Measuring Performance of Direct Support Shift Changes in Independent Living

Scott Guirlinger, Andrew McFarland, Eric Weflen

November 24, 2020

Abstract

Shift changeover, or shift handover, is a critical transition task that occurs multiple times every day in industries all over the world. Effective shift changeover often has a direct impact on the safety and satisfaction of both workers and clients. In our study, we predicted that frequency of questions, relative duration of eye contact, and IBM Watson's "joy" score would correlate, and therefore measure, the effectiveness of a shift changeover. To test this theory, we observed teams of direct support professionals (DSPs) who care for adults with disabilities in their own homes. Four shift changes across eight individuals on two teams were recorded and the video analyzed for verbal and non-verbal cues, particularly questions and eye contact. Audio was transcribed and analyzed by IBM Watson to produce a "joy" score. Contrary to our hypothesis, our results indicated a weak negative correlation between these three metrics and shift changeover effectiveness, as self-reported by the workers on post-changeover questionnaires. We discuss the biases and other factors that we believe led to this result and suggest important considerations for future research in this subject area.

Introduction

Shift changeover, also called shift handover, is the task that takes place anytime one worker's shift ends and another worker's shift begins. To ensure a smooth and successful transition, shift changeover typically requires the team members to efficiently communicate the current state of the system, particularly anything abnormal, and any pending action items. This information helps to align the mental model of the incoming worker with the mental model of the outgoing worker. Communication can be verbal, non-verbal, or both. Non-verbal communication can include body language as well as documentation or other shared artifacts.

Industries and roles that operate on shifts most often have coverage 24 hours a day, with multiple shifts taking place each day. Multiple shifts means that multiple changeovers must take place each day as well. Shift changeovers represent an overhead cost where productive work is not being performed, and so changeovers are typically performed over a relatively short period, accounting for only a small fraction of the overall shift time. As seen in Figure 1, handover

accounts for only 5% of operations time but 40% of incidents. This comparison underscores the relative importance of effective shift changeover to successful operations and why our research team found this topic worthwhile to study.

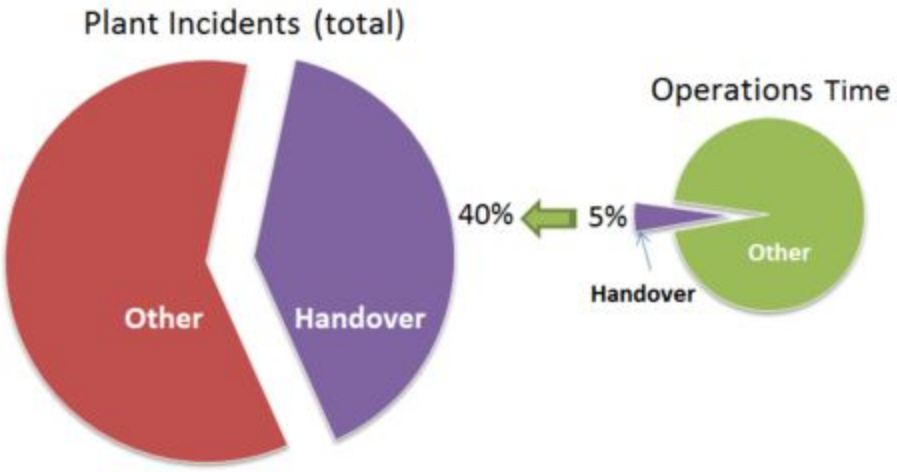


Figure 1: Plant incidents and operations time in the oil industry (Ford, 2015).

There are at least two roles in each shift change: the incoming worker and the outgoing worker. The outgoing shift worker brings knowledge of recent events on the preceding shift and must transfer this knowledge to the incoming shift worker, who brings fresh perspective and new energy. As seen in Figure 2, the incoming shift typically asks the majority of the questions and the outgoing shift provides clarification, although we would see in our own observations that this communication flow could reverse at times in the discussion. In many industries, shift changes take place between outgoing teams and incoming teams consisting of multiple members. Automated machine agents may be used in some cases, such as overnight surveillance shifts. For the sake of our study, we focused only on shift changes between one outgoing human worker and one incoming human worker.

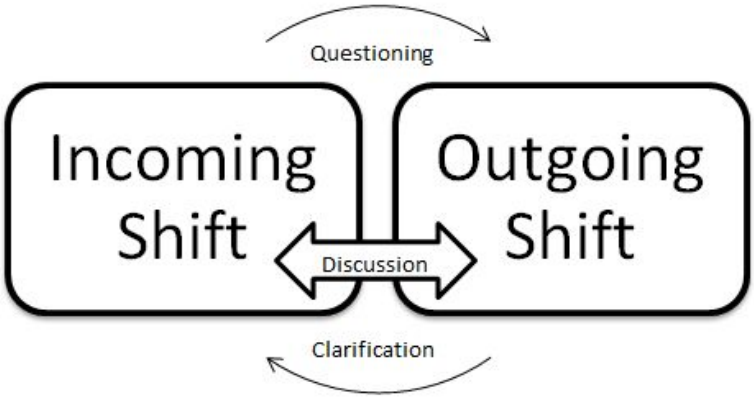


Figure 2: Simple communication diagram for typical shift changeover (Ford, 2015).

Often a manager is also present for a shift change. When involved, management typically provides oversight, including process control and quality assurance, but may also actively participate in the knowledge transfer, particularly if they have a perspective that may augment the knowledge being shared by the outgoing worker. Other stakeholders may also be present during a shift change, usually as observers.

Our research team had prior experience studying elements of shift changeover in other industries, such as military intelligence, manufacturing, and industrial facility operations. For this study, we had access to physically observe teams of direct support professionals (DSPs) in their workplace. DSPs work with people with disabilities (physical or intellectual) to help keep them safe, nurtured, and integrated with their community as much as possible. A goal of many direct support teams is to, as much as possible, enable their clients with disabilities to live independent and fulfilling lives.

Because many disabilities require supportive care 24 hours a day, DSPs often work in shifts on small teams. The quality of overall support given to clients is affected by the success or failure of transitions between staff members. For this reason, DSP shifts are often not strictly sequential but contain some overlap in time where both outgoing and incoming workers are present before the outgoing worker eventually departs. Also, while shift and changeover times may be specifically planned and scheduled, the actual times that these take place tend to fluctuate around the immediate needs of clients. Inevitably, an outgoing DSP may still be with a client at the prescribed time that an incoming DSP arrives and the shift changeover discussion between the two workers must wait until a mutually convenient time.

Success in shift changeover is typically measured by an absence of incidents in the subsequent shift. High performance looks like operations running smoothly, while low performance can cause dramatic and unexpected events. In the direct support field, the definition of success is often expanded to include client satisfaction. Examples of consequences of poor changeover that were shared by the participants in our study included medication errors, missed appointments, incomplete tasks (e.g. laundry, cleaning, or meals), duplicated tasks (e.g. two dinners, leading to overfeeding), and upset clients. The impact of these consequences range from simple emotional instability to serious physiological danger.

Still, DSPs are not often evaluated on their performance in the shift changeover, but rather on their performance during their actual shift, i.e. how few incidents occurred with clients under their care. So while shift changeover plays an important role in performance (refer back to Figure 1), that role is almost always indirect in nature. As mentioned above, the goal of a shift changeover is to obtain the most accurate mental model possible to then be able to effectively perform during the shift.

Background

To form a foundation for our project, we looked to prior research to determine what others measured when analyzing shift changeover. We also sought to determine how previous

research defined changeover effectiveness and to learn about factors that may lead to effective changeover. The information we found was instrumental in our development of metrics to measure in the analysis of DSP changeover.

Upon reviewing existing research, we found that the primary measures used to determine changeover effectiveness have historically been participant confidence, or reported effectiveness of the changeover, and error rate (Anders et al., 2016). The greatest inhibitors of effective changeover are factors that prevent effective communication generally. One of the primary detractors of effective changeover that has been observed is the absence of questions, whether that absence is due to a lack of opportunity, overconfidence, or a simple failure to ask (Anders et al., 2016; Patterson et al., 2005). The knowledge that the number of questions asked can have a positive correlation with shift change effectiveness led us to use the number of questions asked as the metric to measure effectiveness in verbal communication.

We also found that interruptions and distractions have shown a negative correlation with shift change effectiveness (Anders et al., 2016, Patterson et al., 2005). Interruptions from individuals or objects not directly involved in changeover or a distraction that removes participant focus from the items relevant to changeover can cause those relevant items to be missed or forgotten (Anders et al., 2016, Patterson et al., 2005). To measure non-verbal distractions and inhibitors to communication during changeover, we measured the amount of eye contact during the shift change sessions. If eye contact is maintained, it is unlikely that participants are distracted by irrelevant factors. Shift change has also been shown to be an effective way to mitigate employee anxiety (Mower et al., 2008). To capture this reduction in anxiety, we gathered employee confidence levels through a questionnaire. This level of confidence in the shift change could be a good way to determine its overall effectiveness (Mower et al., 2008). Some other metrics used in the analysis of shift change have been the duration of the shift change, the location that the participants conduct the changeover at, the mode of the change (written vs verbal), and the type of communication used. However, most of these metrics did not fit the scope or goals of our study, leading us to eventually exclude them from consideration.

