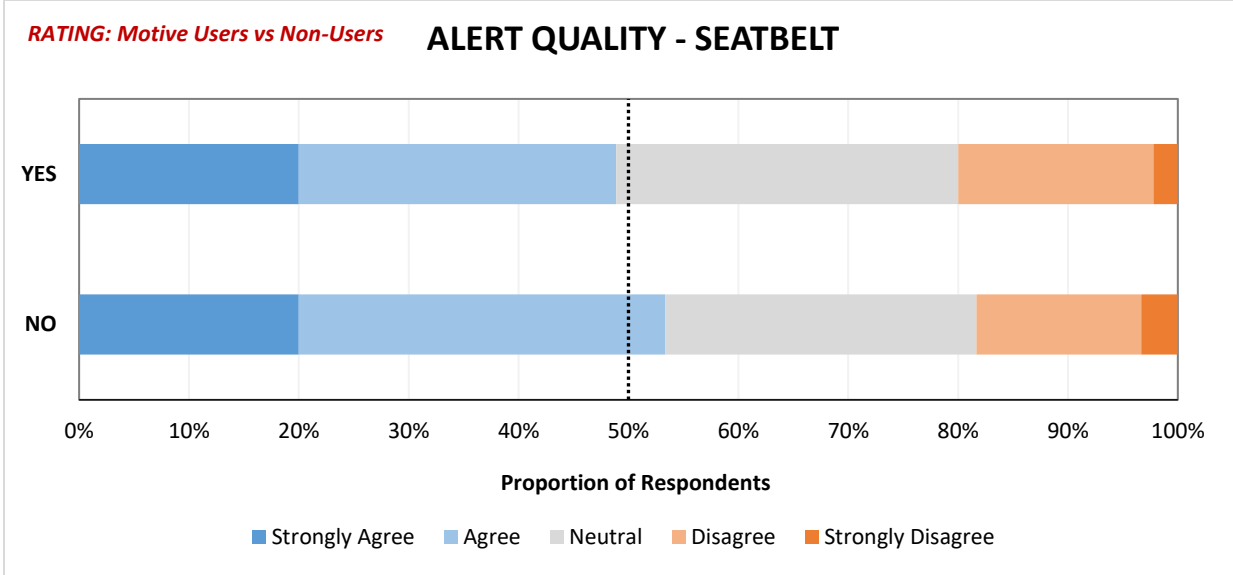


Figure 42 – Likert scale ratings for seatbelt use alert quality comparing current Motive users and non-users.

Close Following Distance Alert Quality Ratings

Table 95 – Frequency counts for Likert scale ratings of close following distance alert quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	6	11	29	37	20	103
Lytx	2	6	29	46	21	104
Samsara	5	12	32	40	15	104

Table 96 - Model results for Likert scale ratings for close following distance alert quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.4546	0.1184	3.2226	3.6866	851.71	<.0001
System Evaluated	LytX	1	0.2240	0.1405	-0.0515	0.4994	2.54	0.1110
System Evaluated	Samsara	1	-0.0645	0.1405	-0.3399	0.2110	0.21	0.6463
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	0.1259	0.1156	-0.1006	0.3524	1.19	0.2759
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	1.0109	0.0405	0.9345	1.0936		

Table 97 - Model results for Likert scale ratings of close following distance alert quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.8323	0.1591	27.3649	<.0001	0.435
System Evaluated	LytX	1	0.2411	0.2199	1.2014	0.2731	1.273
System Evaluated	Samsara	1	-0.0509	0.2256	0.0509	0.8216	0.950
Current Motive User	No	1	0.00124	0.0914	0.0002	0.9891	1.001

Table 98 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in close following distance alert quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	0.7858	0.1728	0.05	0.5106	1.2092	1.2014	0.2731
Samsara vs Motive	1.0522	0.2373	0.05	0.6762	1.6372	0.0509	0.8216

Ratings of close following distance alert quality were further stratified by DMS experience, with assessment of Motive users' opinions of all three systems and a comparison of Motive users to non-Motive users. Table 99 and Table 100 present the frequency counts of close following distance alert quality ratings for the stratified assessment.

Figure 43 illustrates the distribution of close following distance alert quality ratings for all three systems, as rated by current Motive users. Figure 44 illustrates the distribution of ratings for current Motive users and non-Motive users for close following distance alert quality.

Table 99 – Frequency counts for Likert scale ratings of close following distance alert quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	3	8	6	19	10	46
Lytx	1	5	13	20	6	45
Samsara	2	6	14	17	6	45

