"I need to respond to this" – Contributions to group creativity in remote meetings with distractions

Alberta A Ansah
University of New Hampshire
Durham, USA
alberta.ansah@unh.edu

Yilun Xing
University of Washington
Seattle, USA
yilunx19@uw.edu

Amudha Varshini Kamaraj
University of Wisconsin - Madison
Madison, USA
akamaraj@wisc.edu

Diana Tosca
New York University
New York, USA
dt2290@nyu.edu

Linda Boyle
University of Washington
Seattle, USA
linda@uw.edu

Shamsi Iqbal
Microsoft Research
Redmond, USA
shamsi@microsoft.com

Andrew L Kun
University of New Hampshire
Durham, USA
andrew.kun@unh.edu

John D Lee
University of Wisconsin - Madison
Madison, USA
john.d.lee@wisc.edu

Michel Pahud
Microsoft Research
Redmond, USA
mpahud@microsoft.com

Orit Shaer
Wellesley College
Wellesley, USA
oshaer@wellesley.edu

ABSTRACT

Remote meetings have become more prevalent due to the COVID-19 pandemic and technology that facilitates remote work. There is limited research on the effect of remote meetings on group performance and the goal of this study is to identify how distractions affect the individual and group creativity in remote work meetings. A virtual study was conducted where groups of four people participated in divergent and convergent thinking tasks. One group member was assigned an additional non-meeting task while another was assigned as a scribe. Measures of creative performance (e.g., uniqueness of idea) of the distracted members and the group were analyzed. The results show that the distractee contributed (on average) less time and ideas when compared to monotaskers and those assigned as a scribe. The study highlights ways that remote meetings can facilitate creativity.

CCS CONCEPTS

• **Human-centered computing → Collaborative interaction**.

KEYWORDS

Collaboration, Tasks/Interruptions/Notification, Workplaces

CHIWORK '22, June 8–9, 2022, Durham, NH, USA

ACM ISBN 978-1-4503-9655-4/22/06…$15.00
https://doi.org/10.1145/3533406.3533411

ACM Reference Format:
Alberta A Ansah, Yilun Xing, Amudha Varshini Kamaraj, Diana Tosca, Linda Boyle, Shamsi Iqbal, Andrew L Kun, John D Lee, Michel Pahud, and Orit Shaer. 2022. "I need to respond to this" – Contributions to group creativity in remote meetings with distractions. In *2022 Symposium on Human-Computer Interaction for Work (CHIWORK '22), June 8–9, 2022, Durham, NH, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3533406.3533411

1 INTRODUCTION

The COVID-19 pandemic has forced many knowledge workers to transition from working in the office to working from home [21, 26, 34]. Several meeting platforms such as Zoom and Microsoft Teams became widely used for remote collaboration. For example, the meeting platform Zoom had about 10 million daily meeting participants in December 2019, but by April 2020, that number increased to over 300 million [8]. Similarly, Microsoft Teams had 13 million daily active users in July 2019 and 145 million in April 2021 [2].

The rise in the number of meetings has corresponded to a rise in multitasking during meetings [4, 12, 13, 22, 33]. Multitasking during meetings reflects a growing number of unplanned interruptions that have reduced the time people have for focused work [10]. In addition, longer meetings, meetings with many participants, and recurring meetings have encouraged passive participation in the meeting discussions and active participation in other meeting and non-meeting related activities [4]. For example, most meeting platforms allow individuals to post text, images, files, and links while others are in conversation mode during a meeting. While these activities are considered acceptable during a meeting [30], there are concerns that other non-meeting distractions can impact the flow of conversation, group creativity, and ability to connect socially. And unlike in-person meetings, workers can be multitasking without the other attendees noticing.

Metrics for a successful meeting depend on the meeting purpose. Success can be measured by the ability to transmit information among group members, ability to get a task activity accomplished, or the ability to foster creative thinking [31]. The goal of this study is to examine group performance and creativity in remote meetings. In particular, we examine the effect on individual and group performance when some meeting participants experience frequent interruptions. This study goal is achieved with remote meetings consisting of groups of four people that completed two sets of two tasks: one set depicted a more controlled setting while the other provided a more naturalistic (real world) example of group creativity. The two tasks included a divergent thinking and a convergent thinking task. Divergent thinking is used to explore multiple possible solutions or paths of action. Convergent thinking allows narrowing down solutions from many options to a few. Both types of thinking can be required for creative work [25]. To study the impact of multitasking on the group's performance, one person in the group is given an additional task, where they are asked to perform simultaneously without the knowledge of the other group members. The group and individual roles are examined in terms of how they generate ideas and converge on a decision. We also examined the impact of non-meeting tasks on the overall meeting performance. Our findings can be used to inform emerging practices for effective meetings and have implications for providing new experiences in remote meeting platforms.