Armed with this information from previous studies and the introduction to typical operations of a DSP team, our research team sought to determine if verbal and non-verbal cues could be measured to predict the effectiveness of a shift changeover. We also wanted to analyze if verbal and non-verbal cues could indicate the level of comfort between incoming and outgoing shift workers.

Methods

Hope House Foundation is a charitable organization in Norfolk, Virginia providing support services to adults with disabilities for more than 50 years. As of the publishing of this paper, they supported over 120 individuals in their own homes. DSPs constituted the majority of their staff and were grouped into 15 teams by location. These teams provided some level of support 24 hours a day, so shift changeovers took place multiple times each day.

Direct verbal conversations between outgoing and incoming DSPs occurred in the team office during each shift changeover. Additionally, DSPs shared information through three primary artifacts: official paper records used for billing, informal online logs, and wall calendars or whiteboards. The online logs consisted of posts, including pictures, on Yammer sites for each team. Incoming DSPs frequently would review these online logs before arriving to increase their situational awareness. Operating procedures, assessments, and other documents are also readily available electronically through this Yammer interface so that staff can access them from any location in the field.

We studied eight staff members across two different DSP teams. Both of these teams typically operated in three shifts each day: morning (7am ~ 3pm), evening (3pm ~ 11pm), and overnight (11pm ~ 7am). (Times are not exact and are shown for example purposes only.) The participants in our study were all native English speakers. Seven participants were female, and one was male. Five participants identified as Black or African-American; the other three participants identified as white. Participants self-reported having an average of 3.5 years (SD = 2.2 years) of experience in DSP work, and the majority had known the other worker in the shift changeover between 6 and 24 months. While the gender demographics in our study may seem skewed towards females, this distribution (88%) is consistent with other studies of DSPs (Bossink et al., 2019).

We began the study of these staff members by observing and recording four shift changes. The participants were aware that they were being recorded and each had given consent to be recorded prior to the observed changeover. The video recording devices were new introductions to the environment, and could have slightly influenced results, particularly in questionnaire responses. After each shift change, a questionnaire was given to both the incoming and outgoing shift workers (see Appendix A). The purpose of this questionnaire was to gather the basic demographic information referenced above, as well as capture participants' perceived effectiveness of the shift change process and their level of comfort with the other worker. The questionnaire also attempted to measure the quality of their relationship with the other participants in the shift change by asking how well they knew their co-worker on a scale ranging from "I hardly know them" to "they are like family". Another measure of the quality of their relationship was the question asking how comfortable they are sharing information with their co-worker ranging from "just what is necessary" to "private matters". In the survey, each DSP reported how effective they felt their shift change was on a scale of 1 to 5. Only DSP perception of effectiveness was captured. Ideally, we would have also captured managers' and clients' perceived effectiveness but were limited in our observation time, requests of managers' time, and access to clients.

After these artifacts were collected, the videos and questionnaires were shared with the project team members. The videos were analyzed using BORIS, a software tool that allowed us to tag each occurrence of a question, laughter, an interruption or distraction, an instance of social support, a reference to an artifact, or a gesture. BORIS also allowed us to track the amount of time that participants were facing each other or making eye contact. Multiple team members tagged each of these interactions to ensure that we were as accurate as possible in our data

collection. Most of these interactions were not used in our final metrics, as previous research did not seem to indicate that they would be useful in measuring changeover effectiveness. Our strategy was to capture as much data as possible, but we decided to focus on the three metrics that we selected because, based on prior studies, we felt they would carry the most weight in predicting outcomes. The data is available for future research on these other metrics if someone else decides to study a similar subject.

After tagging each of these interactions seen in the videos using BORIS, we transcribed the audio from the videos using Otter.ai. It is worth noting that this application was not perfect, causing some mistranscribed sections, as discussed later in this paper. After the audio was transcribed, we analyzed the audio using IBM Watson and Grammarly's Tone Detector tool to group statements into categories identifying the primary tones and attitudes expressed in the videos.

Table 1: Metrics chosen to predict shift changeover effectiveness.

Measure	**Verbal Communication**	**Non-Verbal Communication**	**Verbal Content Cues**
Metric	Frequency of questions	Relative duration of eye contact	IBM Watson "Joy" score
Variable Type	Ratio	Percentage time duration	Text analysis score
Data Source	Video recordings	Video recordings	Transcribed video recordings
Collection Frequency	Each time the event triggered	Each time the event started and stopped	After each interview
Formulas & Thresholds	Count number of questions asked in one complete shift change; divide this count by total instances of communication	Eye contact was captured when either the incoming or outgoing workers looked towards one another; sum these instances, then divide by total duration of shift change	IBM Watson uses Natural Language Classifier, which combines inputs from SVMs and a Convolutional Neural Network to classify human language
Comments	Prior research showed that lack of questions can be detrimental to changeover effectiveness (Anders et al., 2016, Patterson et al., 2005)	Lack of attention could cause ineffective changeover; eye contact could indicate levels of attentiveness, an attribute of successful changeover (Anders et al., 2016, Patterson et al., 2005)	Level of joy present in communication could indicate participants' levels of engagement and comfort; levels of comfort and reduced anxiety were shown to be indicators of successful changeover (Mower et al., 2008)

The measures that we chose for this project were verbal communication, non-verbal communication, and verbal content cues. Table 1 contains detailed attributes of these three measures. We needed to normalize our questionnaire results so that we could consistently compare them with our metrics. To do so, we divided the Likert scale responses from question 6 (comfort at work), question 8 (coworker relationship), and question 9 (perceived effectiveness) by the total available response (5 in each case).

Results

There were several different event logs that documented the interactions, allowing us to derive our metrics. From BORIS we had the tagged instances of verbal and non-verbal communication, as well as interruptions that are annotated in the software for future reference. These annotations could be exported in several formats for visualization and further analysis. We also had the otter.ai transcriptions and the tone analysis annotations made by the IBM Watson tone analysis system. In this section, we will review each of these sets of data and apply our metrics.

BORIS Video Analysis Event Logs

For each shift change the video was analyzed for ten different forms of communication, with time stamped tags being used in BORIS. A useful way to visualize the frequency of different communication forms was to mark them on a timeline, as seen in Figure 3 and Figure 4 for the first and third observed shift changes. (Refer to Appendix C for higher resolution annotated event logs for all shift changeovers.) There was a separate timeline for each actor in the shift change. Interruptions and distractions from outside sources were marked with “No Focal Subject”. Figure 3 and Figure 4 provide visualization of the different levels of communication in shift change #1 compared to shift change #3.

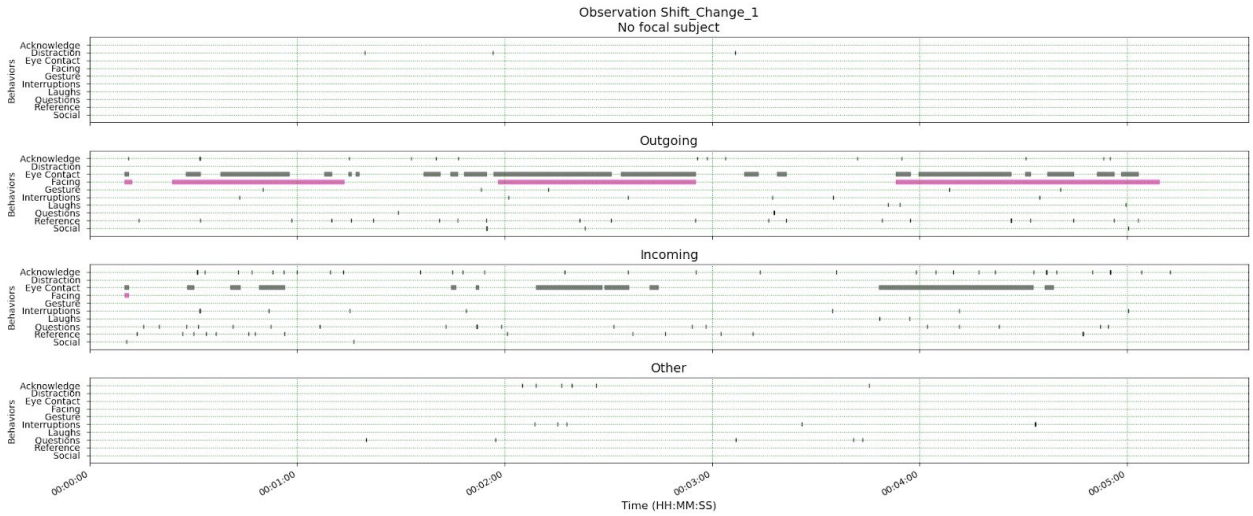


Figure 3: Event log exported from BORIS software for shift change #1.