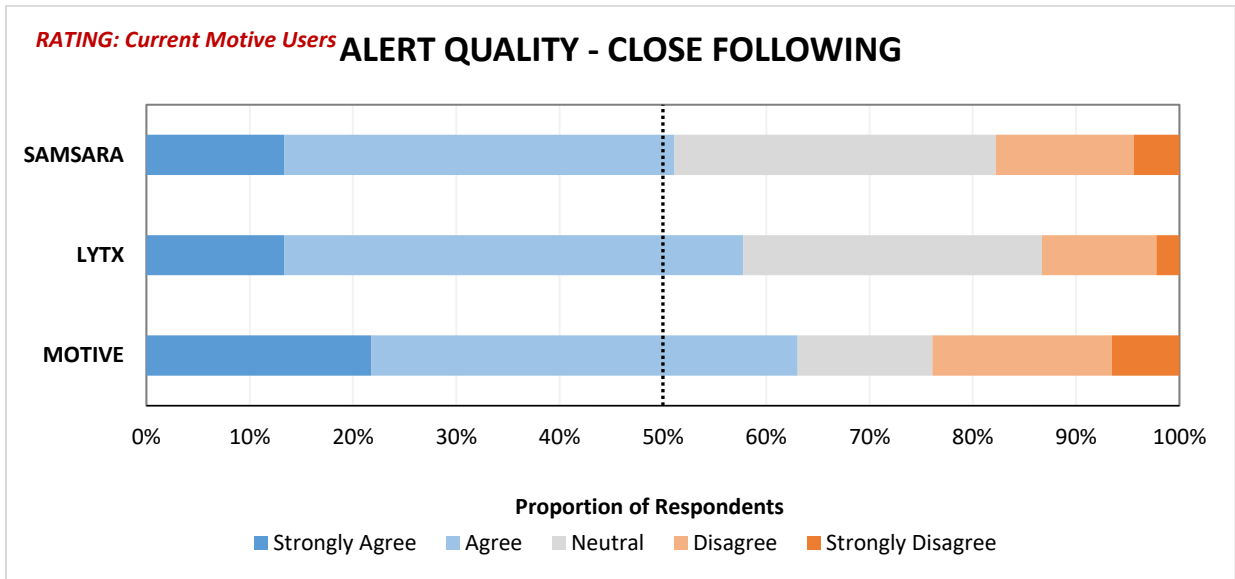


Figure 43 - Likert scale ratings for close following distance alert quality comparing current Motive users' opinions of all three systems.

Table 100 – Frequency counts for Likert scale ratings of close following distance alert quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	3	3	23	18	10	57
Yes	3	8	6	19	10	46

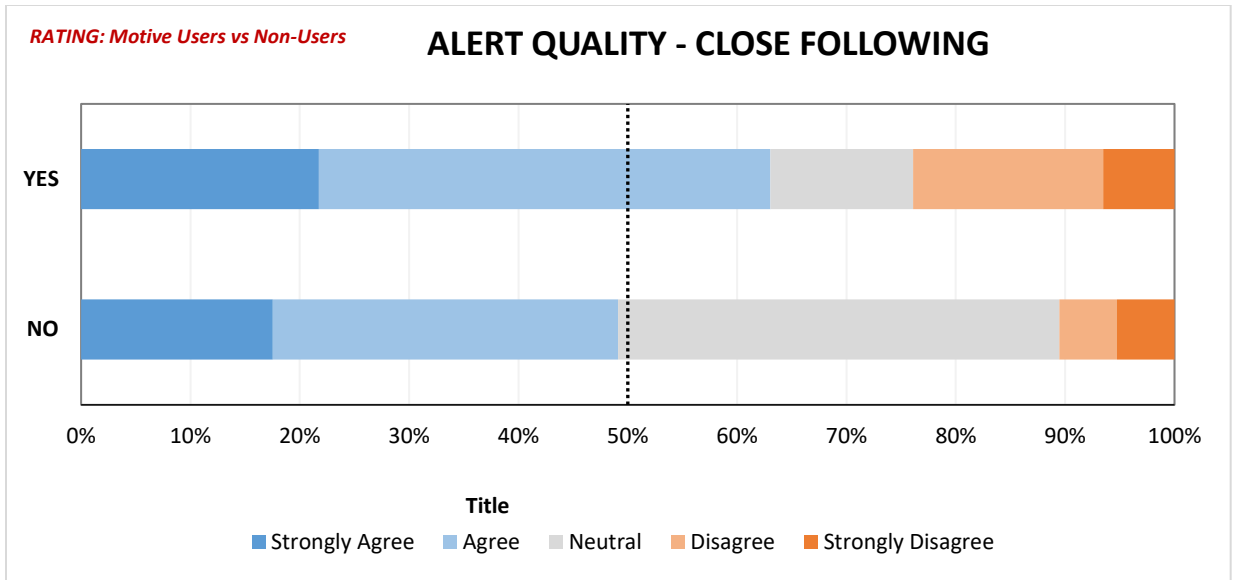


Figure 44 – Likert scale ratings for close following distance alert quality comparing current Motive users and non-users.

Cell Phone Use Alert Quality Overall Order Rankings

Table 101 – Frequency counts for overall order ranking of cell phone use alert quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	48	22	37	107
LytX	22	52	32	106
Samsara	37	32	37	106

Table 102 - Model results for overall order ranking of cell phone use alert quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.7642	0.1579	23.414	<.0001	2.147
System Evaluated	LytX	1	-0.5821	0.2159	7.2654	0.0070	0.559
System Evaluated	Samsara	1	-0.2473	0.2190	1.2755	0.2587	0.781
Current Motive User	No	1	0.0194	0.0886	0.0480	0.8266	1.020

Table 103 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in cell phone use alert quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	1.7898	0.3865	0.05	1.1721	2.7328	7.2654	0.0070
Samsara vs Motive	1.2805	0.2804	0.05	0.8337	1.9668	1.2755	0.2587

Rankings of cell phone use alert quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 104 and Table 105 present the frequency counts of cell phone use alert quality rankings for the stratified assessment. Figure 45 illustrates the distribution of cell phone use alert quality ranking for all three systems, as ranked by current Motive users. Figure 46 illustrates the distribution of ranking for current Motive users and non-Motive users for cell phone use alert quality.