2 RELATED WORK

2.1 Challenges of Remote Work

The COVID-19 pandemic has led to changes in the way people work, one of which is the increased prevalence of remote meetings as a means of communicating or holding work meetings [17]. Some workers and companies are interested in continuing remote work long-term while others are considering a more hybrid working style (part-time remote) [28, 29]. This possible change in people's long-term working style has fostered a growing body of studies on remote meetings. In one study [36] remote meetings were welcomed by individuals living alone who craved social contact but were considered intrusive for working parents given the blurring of work-life boundaries. Karl et al. [17] identified six challenges of videoconferencing from text mining 549 comments posted to a LinkedIn online discussion board "Chewing while on Zoom? Oh dear." Three of the challenges were tied to camera and microphone issues, two involved eating and meeting management issues and one dealt with work-from home issues. Videoconferencing provides more flexibility than face-to-face meetings, but also allows more back-to-back meetings, which can be mentally and physically exhausting [9]. In 2019, the online platform Statista [1] conducted a study with US participants aged 22 to 65. Out of 1,202 responses, the number of individuals that had more than 11 meetings per week rose from 3% to 15% after working remotely. The studies described in this section are all based on self-reported data from surveys, interview and social media comments, and little is known about the overall meeting performance, satisfaction, engagement and effectiveness of remote meetings. In this work we contribute to the field by presenting data from remote meetings.

2.2 Group Creativity and Collective Intelligence

Wescheler [37] defines collective intelligence as intelligence which "involves, or is in some way the result of, group rather than individual mental effort" (p. 904). For more than a century, it has been shown that the average assessment of a group in relation to a number of elements or a reality can be more accurate than estimates by experts [18]. This phenomenon has been studied throughout the 20th century, confirming that the average execution of a group in decision-making is higher than the execution of the subjects individually [19]. Wooley et al. [38] used the statistical approach developed for individual intelligence [7] to systematically measure the intelligence of groups. They found a converging evidence of a general collective intelligence factor that explains a group's performance on various tasks. A "c factor" was generated as a measure for the groups' collective intelligence, which is not strongly correlated with the average or maximum individual intelligence of group members but correlates with the average social sensitivity of group members.

In terms of what inhibits and facilitates collective intelligence, a design goal to support distributed meetings could enhance the collective intelligence of a group by increasing group awareness. McGrath [23] proposed that most group tasks can be classified into categories that reflect the four basic processes: generate, choose, negotiate, and execute. Creativity tasks include brainstorming or planning tasks which requires idea generation. Intellective or problem-solving tasks require choosing correct answers, and judgement or decision-making tasks require reaching consensus on a preferred answer. Resolving conflicting viewpoints or conflicting interests require negotiation. Execute tasks are those requiring physical movement, coordination or dexterity, such as psycho-motor tasks and athletic contests.

Studies have addressed the collective intelligence in meetings such as in Wooley's paper [38]. However, few studies investigated the collective intelligence in a remote setting. COVID-19, and the associated changes in people's meeting habits, demands research into collective intelligence in remote work settings to help knowledge workers meet the challenges of a new remote or hybrid working mode. The collective intelligence for different types of tasks is also worth examination. The taxonomy of group tasks proposed by McGrath [23] can guide experimental task selection. Our work contributes to the exploration of group intelligence by assessing group creativity for remote meetings and for different types of tasks.

2.3 Interruptions and Task switching

Knowledge workers are often governed by multiple tasks and activities that must be retained in working memory to perform, often in parallel or in rapid succession; this is called multitasking [4, 6]. Prospective memory failure is shown to be a significant problem for knowledge workers [5, 27, 32, 35] and interruptions of tasks are one of the most frequently cited reasons for prospective memory failures during the work day [27]. External interruptions can cause the knowledge workers to enter into a "chain" of distraction where stages of preparation, diversion, resumption, and recovery can describe the time away from an on going task [14].

Multitasking occurs frequently in remote meetings. The work-from-home regulation due to the COVID situation forced knowledge workers to work at home. This new remote working mode has already received many complaints on interruptions during remote meetings from knowledge workers [9, 17, 36]. A large-scale analysis of multitasking behavior during remote meetings showed that 30% of meetings involved email multitasking [4].

Previous studies were conducted largely with in-person meetings with some participants being distracted [12, 13]. Iqbal et al. [13] investigated the perceptions and preferences regarding the use of computing devices for potentially extraneous tasks in settings such as viewing presentations at seminars. Results showed that audience members who use devices believe that they are missing content being presented and are concerned about social costs. Karl et al. [12] compared the performance of students allowed to use laptops to browse, search or social computing during a lecture and those who are not. They found that students in the laptop condition remembered less lecture content.