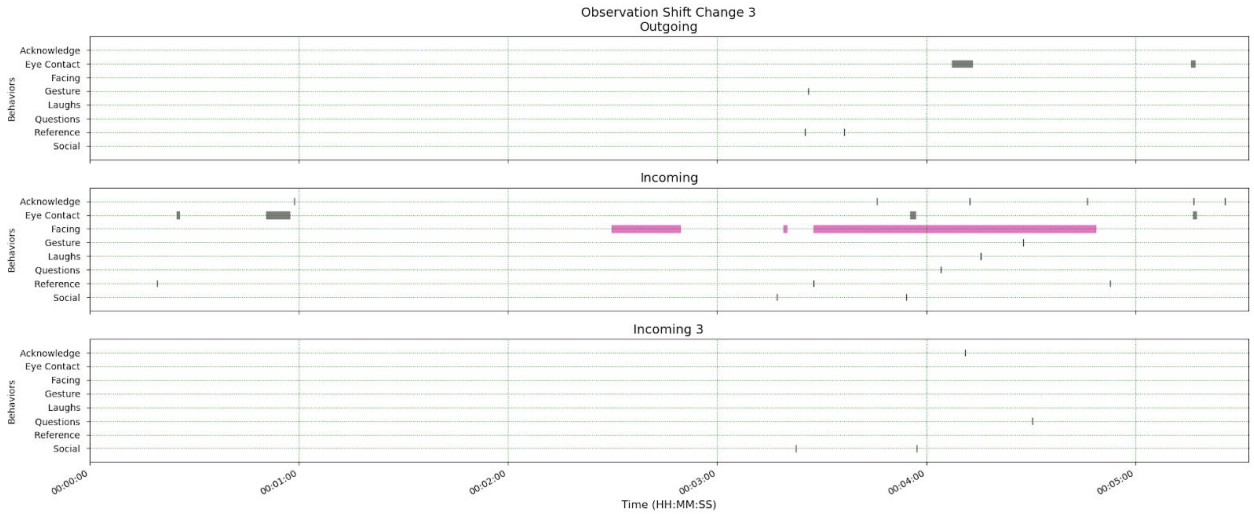


Figure 4: Event log exported from BORIS software for shift change #3.

The total number of communication instances was much greater in shift change #1 than in shift change #3, having 178 instances compared to just 30, respectively. While total instances of communication could have shown us a level of activity, we were more interested in our specific metrics of frequency of question asking, duration of eye contact, and the presence of the Joy tone to better understand the effectiveness of the communication, not just the quantity. In a later section, we will take a closer look at these metrics.

Dialog Tone Analysis Event Logs

We attempted to simulate a virtual assistant that listens in to a shift change conversation to analyze the tone to see if there is potential for such a system to give the participants feedback on the effectiveness of their communication. To achieve this, audio from each of the four shift changes was uploaded to otter.ai for transcription (see Appendix D). No corrections were made to the automatically generated voice-to-text files so that the current state of this technology could be evaluated. A review of the transcripts showed that otter.ai struggled to correctly transcribe much of the audio and in certain cases completely missed several of the DSPs' statements in the interaction. This may have been due to the quality of audio, which at times was relatively quiet or obfuscated by background noise and therefore could be difficult for the authors to clearly understand. A real implementation of such a system would likely require work to improve the automatic transcription process. However, the IBM Watson Tone Analysis model was understood to evaluate, to some extent, the context of the words, and not just the presence of words, and may have returned usable information even for low quality transcriptions. A more in-depth analysis would be needed to evaluate how errors impact the accuracy of this system.

The output of the IBM Watson Tone Analysis was a JSON file that gave a document level score for tones having a value over 0.5, and a sentence-by-sentence analysis evaluation of each of the tones (see Appendix E). It also provided a visual output, but did not export results in this format. Figure 5 shows an excerpt of the visual output of the Joy and Tentative tones for the

higher rated Shift Change #1 (Figure 5a & 5b) and for the lower rated shift change #3 (Figure 5c & 5d). Each sentence received a strength rating that was signified by the level of shading. A document score was derived from the number of instances and the strength of each instance. Shift change #1 had more instances with higher ratings, and therefore an overall higher rating, than shift change #3.

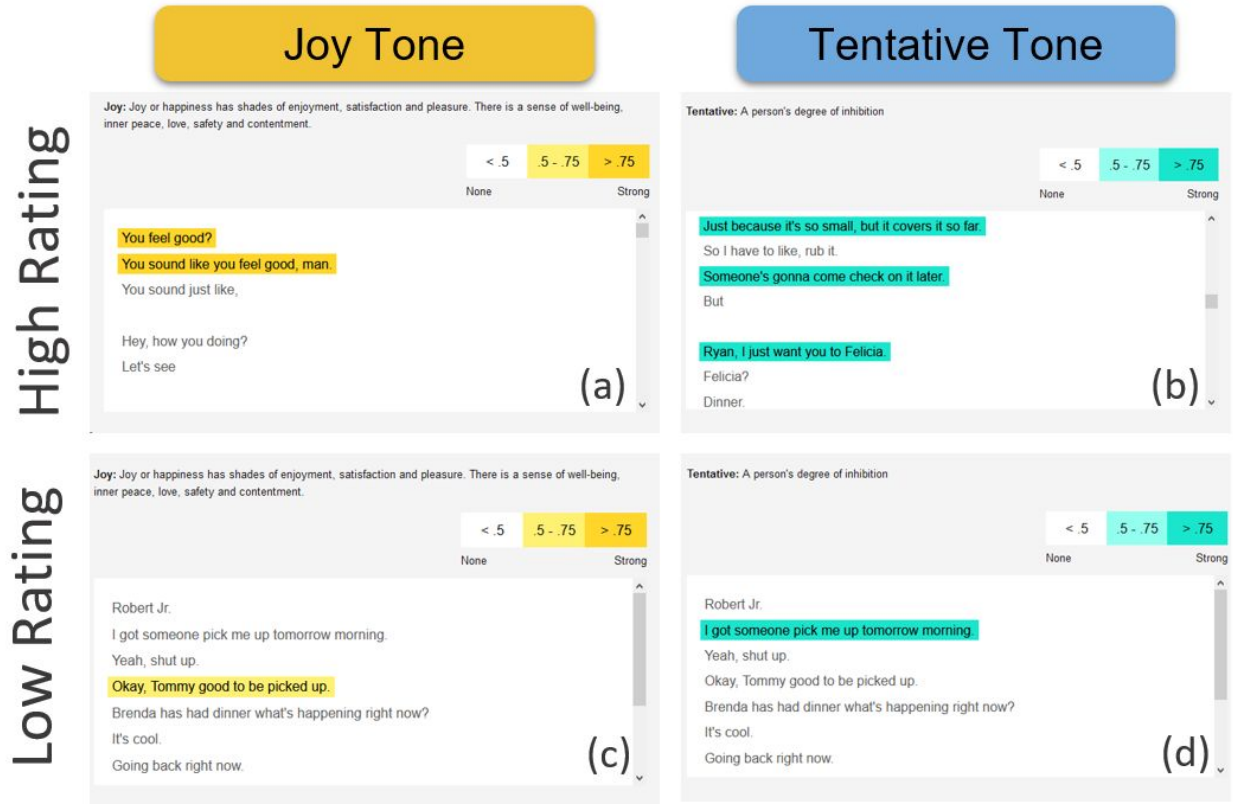


Figure 5: Excerpt from event log for IBM Watson tone analysis. (a-b) Examples of a high rating for both the Joy and Tentative tones from shift change number #1. (c-d) Examples of a low rating for the Joy and Tentative tones from shift change number #3.

Comparing the results of the IBM Watson tone analysis to the Grammarly tone analysis of the same otter.ai transcript showed disagreement between the two systems (Figure 6). While it could be difficult to compare the two systems because they use different metrics, there seemed to be a possible relationship between Grammarly's "Informative" tone and IBM Watson's "Tentative" tone which are only found in the first and fourth observed shift changes. The IBM Watson Tone Analysis system is more well documented than the Grammarly system, which potentially made the results easier to interpret. In addition, the ratings were more transparent since the model identifies the areas that drive the tone, and specific because it provides numerical values. These attributes made Watson potentially better suited than Grammarly for use as a metric to measure shift change effectiveness.