Table 104 – Frequency counts for overall order ranking of cell phone use alert quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	22	9	16	47
LytX	9	25	14	48
Samsara	16	14	16	46

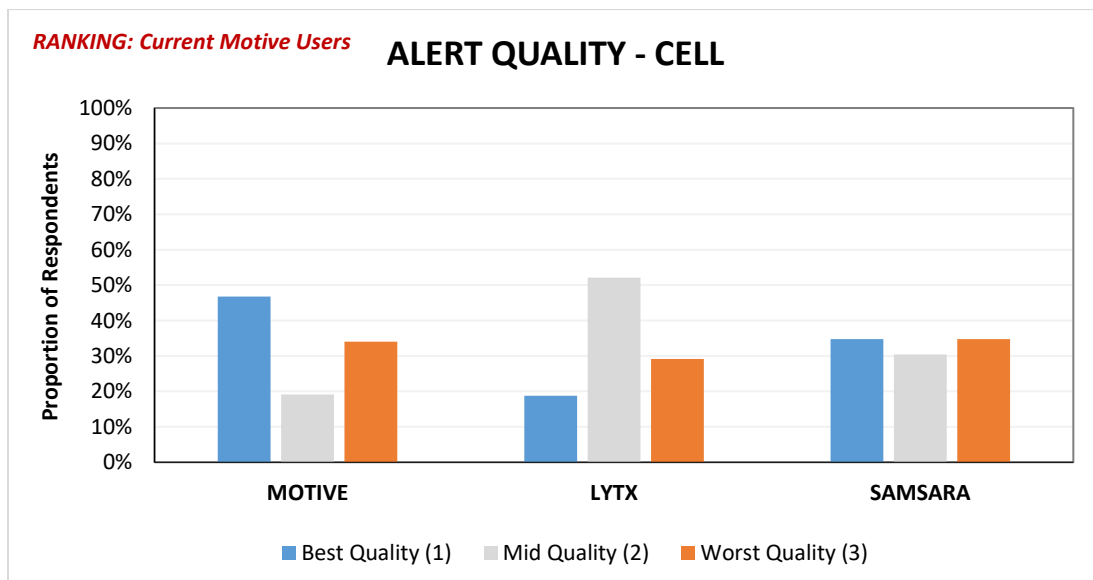


Figure 45 - Ranking for cell phone use alert quality comparing current Motive users’ opinions of all three systems.

Table 105 – Frequency counts for ranking of cell phone use alert quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	26	13	21	60
Yes	22	9	16	47

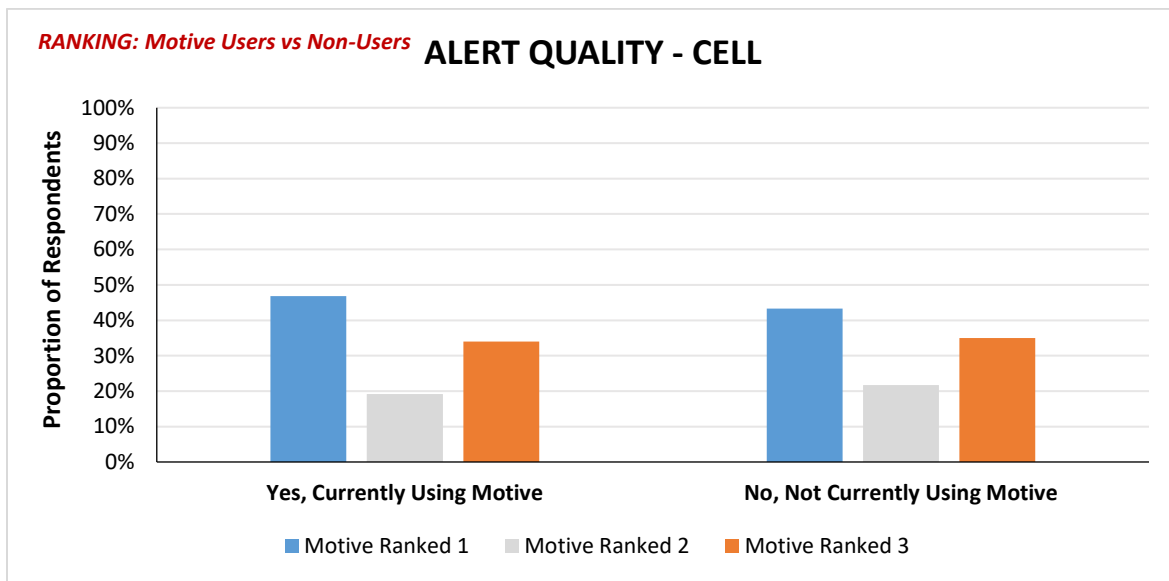


Figure 46 - Ranking for cell phone use alert quality comparing current Motive users and non-users.