Few studies have investigated multitasking in remote meetings with controlled experiments. A online study showed that the technology used for multitasking has a significant effect on others' assumptions of what secondary activity the multitasker is likely engaged in, and that this assumed activity in turn affects evaluations of politeness and appropriateness [22]. They also show that different layouts of the video conferencing UI can affect perception of engagement in the meeting and ratings of polite or impolite behavior. However, this study focused on single participants engaged in one-on-one meetings with a researcher. These meetings are somewhat artificial and are unlikely to mimic the consequence of multitasking on the organic development of conversations in actual meetings. In this work we explored tasks that involve four remote participants – the researchers did not participate in the completing tasks.

Multitasking can involve activities that are unrelated to the meeting (e.g., text or email communications) and related (e.g., notes taking) [13]. Tasks related to the meeting can help attendees to concentrate on the group task, while the tasks unrelated to the meeting does the opposite. To improve the performance and effectiveness of a remote meeting, both types of multitasking merit investigation. Our work explores the effects of one group member being engaged in a distracting task that is unrelated to the group task.

3 METHOD

This study examined the effect of a group's collective performance in a remote meeting when one group member is engaged in a non-meeting-related task. For the purpose of the study, each remote meeting consisted of groups of four participants from different geographical areas. Participants were put into groups based on scheduling and availability. Participants did not know who they would be paired with.

3.1 Participants

There were 23 groups that participated in the study with a total of 93 participants. Due to scheduling conflicts, technical issues and no-shows, some groups had only 3 participants instead of 4. In this paper we present data for the 12 groups that had all four participants present (n=48).

Participants from the 12 groups ranged in age from 18 to 64 with 52.1% within the ages 25–34. 68.8% of the participants identified as women, 22.9% identified as men and 8.3% of the participants either identified as non- binary, transgender or preferred not to disclose.Each group comprised of at least two individuals who identified as women. Since the study was conducted virtually, the inclusion criteria required that participants have prior experience with virtual meetings and that they have access to a laptop with video calling capabilities. The virtual meeting took place using the Microsoft Teams application. To recruit participants, we posted flyers on social media (Facebook, Twitter, LinkedIn) and contacted knowledge workers in various industries. Interested participants first completed a screening survey. The screening criteria required willingness to participate in a recorded group meeting and be physically present in the US. Participants who met the criteria were sent an email with more details about the study and a link to schedule the meeting. The participation time was approximately one hour, and participants were compensated with gift cards valued at US $25. The research was approved by the institutional review board (IRB) at the University of New Hampshire.

3.2 Tasks

All groups partook in the preliminary and main study tasks, which were designed to assess the group's creativity. The preliminary tasks included the Remote Associates Test (RAT – task 1) and the Alternate Uses Task (AUT – task 2). RAT was used to assess the group's convergent thinking and AUT was designed to assess divergent thinking [24]. This paper focuses on analyzing only the main study tasks. Future work will explore the analysis of the preliminary tasks.

3.2.1 Task 1 (Convergent thinking task — RAT). Participants were given three words and asked to provide the word that connects the three. For example, given the three word prompts "cottage", "swiss", and "cake", the word that connects them is "cheese". During the task, facilitators provided prompts and the group was tasked with finding the connecting word. The goal of the task was to get as many correct answers in three minutes. Participants were given one practice round.

3.2.2 Task 2 (Divergent thinking task — AUT). The goal of the task was to provide as many alternative uses of a given object. For example, alternative uses of the object 'brick' could be to used as a doorstop, or to prop something up. Participants had one practice round and two rounds for this task each timed for two minutes. The object "brick" was used for the practice round. The objects "shoe" and "paperclip" were used during the real task.

3.2.3 Task 3 (Real world divergent thinking task). For the real world divergent task, all groups were first asked to assume that they were a philanthropic committee that had a US $100,000 to give to worthy causes. The goal for each group was to generate a list of philanthropic causes which the $100,000 could benefit. The groups were instructed to generate a diverse set of causes within 5 minutes. By asking groups to generate diverse set of causes, the task aimed to measure divergent thinking in groups for real-world problems.

3.2.4 *Task 4 (Real world convergent thinking task).*

The final task required groups to select three causes from the list of charitable causes generated in the divergent task and asked groups to reach a consensus on the amount of funds to be allocated to each cause. The $100,000 was divided into three funds, one large fund of $70,000, a medium sized fund of $25,000, and small fund of $5000. Groups were given 15 minutes to arrive at a consensus. Here, the goal was to measure convergent thinking in groups in a real-world setting by having groups discuss and arrive at a consensus.