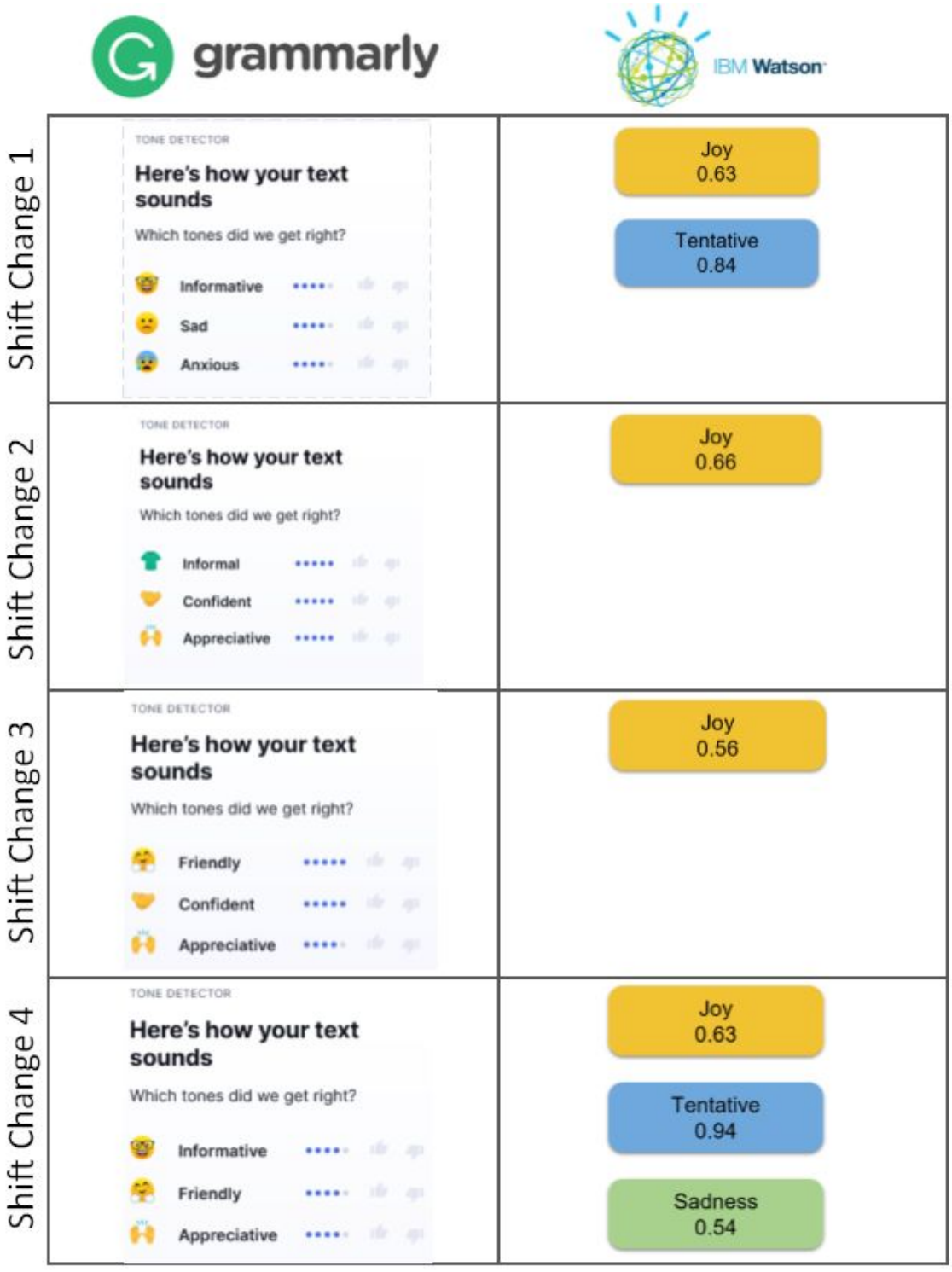


Figure 6: Comparison of Grammarly and IBM Watson Tone analysis.

Questionnaire Results

Figure 7 plots the individual responses of the post-shift-changeover questionnaires completed by the DSPs (see Appendix B for raw data) against the self-reported effectiveness of the shift change. There were no strong relationships in the data, but there are some general trends, such as DSPs who had known each other longer tended to report the shift change as being more effective (Figure 7b). The weak relationships may have been due to the low number of data points. The DSPs reported high values for the questions meant to judge the quality of their interpersonal relationships, with the exception of one worker (Figure 7c & 7d).

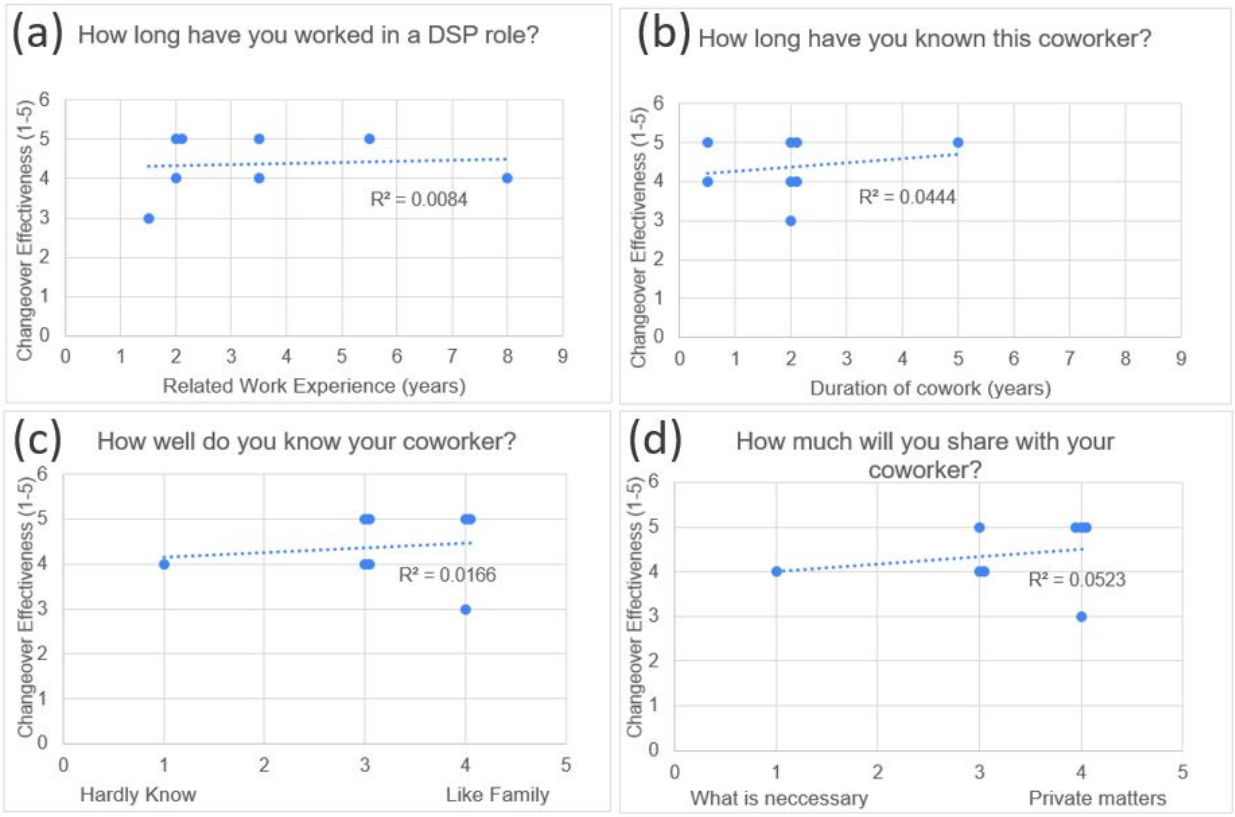


Figure 7: Results from the survey data for individual participants in all shift changes. (a) Work experience in DSP role compared to perceived effectiveness of the current shift change. (b) Duration of time coworkers knew each other against effectiveness. (c) Familiarity with coworker on a scale of 1 (hardly know them) to 5 (they are like family) against effectiveness. (d) How comfortable coworkers are sharing private matters against effectiveness.

Shift Change Effectiveness Metrics

The data collected in BORIS for the number of questions asked was divided by the total number of instances of verbal and non-verbal communication to get the ratio of all communication that was in the form of questions. This is compared with the ratio of time the DSPs had eye contact, and the IBM Watson Joy metric, all on a scale of zero to one (Figure 8). There was a general

trend that as the eye contact metric increased, the questioning metric also increased. In addition, shift change #3 had the lowest scores for all three metrics, suggesting that this shift change may have been the least effective. The otter.ai transcription struggled to pick up the DSPs' speech in shift change #3, which may be the cause for the low score in that shift change. The IBM Watson Joy metric did not follow the trend of the other metrics. For example, shift change #1 had higher scores for the questioning and eye contact metrics as compared to shift change #2, but a lower score for the Joy metric.