Seatbelt Use Alert Quality Overall Order Rankings

Table 106 – Frequency counts for overall order ranking of seatbelt use alert quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	34	26	42	102
Lytix	21	50	31	102
Samsara	46	27	27	100

Table 107 - Model results for overall order ranking of seatbelt use alert quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.5478	0.1527	12.8631	0.0003	1.729
System Evaluated	Lytx	1	-0.2912	0.2123	1.8816	0.1702	0.747
System Evaluated	Samsara	1	0.3276	0.2223	2.1725	0.1405	1.388
Current Motive User	No	1	0.0250	0.0894	0.0784	0.7795	1.025

Table 108 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in seatbelt use alert quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytx vs Motive	1.3380	0.2840	0.05	0.8826	2.0283	1.8816	0.1702
Samsara vs Motive	0.7207	0.1602	0.05	0.4662	1.1141	2.1725	0.1405

Rankings of seatbelt use alert quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 109 and Table 110 present the frequency counts of seatbelt use alert quality rankings for the stratified assessment. Figure 47 illustrates the distribution of seatbelt use alert quality ranking for all three systems, as ranked by current Motive users. Figure 48 illustrates the distribution of ranking for current Motive users and non-Motive users for seatbelt use alert quality.

Table 109 – Frequency counts for overall order ranking of seatbelt use alert quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	13	8	23	44
Lytx	8	27	11	46
Samsara	23	11	10	44

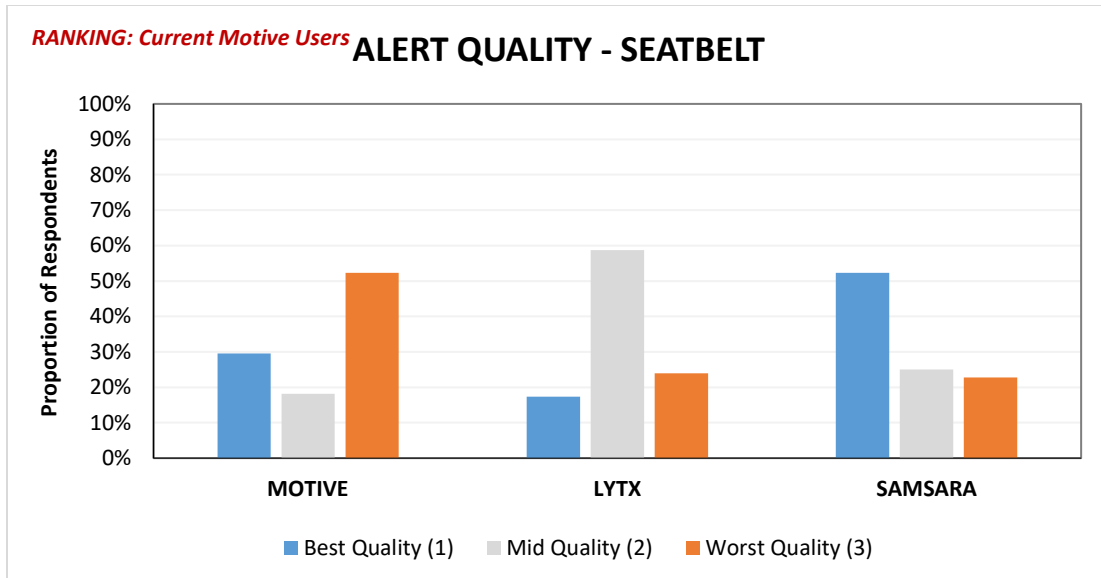


Figure 47 - Ranking for seatbelt use alert quality comparing current Motive users' opinions of all three systems.

Table 110 – Frequency counts for ranking of seatbelt use alert quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	21	18	19	58
Yes	13	8	23	44

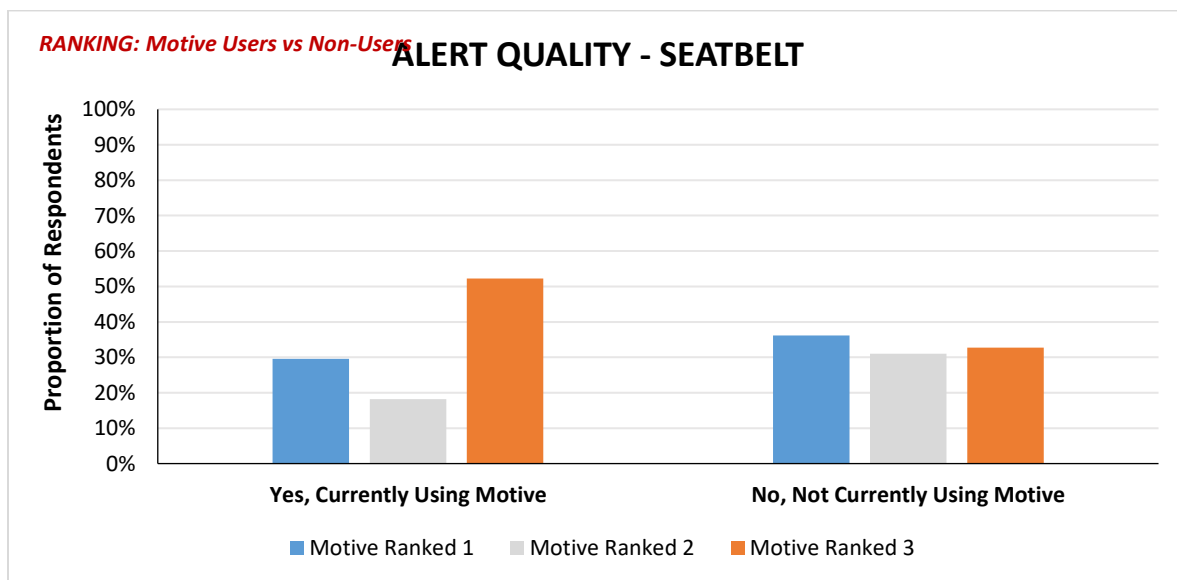


Figure 48 - Ranking for seatbelt use alert quality comparing current Motive users and non-users.