3.2.5 *Phone task.*

The phone task was only performed by the designated distractees during the group tasks (tasks 1–4). The task was to find the sum of two numbers. Participants initiated the task by texting "Start" to a US-based cell number. The system responds with a prompt "Please wait for the next text", followed by the question "What is the sum of these two numbers? a & b", where a and b are whole numbers between 0 and 100. The distractee was responsible for responding to the prompts accurately and as quickly as possible. Both of these goals were equally important. After the participant provides an answer the next question is given. After five rounds of questions the distractee receives the prompt "Thanks for completing the task. To start again, text "Start". This was done to ensure that the phone task could be completed regardless of the duration of each creativity task. Distractees were required to do the phone task as long as the main task was going on. Phones were adopted to deliver the distraction task so that the medium for the remote meeting was separate from the medium for distraction. Also, simple arithmetic was selected to ensure that the distraction tasks were manageable for anyone who signed up for the study.

3.3 Participant Roles

Once the virtual meeting was scheduled, participants were assigned one of three roles based on the tasks they would perform during the study: scribe, distractee, or monotasker.

3.3.1 *Scribe.*

The role of the scribe was to record the answers from the group by typing them into the chat. Scribes were assigned their roles prior to the start of each task. There were two scribes in each group- one scribe for the preliminary tasks (1 and 2) and the other for the main study tasks: divergent (task 3) and convergent (task 4).

3.3.2 *Distractee.*

The distractee is the group member who was assigned a non-meeting task during the study. The identity of the distractee was not disclosed to the group. Those who were assigned the distractee role were informed via email prior to the study. The email included detailed instructions on how to perform the non-meeting task and requested that distractees not disclose their role to other group members. There was one distractee assigned for the preliminary tasks and another for the main study tasks.

3.3.3 *Monotasker.*

There were two participants assigned as monotasker. These individuals were not given any additional tasks. The expectation of the monotaskers was that they would be fully engaged in the primary tasks that was assigned to the group.

3.4 Procedure

Prior to the virtual study, each group member was sent an email providing information on the time of the meeting and instructions for accessing the online meeting room. Additionally, two of the four members were randomly selected to play the role of distractee. These members were notified via email of their role along with instructions to perform a non-meeting task (the phone task) as inconspicuously as possible when asked to during the meeting.

Before beginning the tasks, all group members were asked to provide a brief verbal introduction of themselves and asked to state how often they participate in remote meetings. The virtual meeting was split in to four segments with each segment having the group perform a different task that tests the group's creative performance. As noted earlier, each group had two assigned distractees, one for the preliminary tasks and another for the main real world study tasks.

Each task also had an assigned scribe whose role was to collect the answers from the group and enter it in the chat interface. The scribes were assigned the same way as the distractees. At the beginning of each task, the distractee is reminded (without revealing their role to the rest of the group) to start the non-meeting task (phone task) that was detailed in the email that was sent to them. After each task, group members are required to complete a survey before proceeding to the next task.

Each meeting had 2–3 researchers whose cameras were turned off and muted as needed. Their roles were to moderate and observe the study, and keep track of the task times.

3.5 Questionnaires

3.5.1 *NASA Task Load Index (TLX).*

After each task, participants completed a NASA TLX questionnaire [11]. This is a subjective indication of the mental workload for the given task. It is rated along six dimensions which are the (1) mental demand, (2) physical demand, (3) temporal demand, (4) effort, (5) performance, and (6) frustration level. Participants assessed workload on a 7-point scale from very low to very high.

3.5.2 *Personality Questions.*

Participants were required to complete a personality questionnaire to understand the types of people in the group. The questionnaire was completed prior to task 1.

3.6 Performance and creativity measures

The video recordings and chat files from Microsoft Teams were examined for the following performance measures:

- Uniqueness of the three selected causes (convergent)
- Task completion time (convergent)
- Number of causes each group member contributed (divergent)
- Percent of time each group member contributed to the discussion (divergent and convergent)
- Number of new ideas, following idea, and revisited ideas

3.7 Phone task

The following measures were recorded for the phone task:

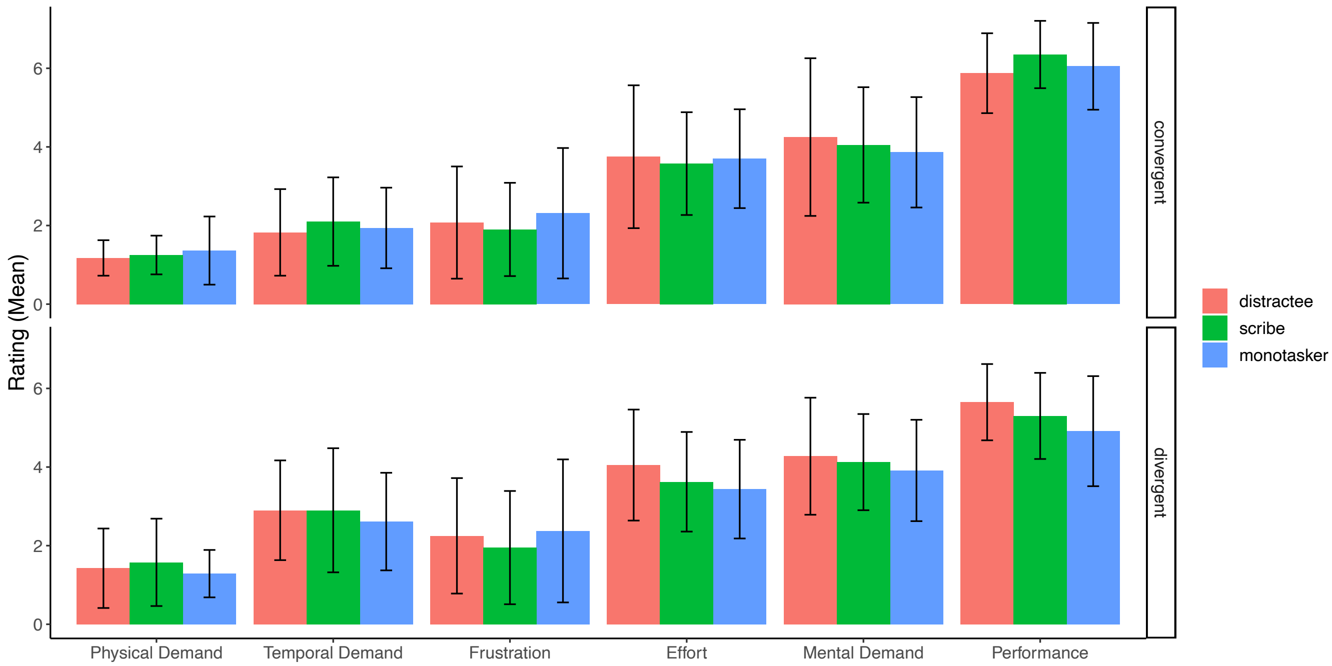

Figure 1: NASA TLX ratings across six component dimensions.

- Response time: This is calculated as the time period between when the question was sent and when distractees responded with an answer.
- Response accuracy: The accuracy of the distractee's response was marked as correct if the answer provided by the distractee is the correct sum of the two numbers.
- Number of responses.

4 RESULTS

The analysis presented in this section focuses only on the real-world tasks (tasks 3 and 4). The RAT and AUT tasks functioned as ice-breaking tasks and were used to give group members an opportunity to get acquainted with each other. Means and standard deviations (SD) are reported where appropriate. Analysis of variance (ANOVA) was conducted to examine differences in roles (F-test) and pairs were examined using t-tests. All analysis was conducted in R version 4.03. All trends (increasing and decreasing) are reported and significance was assessed at $\alpha = 0.05$.

4.1 NASA TLX

The NASA TLX unweighted score provides a measure of workload as experienced by each group member with the six dimensions contributing to the overall workload. Overall scores show that in the divergent task, the distractee scored the highest (mean=3.4 and SD = 0.7) followed by the scribe (mean=3.3, SD = 0.7) and the monotaskers (mean=3.1, SD = 0.6). In the convergent task, the scribe (mean=3.2, SD = 0.6) and the monotaskers (mean=3.2, SD = 0.6) scored the same while the distractee scored lower (mean=3.2, SD = 0.9). While there was no significant effect, the results suggest distractees experienced the highest workload in the divergent thinking task and experienced the lowest workload in the convergent thinking task.

In examining the dimensions of the NASA TLX (see Figure 1), minimum physical demand was reported (mean=1.3, SD=0.8) compared to all other dimensions for both the divergent and convergent task. For the temporal demand, group members report significantly higher temporal demand ($t(89.6)=3.31$, $p = 0.0013$) for the divergent task when compared to the convergent task. The scores from the frustration dimension show that monotaskers report a higher frustration (mean=2.4, SD=1.7) in the divergent and convergent thinking task compared to the distractee and the scribe. In the mental demand dimension the scribes (mean=4.3, SD=1.4) and distractees (mean=4.3, SD=1.5) reported a higher mental demand than the monotasker for both the divergent and convergent tasks. Finally, significantly lower performance demands for the divergent task was reported when compared to the convergent task ($t(90.4)=-3.816$, $p < 0.001$). No significant results were detected between each pair of roles for any NASA TLX dimension or the overall score.