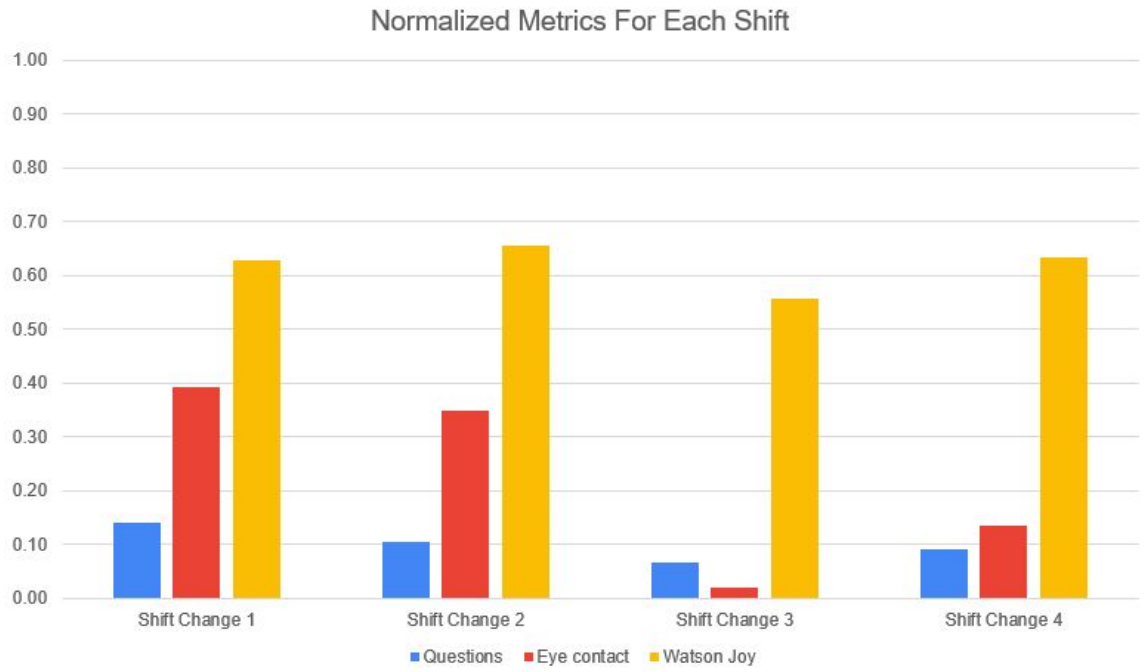


Figure 8: Metrics for shift change effectiveness.

While our metrics in Figure 8 suggest that shift change #3 was the least effective, when we compared those results to normalized self-reported shift change effectiveness, we see that it scored the highest (Figure 9). This shift change had three participants, but only one survey result, which may be influencing this result. Looking at the other shift changes, there appeared to be no clear relationship between self-reported effectiveness and the questioning and eye contact metrics. However, the Joy metric moved in the opposite direction as the self-reported effectiveness, which is unexpected. The comfort at work value, which is the normalized average response to the years of work experience questions for all DSPs in the shift, did not seem to follow along with our three metrics either. The coworker relationship value, which is the average response by all DSPs in the shift change to the question about what type of information they are comfortable sharing with their coworker, normalized to a scale of 0 - 1, also showed little relationship with our metrics.

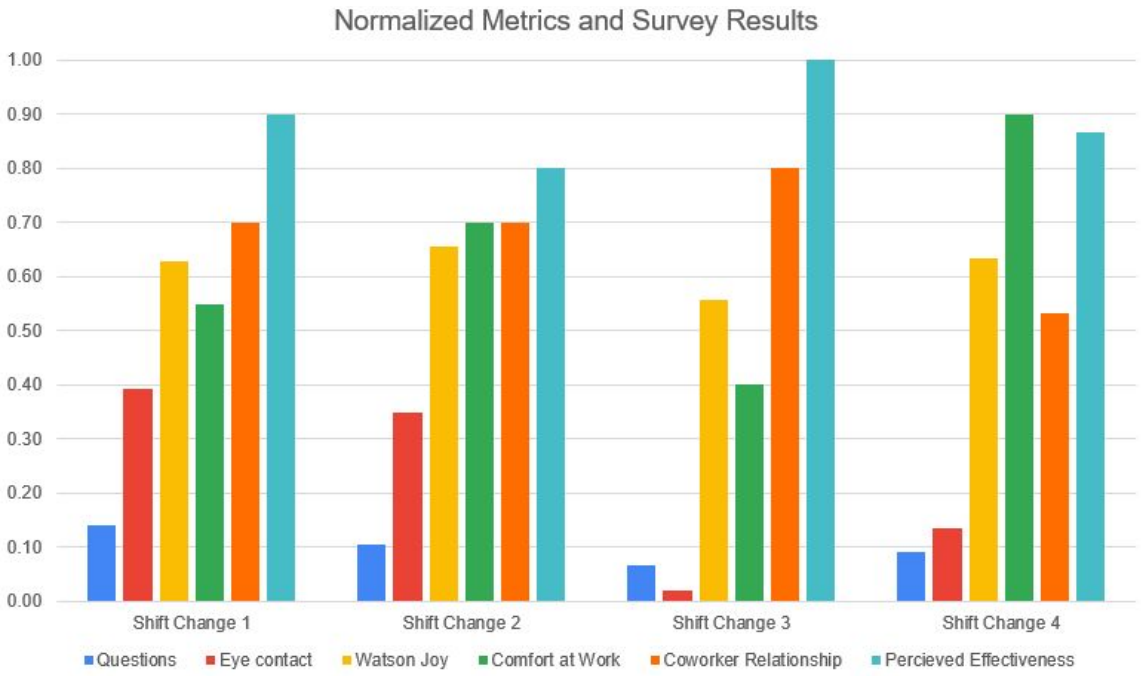


Figure 9: Shift change effectiveness metrics and normalized survey results.

A simple linear regression was used to model the relationship between each metric and the normalized average self-reported effectiveness of each shift change (Figure 10). In every instance, we observed a negative relationship between our metric and the self-reported effectiveness. The Watson Joy rating had the steepest slope and the strongest relationship. Eye contact had a moderate relationship, and questioning was the weakest.

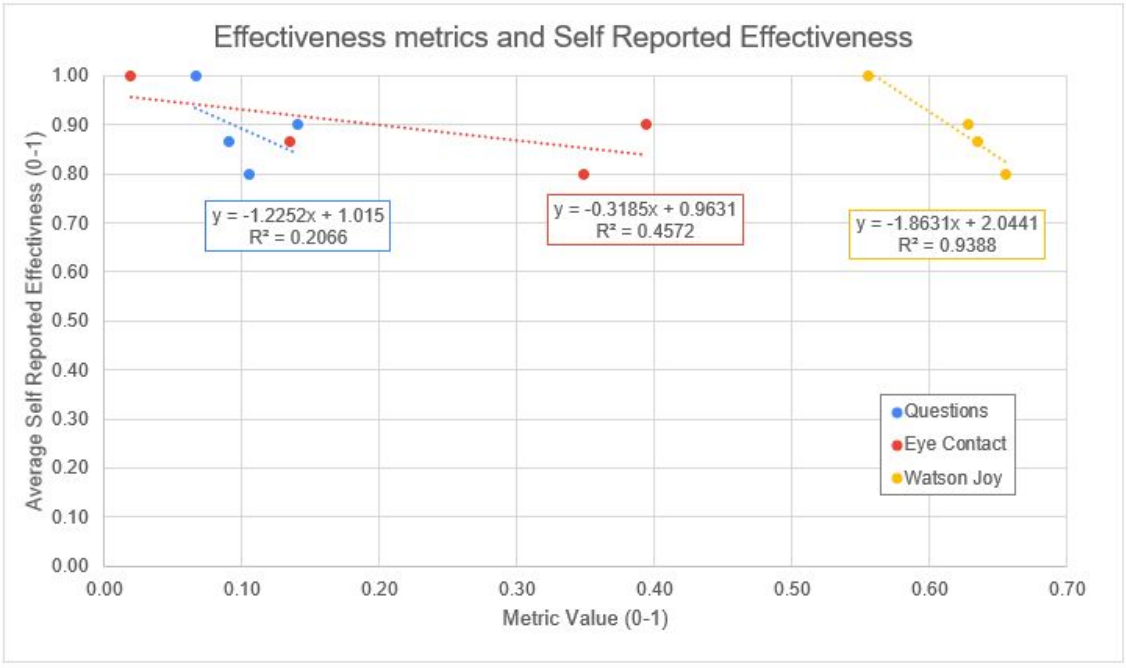


Figure 10: Predicting self-reported effectiveness of the shift change using the three metrics.

Discussion

The results suggested that the three chosen metrics had potential for predicting effectiveness of a shift change, with the Watson Joy metric showing the strongest relationship with self-reported effectiveness. It was difficult to determine the reliability of the metrics without an objective measure of how effective the shift change was, which would require the measurement of errors that occur during the shift and determining if the errors were related to poor information transfer during the shift change. This metric would require significantly more time to measure, as it would require observation of the entire shift. In addition, by informing the participants that they were being observed and recorded, we likely influenced both their behavior during the shift change and their responses afterwards to some degree. If possible, future studies would be conducted without this form of bias.

The question and eye contact metrics both showed weak relationships to the self-reported shift change effectiveness. The small sample size of this study may have contributed to that result, and the authors recommend that a more extensive study be conducted to either validate or invalidate these findings. Our review of past research found that sample sizes were often limited in studies evaluating the effectiveness of shift changes, which may be due to challenges similar to those that we faced with coordinating observations and the time consuming nature of gathering the data.