Close Following Distance Alert Quality Overall Order Rankings

Table 111 – Frequency counts for overall order ranking of close following distance alert quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	47	24	28	99
Lytix	29	49	21	99
Samsara	23	25	49	97

Table 112 - Model results for overall order ranking of close following distance alert quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.9431	0.1637	33.1807	<.0001	2.568
System Evaluated	Lytix	1	-0.4415	0.2227	3.9292	0.0475	0.643
System Evaluated	Samsara	1	-0.5502	0.2220	6.1405	0.0132	0.577
Current Motive User	No	1	0.0275	0.0900	0.0934	0.7598	1.028

Table 113 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in close following distance alert quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytix vs Motive	1.5550	0.3463	0.05	1.0050	2.4062	3.9292	0.0475
Samsara vs Motive	1.7337	0.3850	0.05	1.1219	2.6790	6.1405	0.0132

Rankings of close following distance alert quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 114 and Table 115 present the frequency counts of close following distance alert quality rankings for the stratified assessment. Figure 49 illustrates the distribution of close following distance alert quality ranking for all three systems, as ranked by current Motive users. Figure 50 illustrates the distribution of ranking for current Motive users and non-Motive users for close following distance alert quality.

Table 114 – Frequency counts for overall order ranking of close following distance alert quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	18	16	10	44
Lytix	17	20	7	44
Samsara	9	7	27	43

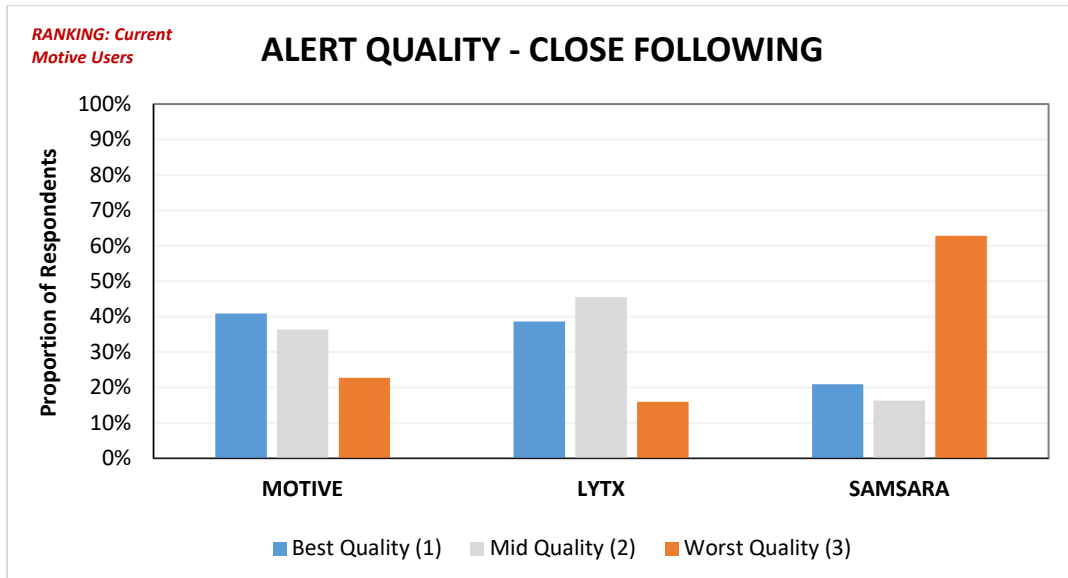


Figure 49 - Ranking for close following distance alert quality comparing current Motive users' opinions of all three systems.

Table 115 – Frequency counts for ranking of close following distance alert quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	29	8	18	55
Yes	18	16	10	44

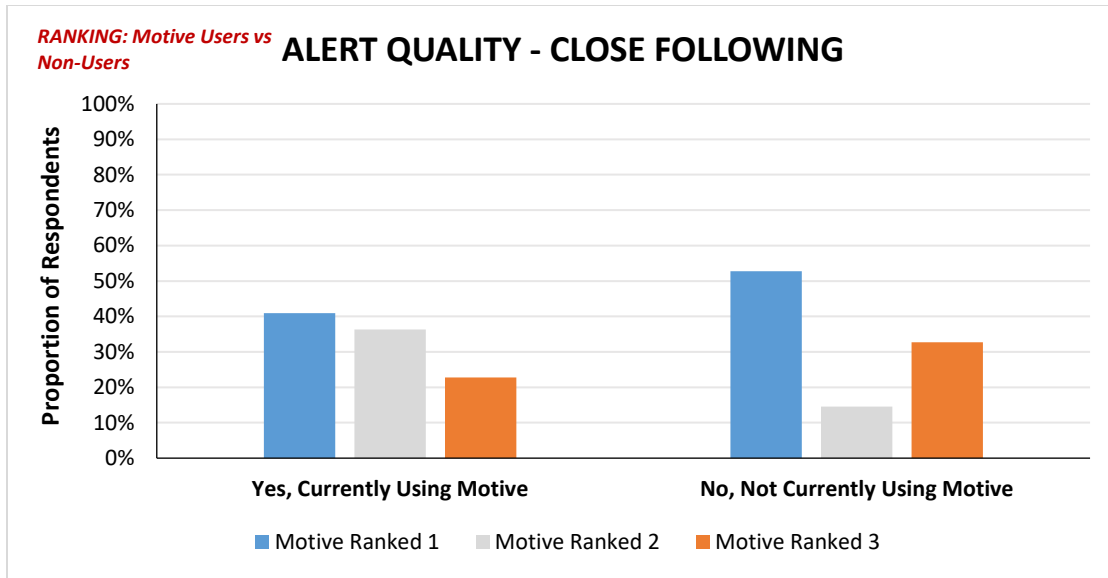


Figure 50 - Ranking for close following distance alert quality comparing current Motive users and non-users.