4.2 Group Performance

4.2.1 Real world divergent thinking task . In the divergent thinking task, groups were instructed to create a list of philanthropic causes and were given 5 minutes to accomplish the task. Group task performance is assessed using the number of causes generated. To assess the performance across the group member roles, the proportion of causes generated across each role and the proportion of time group members were engaged during the task is estimated.

The number of causes generated by each group in descending order is shown in a stacked bar plot in Figure 2(a) (mean = 21.6 and SD = 6.8). The stacks in the plot indicate the contribution of each

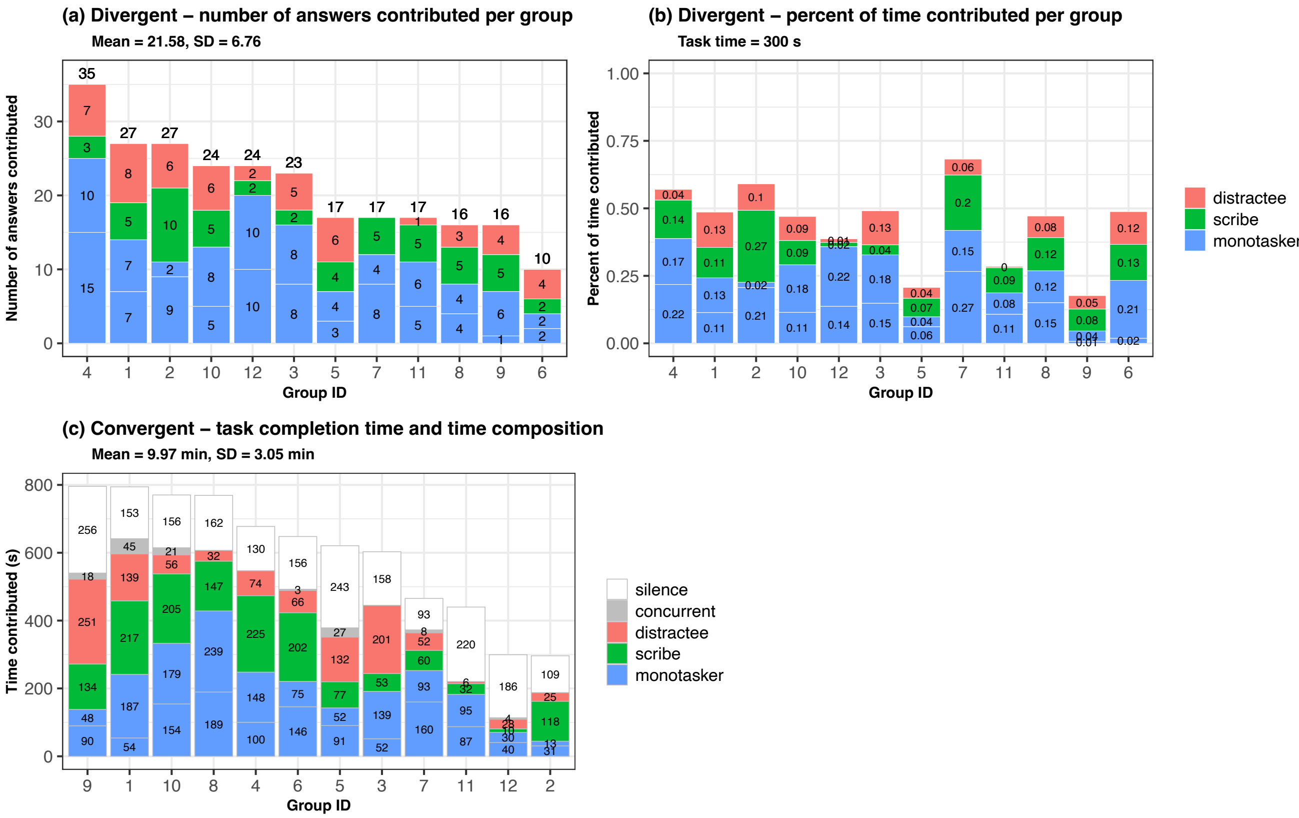

Figure 2: Task performance per group stacked by different participant roles.

role. The percentage of time contributed stacked by each role is shown in Figure 2(b). From 2(b) it is observed that most groups only spent half their time engaged in discussion and the other half of the time silent. Recall that for this task, the time allotted was the same across groups which is 5 minutes. The proportion of time when group members were engaged in the task is plotted in Figure 2(b). Thus, the time when group members were silent (i.e., not engaged in discussion) equals one minus the value of each bar.

Based on the number of causes generated shown in Figure 2(a) and the proportion of time members were engaged in conversation (shown in 2(b)), it appears that groups which generated more causes tend to engage more in the conversation.