It was unexpected that the relationship we found for all metrics exhibited a negative correlation, even though our literature review suggested we should find a positive correlation. The negative relationship may suggest that our metrics were able to make some predictions of shift change performance, but in the opposite direction as expected, and as reported by other researchers. What may be more likely was bias in the self-reported shift change effectiveness value. It may have been possible that a shift change that feels effective to the DSP, may in fact be less effective at transferring the necessary information to avoid errors. This suggests that there would be value in digging deeper by measuring errors during a shift in addition to our metrics to determine if there is a positive relationship. It was concerning that our metrics suggested that when a DSP is confident in the effectiveness of a shift change, that the shift change may be less effective. This could lead to overconfidence as they conduct their work throughout the shift.

The Watson Joy metric may have had reduced effectiveness due to the poor quality of voice-to-text transcription performed by otter.ai. Figure 11 shows an excerpt from shift change #1 that was difficult to comprehend, even though the conversation was understandable when the authors reviewed the video. Previous experience with the otter.ai software suggested that it had the ability to transcribe when there is high quality audio and clear annunciation. In several of our videos there were times when it was difficult for the authors to comprehend exactly what the participants were saying, and in those instances otter.ai completely failed to transcribe any of the conversation. If this metric is going to be studied more or put into practice, it may require specialized audio equipment or an environment with less background noise. Since IBM Watson is likely taking the sentence structure into some consideration, it may have a level of fault

tolerance, but this fault tolerance would need to be defined to ensure that the transcription process meets those requirements.

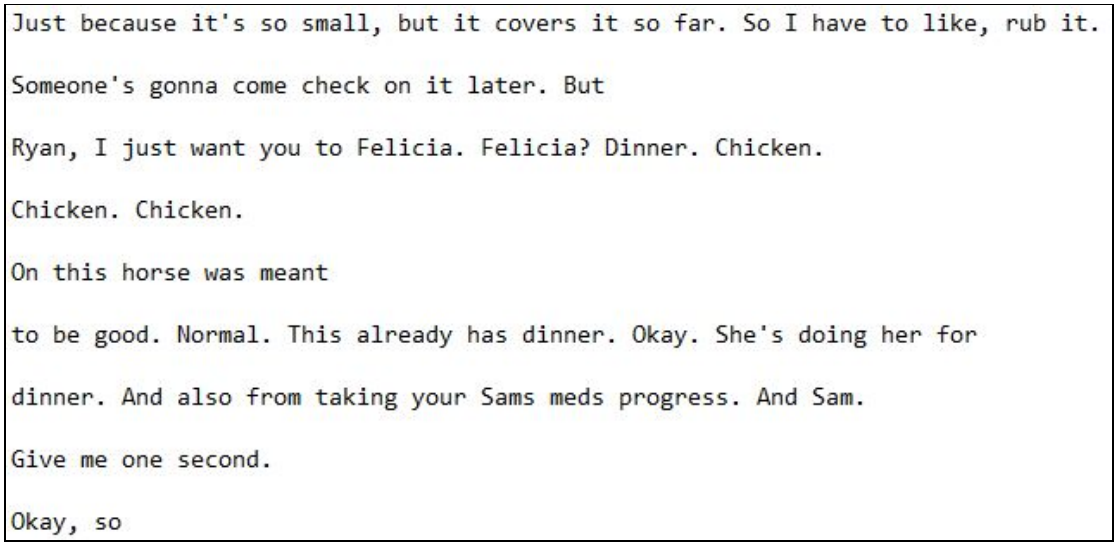


Figure 11: Excerpt from otter.ai transcription of shift change #1.

Our literature review revealed several other possible metrics for which data was gathered in this study, but due to limited time were not evaluated. There may be potential for blended metrics to outperform our simple metrics by combining the eight different verbal and non-verbal communication forms, interruptions and distractions, and the IBM Watson tone analysis. In addition, there may be other performance measures that could be used to validate these metrics, such as patient satisfaction or manager rating of effectiveness. While these measures would still be subjective, they should have different bias with a less direct effect in evaluation against the objective metrics. After all, client and manager satisfaction are likely important to the DSPs' success.

If a relationship could be more clearly shown between our metrics and performance, there is potential for their use in a broad range of applications. Because the metrics are not specifically tied to this application, there is potential to be used in any transfer of information between individuals or groups. This could include nursing shift changes in a hospital, retail worker shift changes, or even to evaluate the shift change stand-up meetings in manufacturing. There are also possibilities beyond shift change situations. For example, a project team meeting would consist of transfer of information that could be analyzed using these metrics to possibly predict their ability to deliver results.

We briefly presented results looking at the DSPs' level of expertise, through years of experience, and did not see a clear trend that would suggest that our metrics could differentiate an expert and a novice. However, an objective study of error rates during a shift may uncover a different relationship than we found. It is likely that an expert and novice have different communication needs. For example, an expert may be able to still avoid an error that a novice could not identify after experiencing poor communication during a shift change.

If we were to conduct the study over again and had more time and resources, we would like to take into account the other communication tools used by the DSPs, such as Yammer chat, written reports, and the shared team calendar. The use of these tools may reduce the need for effective verbal communication during the shift change, or reduce errors in the presence of ineffective verbal communication. We would also gather a much larger data set that would be able to more clearly confirm or reject the relationships that we are trying to study.

Conclusions

This study consisted of observations of four shift changes between incoming and outgoing direct support professionals at two Hope House Foundation team locations. Recorded observations were evaluated for verbal and nonverbal communication. These observations were transformed into three metrics that may have potential to measure human performance during the shift change. The frequency of questioning, duration of eye contact, and a textual analysis of tone using machine learning methods were evaluated and found to have some potential as metrics.

Measuring effectiveness of shift changeover was found to be critical for client safety and satisfaction, but data collection consumed significant time and still did not provide the measurements and correlations that we had anticipated. Verbal and non-verbal cues showed a weak, but potential inverse relationship with self-reported effectiveness of a shift changeover. Further, verbal and non-verbal cues showed a weak, but potential relationship with level of comfort between incoming and outgoing shift workers. Watson Tone Analysis showed a strong inverse relationship with self reported effectiveness of the shift change. These results suggest that there is potential to use these metrics to evaluate shift change performance. However, the inverse relationship was unexpected, and may be related to the use of subjective self-reported effectiveness instead of comparing results to actual performance measures.

Given more time, we would have liked to collect more feedback from the shift workers, their management, and their clients. Specifically, we would like to collect errors made and categorize those errors to identify mistakes that could have been avoided with a more effective shift changeover. Ideally, we would suggest that shift workers complete questionnaires at both the start and the end of their shift, managers provide incident reports and employee evaluations, and clients be surveyed to gauge their overall satisfaction from each shift. And we believe it would be beneficial to establish much longer periods of observation so that workers become accustomed to the observers and the Hawthorne effect might factor less into the results. We would also like to extend this research to understand if external communication tools are necessary to create a positive perception of a shift changeover and how often these external communication tools should be referenced.

Acknowledgements

Scott Guirlinger coordinated and conducted the observations of the DSP teams at Hope House Foundation, and collected the questionnaire data. Scott authored the introduction. Andrew McFarland assisted with the literature review, authored the background and methods sections, and produced the final version of the accompanying presentation. Eric Weflen championed the literature review and dialogue tone analysis and authored the results and discussion. All of the authors participated in the BORIS video analysis and regular team discussions.

References

Blattgerste, J., Strenge, B., Renner, P., Pfeiffer, T., & Essig, K. (2017). Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks. *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*. doi:10.1145/3056540.3056547

Bossink, L. W. M., Putten, A. A. J., Paap, M. C. S., & Vlaskamp, C. (2019). Factors associated with direct support professionals' behaviour in the physical activity support provided to people with intellectual disabilities. Journal of Intellectual Disability Research. https://doi.org/10.1111/jir.12616

Ford, B. (2015, December 7). Shift-Handover: 5% of Operations Time, 40% of Incidents. *Yokogawa*. https://blog.yokogawa.com/blog/ensuring-quality-shift-handover

Mower, A., Nguyen, R., & Frank, K. (2020). Evaluation of Technology Accessibility and User Sentiment in Learning Through Virtual Reality Modality. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. Retrieved October 12, 2020, from https://dl.acm.org/doi/10.1145/3334480.3375205

Yoo, S., Lee, Y., & Kim, S. (2020). Learning by Doing: Evaluation of an Educational VR Application for the Care of Schizophrenic Patients. *CHI EA '20: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. Retrieved October 12, 2020, from https://dl.acm.org/doi/10.1145/3334480.3375205

Appendix A: Research Questionnaire

HCI 681, Fall 2020, Iowa State University

Please circle your responses to multiple choice questions. All responses will be kept anonymous.