Hardware Design Quality

The following section presents tables relevant to the analysis of hardware design quality, by rating and overall order ranking questions.

Hardware Design Quality Ratings

Table 116 – Frequency counts for Likert scale ratings of hardware design quality.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	3	7	21	50	20	101
LytX	3	4	21	49	24	101
Samsara	8	8	32	40	13	101

Table 117 - Model results for Likert scale ratings for hardware design quality assessed as continuous variables, with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald 95% Lower Confidence Limit	Wald 95% Upper Confidence Limit	Wald Chi-Square	p-value
Intercept		1	3.8751	0.1153	3.6490	4.1012	1128.61	<.0001
System Evaluated	Lytx	1	0.0990	0.1371	-0.1698	0.3678	0.52	0.4703
System Evaluated	Samsara	1	-0.3465	0.1371	-0.6153	-0.0777	6.38	0.0115
System Evaluated	Motive	0	0	0	0	0	.	.
Current Motive User	No	1	-0.2033	0.1126	-0.4241	0.0175	3.26	0.0711
Current Motive User	Yes	0	0	0	0	0	.	.
Scale		1	0.9746	0.0396	0.9000	1.0554		

Table 118 - Model results for Likert scale ratings of hardware design quality assessed as binary values (4 or higher vs. 3 or lower), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	-0.5056	0.1515	11.1400	0.0008	0.603
System Evaluated	Lytx	1	0.0681	0.2131	0.1021	0.7493	1.070
System Evaluated	Samsara	1	-0.4148	0.2220	3.4909	0.0617	0.660
Current Motive User	No	1	-0.1522	0.0894	2.8965	0.0888	0.859

Table 119 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ratings of 4 or higher vs. 3 or lower in hardware design quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
Lytx vs Motive	0.9342	0.199	0.05	0.6153	1.4184	0.1021	0.7493
Samsara vs Motive	1.5141	0.3361	0.05	0.9799	2.3394	3.4909	0.0617

Ratings of hardware design quality were further stratified by DMS experience, with assessment of Motive users' opinions of all three systems and a comparison of Motive users to non-Motive users. Table 120 and Table 121 present the frequency counts of hardware design quality ratings for the stratified assessment. Figure 51 illustrates the distribution of hardware design quality ratings for all three systems, as rated by current Motive users. Figure 52 illustrates the

distribution of ratings for current Motive users and non-Motive users for hardware design quality.

Table 120 – Frequency counts for Likert scale ratings of hardware design quality for current Motive users.

System	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
Motive	2	3	7	24	9	45
Lytix	1	0	9	22	13	45
Samsara	3	3	12	19	8	45

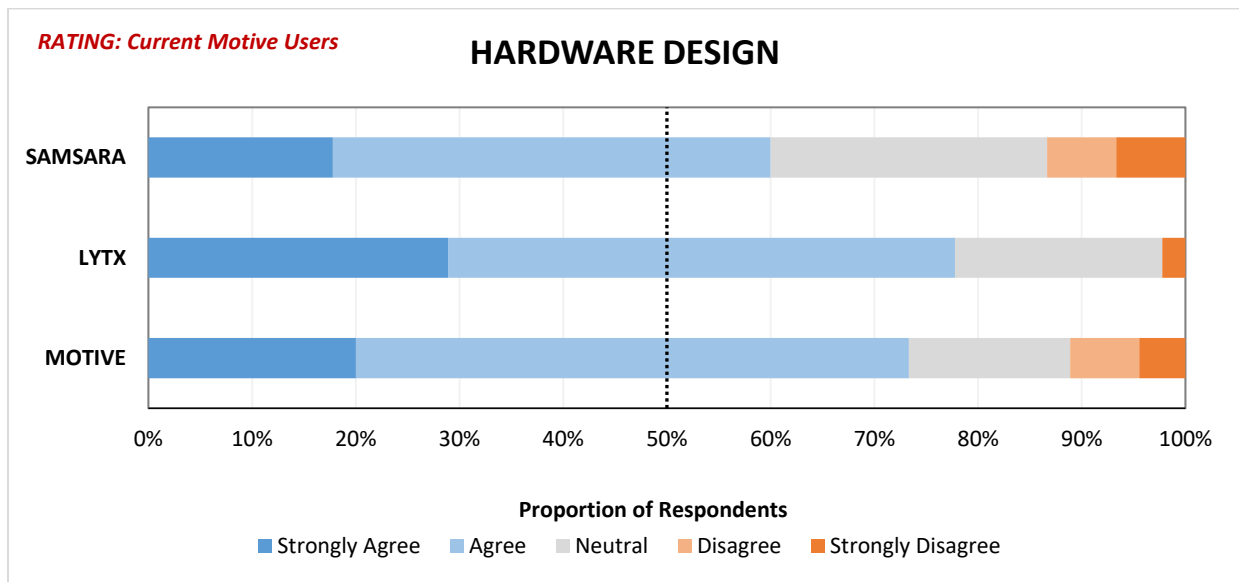


Figure 51 – Likert scale ratings for hardware design quality comparing current Motive users’ opinions of all three systems.