The distribution of the philanthropic causes generated per role is shown in Figure 3(a). This shows that the scribe (mean = 21.8%, SD = 8.9%) generated fewer causes than the distractee (mean = 22.1%, SD = 9.9%). The distractees generated fewer causes than the monotaskers (mean = 28.5%, SD = 10.7%) during the divergent task ($F(2,44) = 2.481$, $p = 0.095$).

The distribution of the percent of time contributed by each role is shown in Figure 3(b) . This shows that the distractee (mean = 7.1%, SD = 4.3%) contributed less than the scribe (mean = 12.2%, SD = 7.2%). The scribe contributed less time than the monotaskers (mean = 12.5%, SD = 7.2%) ($F(2,45) = 2.890$, $p = 0.066$).

The distribution of the percent of words contributed by each role is shown in Figure 3(c). This shows that the distractee (mean = 14.4%, SD = 9.6%) contributed a lower percentage of words than the monotaskers (mean = 24.5%, SD = 12.4%). The monotaskers appears to have contributed significantly less than the scribe (mean = 26.5%, SD = 13.7%) ($F(2,45) = 3.686$, $p = 0.033$). Results from pairwise t-test shows that the distractee contributed significantly fewer words than the other two roles (distractee vs. monotasker: $t(32.13) = -2.769$, $p = 0.009$; distractee vs. scribe: $t(19.89) = -2.155$, $p = 0.044$; monotasker vs. scribe: $t(24.94) = 0.127$, $p = 0.900$).

4.2.2 Real world convergent thinking task. In the convergent thinking task, groups narrowed down the list of philanthropic causes to three causes following a group discussion. The time allotted for this task was 15 minutes and groups were allowed to conclude the discussion earlier if a consensus was achieved. Group performance for this task was assessed based on the task completion time. The performance across each group member role is assessed using only the proportion of time contributed and the number of words spoken by each group member role. Note that the performance measure used in the divergent thinking task (the number of causes generated) is not applicable here as the goal was to only select causes during this task.

The task completion time (mean = 10.0 minutes and SD = 3.1 minutes) for the convergent task for each group in descending order

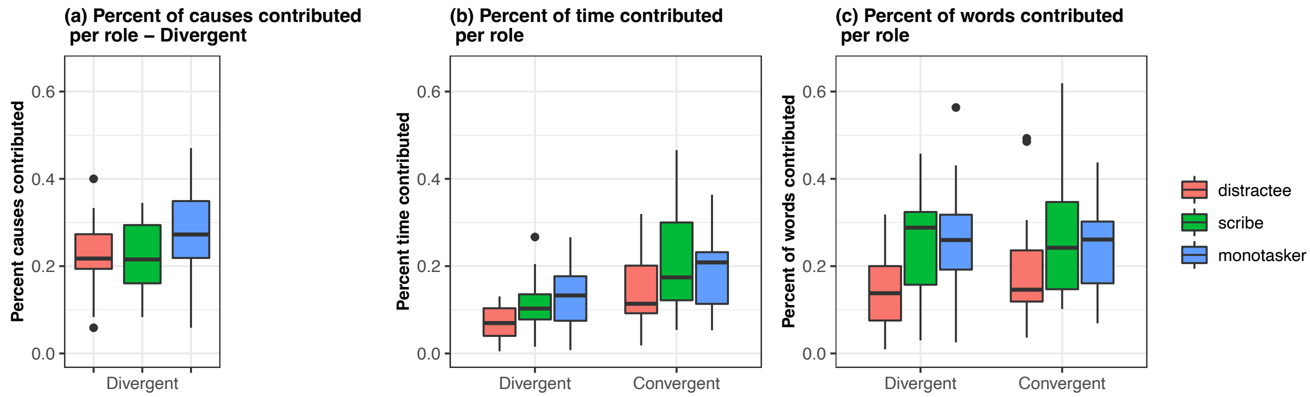

Figure 3: Group performance in divergent and convergent thinking task based on group member role. (a) The percent of causes generated across roles in the divergent thinking task. The convergent thinking task was a consensus task which selected causes and thus answers contributed is not applicable. (b) The percent of time contributed across roles for the divergent and convergent task.(c) The percent of words spoken across roles for the divergent and convergent thinking task.

is displayed in Figure 2(c). The groups where the scribe contributed to the task appear to take more time to reach consensus.

The distribution of time contributed per role during the convergent task is shown in Figure 3(b). On average, the distractee (mean = 14.6%, SD = 9.8%) contributed less time than monotaskers (mean = 18.6%, SD = 8.4%) who in turn contributed less time than the scribe (mean = 20.3%, SD = 12.2%). However, no significant differences were observed ($F(2,45) = 1.074$, $p = 0.350$).