Basic Demographics

1. **Is English the primary language that you speak at home?**

   Yes No

2. **To which gender identity do you most identify?**

   Female Male Non-binary Other _____

3. **Are you of Hispanic, Latino, or of Spanish origin?**

   Yes No

4. **How would you describe yourself?** (circle as many as apply)

   American Indian or Alaska Native
   Asian
   Black or African American
   Native Hawaiian or Other Pacific Islander
   White
   Other _____

Job Experience

5. **Approximately how long have you worked in this direct support professional (DSP) role or a similar role?**

Shift Change

When reflecting on the coworker who is relieving you or to whom you are relieving on this shift change…

6. **How long have you known this particular coworker?**

   Less than 1 month

   1 month - 6 months

   6 months - 2 years

   2 years - 5 years

   More than 5 years

7. **How well do you feel you know your coworker?**

   1 = I hardly know them

   2

   3

   4

   5 = They are like family to me

8. **How much are you willing to share with the other person?**

   1 = I'm only going to share what I have to with them

   2

   3

   4

   5 = I would entrust them with private personal matters

9. **How would you rate the effectiveness of knowledge transfer on this particular shift changeover?**

   1 = Room for improvement

   2 =

   3 =

   4 =

   5 = Nailed it

Thinking about shift changes in general…

10. **What level of impact do you feel a shift change has on your job?**

    1 = No impact
    2 =
    3 =
    4 =
    5 = Mission critical

11. **Why do you feel knowledge transfer at a shift change is required?**

12. **What types of errors have resulted from miscommunication in past shift changes?**

Thank you for your participation! If you have any questions or would like a copy of this research when completed, please feel free to contact Scott Guirlinger at 919-270-5189 or srguirli@iastate.edu.

Appendix B: Questionnaire Responses

ID	Role	Team	1	2	3	4	5	6	7	8	9	10
P1	Incoming	A	Yes	Female	No	Black or African American	3.5	2	3	3	4	5
P2	Outgoing	A	Yes	Female	No	White	2	2	3	4	5	5
P3	Incoming	A	Yes	Female	No	White	1.5	3	4	4	3	3
P4	Incoming	A	Yes	Female	No	Black or African American	5.5	4	3	3	5	5
P5	Incoming	A	Yes	Female	No	Black or African American	2	3	4	4	5	4
P6	Incoming	B	Yes	Female	No	White	8	3	3	3	4	5
P7	Outgoing	B	Yes	Male	No	Black or African American	3.5	3	4	4	5	5
P8	Outgoing	B	Yes	Female	No	Black or African American	2	3	1	1	4	5
Avg							**3.5**	**2.9**	**3.1**	**3.3**	**4.4**	**4.6**
SD							**2.2**	**0.6**	**1.0**	**1.0**	**0.7**	**0.7**

Responses to #11:

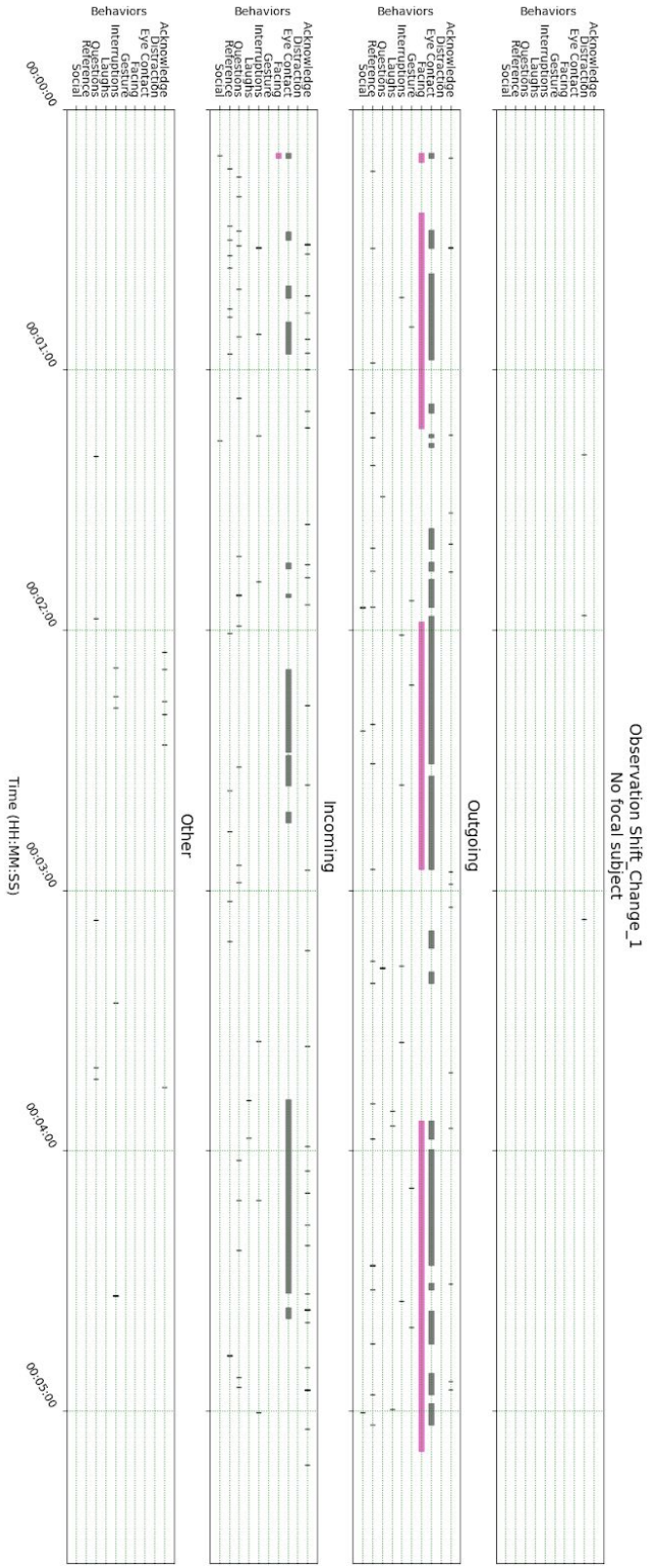
- In order to best support our people, we need to know how their day or night has been going and what are some of the things they need or want outside of the norm.
- I need to know what's going on (med changes, body chart updates, tasks that need to be completed) because that affects how well I am able to do my job during my shift
- It's our communication with one another about the day. With a job like this it's important to know what our clients want to do or what they don't want to do.
- Knowledge at a shift change is required because otherwise mistakes can be made. Most knowledge transfer is pretty routine here but when it's not is when it's most important. Med changes, injuries & PRN's are important info to have when starting your shift. Communication logs are not always accessible & it's much easier to get info at the start of your shift from your co-workers.
- It is important to know what the person ahead of you has / has not done so you know what needs / doesn't need to be done during your shift. It is also important to know about the general well being of the client (sick or not, their mood, etc.)
- Information is critical for services, quality of care, accuracy
- to make sure staff is prepared for the day
- Effective for shift to run smoothly

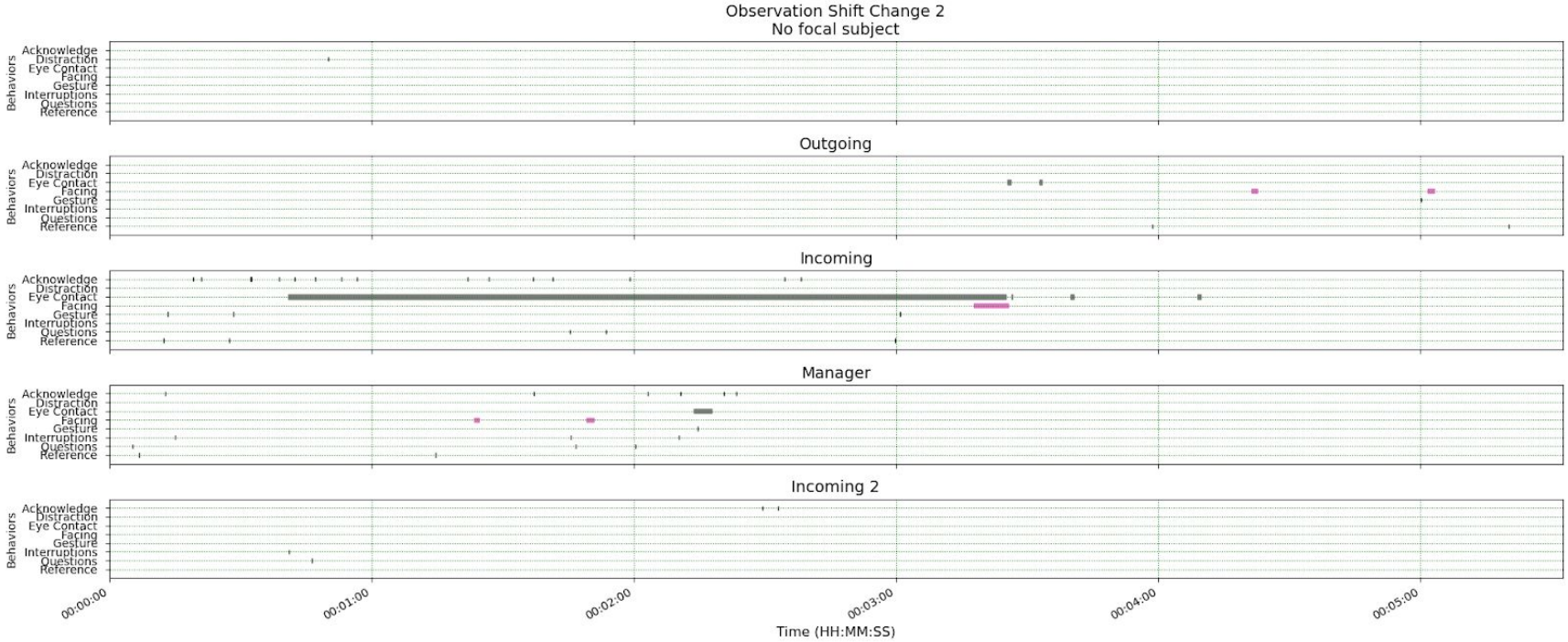
Responses to #12:

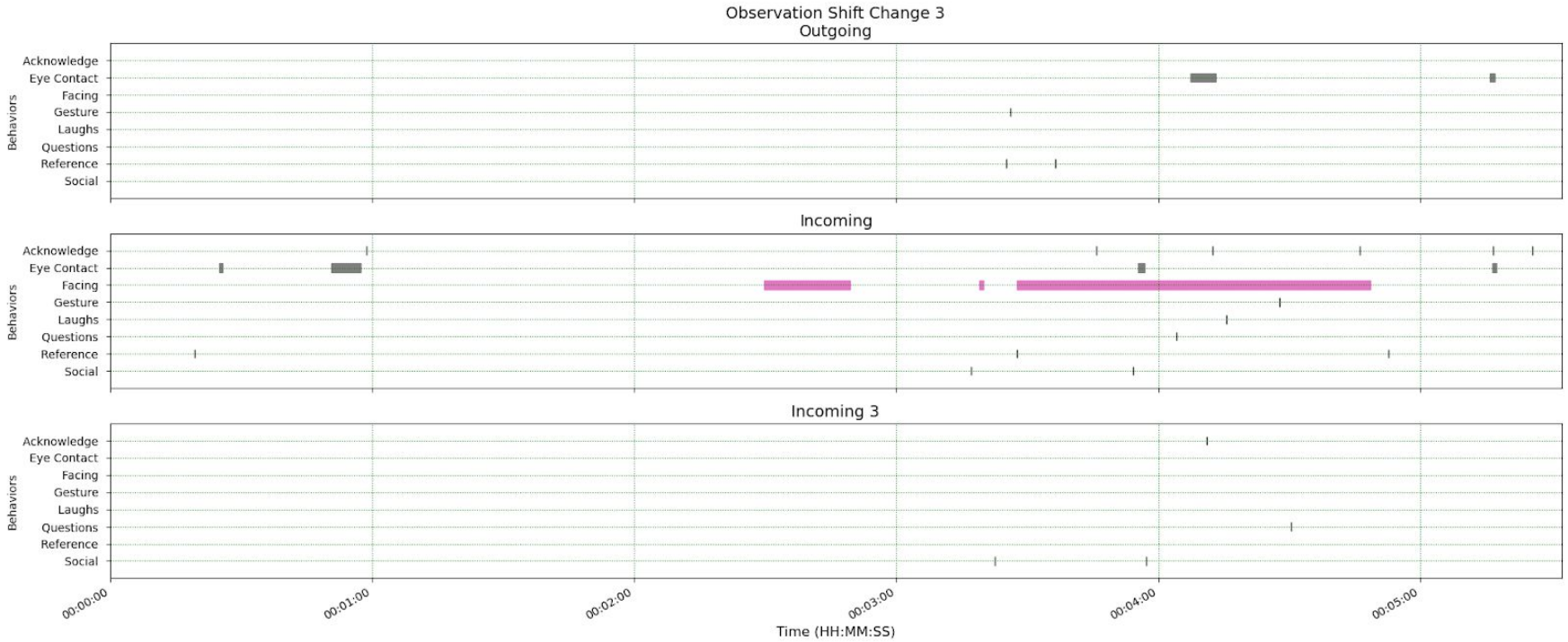
- not knowing why someone was upset with you because an earlier shift didn't tell the oncoming shift that they were expecting something.
- double-doing a dask (ex. making 2 dinners), possible med errors, not completing a task

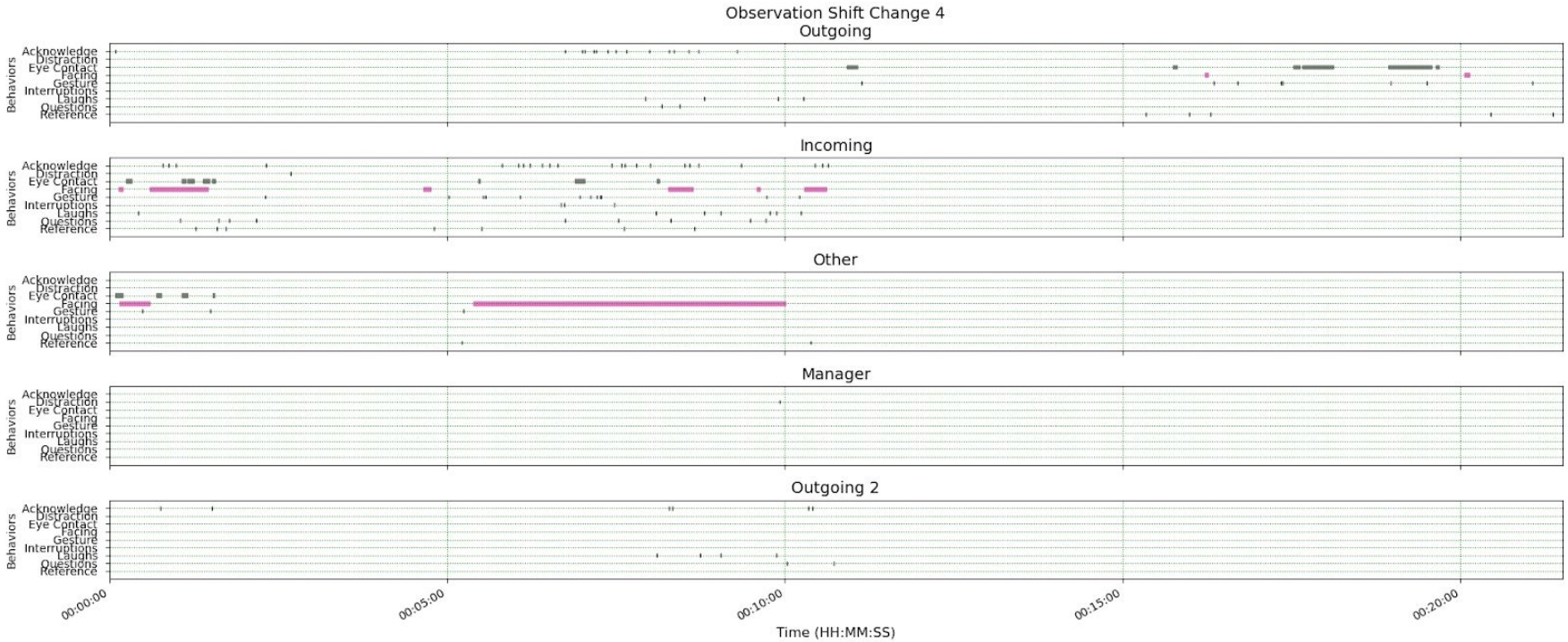
- mainly not knowing about meds being passed during shift change and not having the reliability or someone leaving the site and not telling staff about where they went.
- Medication errors, over feeding, miscommunication on where someone should be, in turn causing behaviors.
- Work that needed to be complete was not completed, support staff rushing, staff completing tasks that were already done, etc.
- substandard care - i.e. laundry, meals, general comfort, etc. are not supported because either misinformation or no information.
- missed medical appointments
- deep cleans and laundry not being done

Appendix C: BORIS Event Logs









Appendix D: Document - Level IBM Watson Tone Analysis JSON

Shift Change 1:


```
{
  "document_tone": {
    "tones": [
      {
        "score": 0.627927,
        "tone_id": "joy",
        "tone_name": "Joy"
      },
      {
        "score": 0.83592,
        "tone_id": "tentative",
        "tone_name": "Tentative"
      }
    ]
  }
}
```


Shift Change 2:


```
 {
  "document_tone": {
    "tones": [
      {
        "score": 0.657087,
        "tone_id": "joy",
        "tone_name": "Joy"
      }
    ]
  }
}
```


Shift Change 3:


```
        {
  "document_tone": {
    "tones": [
      {
        "score": 0.606617,
```

```
      "tone_id": "joy",
      "tone_name": "Joy"
    },
    {
      "score": 0.803444,
      "tone_id": "tentative",
      "tone_name": "Tentative"
    }
  ]
 }
}
```


Shift Change 4:


```
       {
 "document_tone": {
  "tones": [
    {
      "score": 0.584106,
      "tone_id": "joy",
      "tone_name": "Joy"
    },
    {
      "score": 0.505525,
      "tone_id": "sadness",
      "tone_name": "Sadness"
    },
    {
      "score": 0.946081,
      "tone_id": "tentative",
      "tone_name": "Tentative"
    }
  ]
 }
}
```