Table 121 – Frequency counts for Likert scale ratings of hardware design quality for current Motive users compared to non-users.

Motive User	Counts of Strongly Disagree (1)	Counts of Disagree (2)	Counts of Neutral (3)	Counts of Agree (4)	Counts of Strongly Agree (5)	Total Count
No	1	4	14	26	11	56
Yes	2	3	7	24	9	45

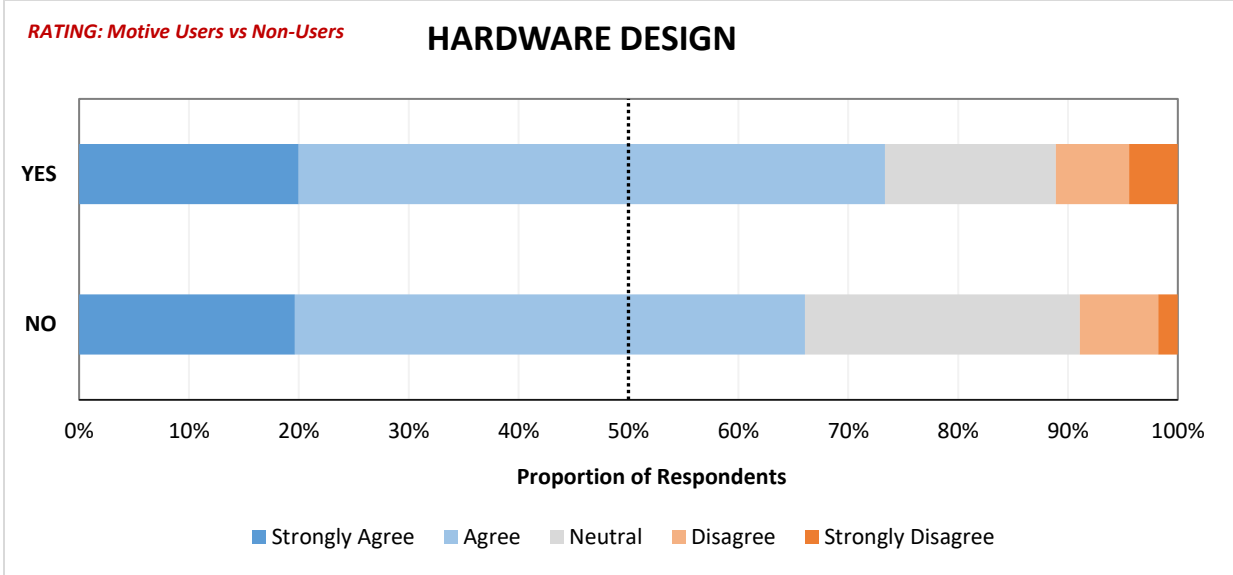


Figure 52 - Likert scale ratings for hardware design quality comparing current Motive users and non-users.

Hardware Design Quality Overall Order Rankings

Table 122 – Frequency counts for overall order ranking of daytime image quality.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	38	48	12	98
LytX	39	39	19	97
Samsara	21	11	64	96

Table 123 - Model results for overall order ranking of hardware design quality assessed as binary values (ranking of 1 vs. ranking of 2 or 3), with parameter estimate values.

Parameter	Parameter Value	df	Estimate	Standard Error	Wald Chi-Square	p-value	Exp(Est)
Intercept		1	0.7398	0.1572	22.1529	<.0001	2.096
System Evaluated	LytX	1	0.0417	0.2230	0.0350	0.8517	1.043
System Evaluated	Samsara	1	-0.3681	0.2171	2.8765	0.0899	0.692
Current Motive User	No	1	0.0185	0.0901	0.0420	0.8375	1.019

Table 124 – Odds ratio calculations comparing Motive to Lytx and Samsara systems for ranking of 1 vs. 2 or 3 in hardware design quality.

Contrast	Odds Ratio Estimate	Standard Error	Alpha	95% Lower Confidence Limit	95% Upper Confidence Limit	Wald Chi-Square	p-value
LytX vs Motive	0.9592	0.2139	0.05	0.6196	1.4849	0.0350	0.8517
Samsara vs Motive	1.4450	0.3137	0.05	0.9443	2.2113	2.8765	0.0899

Rankings of hardware design quality were further stratified by DMS experience, with assessment of Motive users’ opinions of all three systems and a comparison of Motive users to non-Motive users. Table 125 and Table 126 present the frequency counts of hardware design quality rankings for the stratified assessment. Figure 53 illustrates the distribution of hardware design quality ranking for all three systems, as ranked by current Motive users. Figure 54 illustrates the distribution of ranking for current Motive users and non-Motive users for hardware design quality.

Table 125 – Frequency counts for overall order ranking of hardware design quality for current Motive users.

System	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
Motive	14	27	3	44
LytX	22	10	11	43
Samsara	8	7	28	43

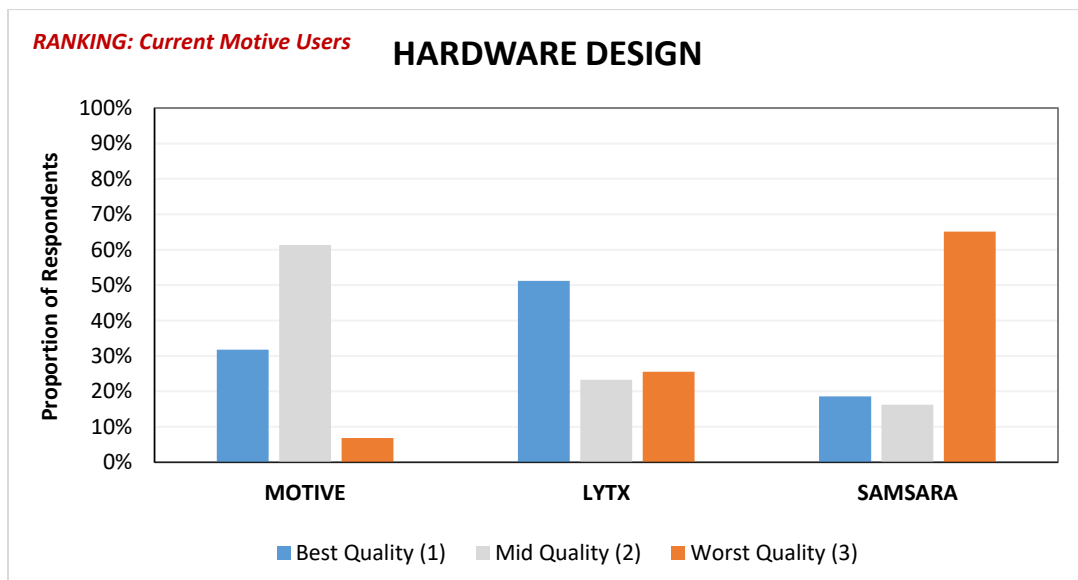


Figure 53 - Ranking for hardware design quality comparing current Motive users’ opinions of all three systems.

Table 126 – Frequency counts for ranking of hardware design quality for current Motive users compared to non-users.

Motive User	Counts of Best Quality (1) Rankings	Counts of Mid Quality (2) Ranking	Counts of Worst Quality (3) Ranking	Total Count
No	24	21	9	54
Yes	14	27	3	44

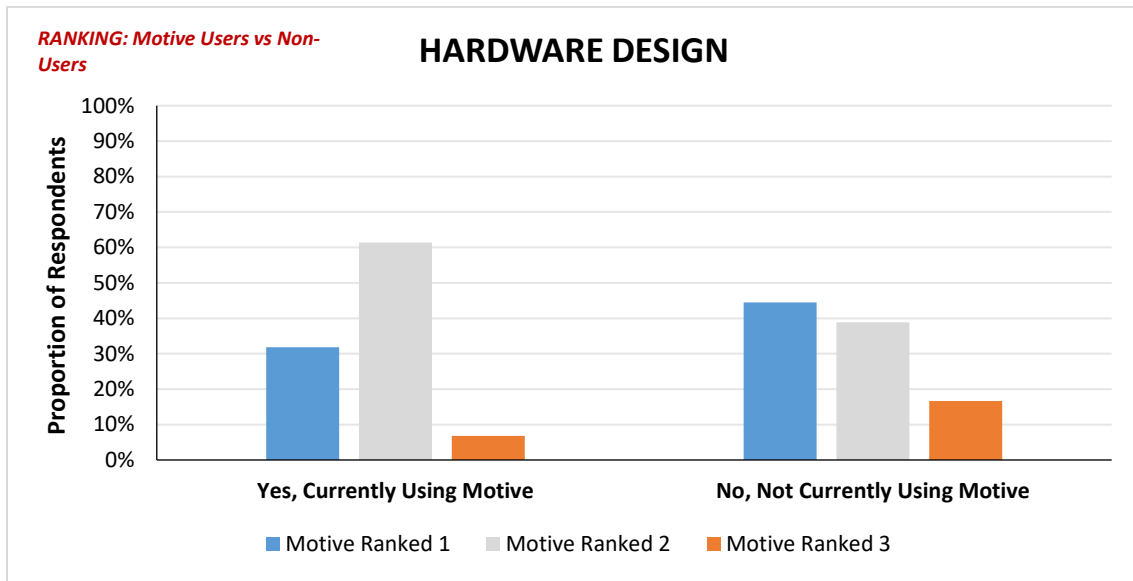


Figure 54 - Ranking for hardware design quality comparing current Motive users and non-users.

Table 127 – Comparison of time to alert and average successful alerts across all systems.

Task Type	# of tests	Motive	Samsara	Lytx
Overall	234	81% (9.4 sec)	26% (12 sec)	34% (17 sec)
Texting	39	92% (6.9 sec)	47% (7.5 sec)	18% (14.5 sec)
Phone Call	39	95% (7.5 sec)	38% (5.5 sec)	28% (21 sec)
Phone in Lap	39	53% (7.0 sec)	15% (8.7 sec)	8% (16.2 sec)
Close Following	39	67% (15.5 sec)	18% (26.1 sec)	28% (13.1 sec)
Seat Belt Use	39	100% (15.5 sec)	0% (n/a)*	100% (20.1 sec)
Rolling Stop	39	77% (3.8 sec)	0% (n/a)	n/a

*VTTI was informed after the study was completed that the seatbelt alert for Samsara’s device was not properly enabled. This task was removed from the calculation of Samsara’s overall percentage.