The distribution of words contributed per role is shown in Figure 3(c). This shows that the distractee (mean = 20.2%, SD = 15.4%) contributed a lower percentage of words than monotaskers (mean = 24.5%, SD = 10.7%) and the scribe (mean = 26.2%, SD = 15.3%). However, no significant differences were observed ($F(2,45) = 0.688$, $p = 0.508$).

4.2.3 Quality of group response in divergent thinking task. Thus far, performance during the divergent task was assessed using the quantity of responses generated and the contribution per group member role. Given the nature of the task, another performance measure is the quality of the responses generated during the task. Since group members were instructed to generate philanthropic causes with little overlap, the uniqueness of response is one way to analyze the responses. Another method of assessment that is adopted here is the manner in which group members interacted to generate each philanthropic cause. For instance, when one member suggests a philanthropic cause related to healthcare, another member may follow up with a either a similar or dissimilar philanthropic cause. Thus, the initial suggestion may influence the following suggestion or not. Analyzing such patterns may provide insight into the factors that differentiate group performance. Here a three step process was followed to analyze the quality of responses shown in Figure 4.

Recall that the responses for this task was a list of philanthropic causes that the group members deemed worthy of financial support. Groups were also instructed to generate philanthropic causes that were as diverse as possible. To analyze these responses, the causes generated from each group were first gathered in the order in which they were suggested. Next, since the causes were asked to be as diverse as possible, the causes were labeled under philanthropic categories to understand how many unique philanthropic categories were present. A publicly available philanthropy taxonomy was used to identify the philanthropic categories that the philanthropic causes belonged to [3]. Three researchers individually labeled the data and convened to resolve any disagreement in the labeling. The uniqueness in the list of philanthropic causes was then assessed based on the number of unique philanthropic categories generated.

The conversational patterns that lead to the generation of causes was then assessed. This was accomplished by analyzing the order in which the philanthropic causes were generated in three ways. The responses were classified as one of three types of ideas:

- *New idea* is the first suggestion of a philanthropic category. Note that the instances of new ideas also indicates the uniqueness of the responses.
- *Following idea* refers to a philanthropic category of a suggested cause being in the same philanthropic category as the previous cause. For example, assume that the first suggestion from the group was 'cancer research' and the second suggestion from the group was 'diabetes'. Based on the philanthropic taxonomy, both these causes are categorized under the philanthropic category of 'health'.
- *Revisited idea* is an idea that is neither new nor a following idea. For example, let's say that health-related causes were previously discussed (e.g., diabetes, cancer research). Another person may then move to a new idea such as 'Black Lives Matter' which could be placed in philanthropy causes – human rights. The next person may then say 'medical research' which was a philanthropic cause already discussed. Hence, 'medical research' would be considered a revisited idea as it is not new and it did not follow from 'Black Lives Matter'.

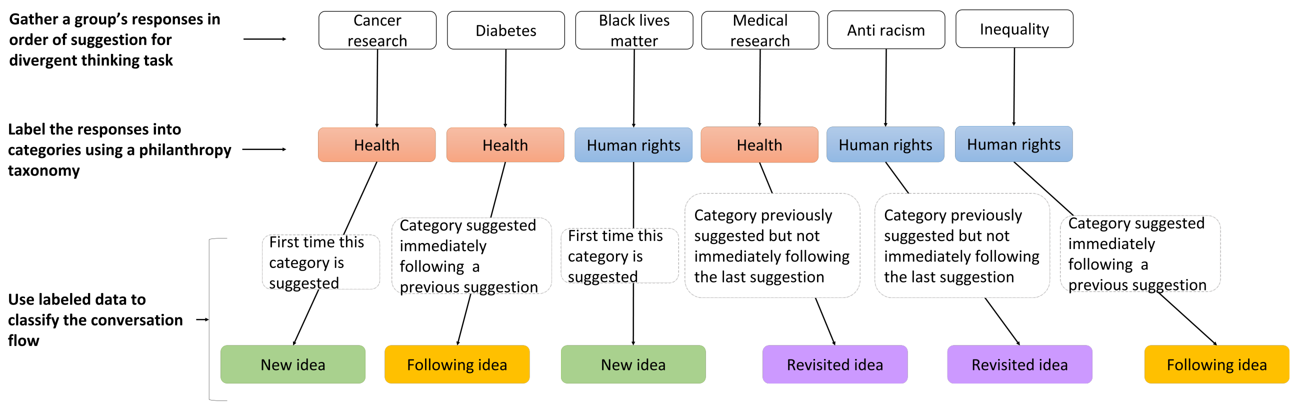

Figure 4: Data analysis procedure for analyzing quality for group responses from divergent thinking task

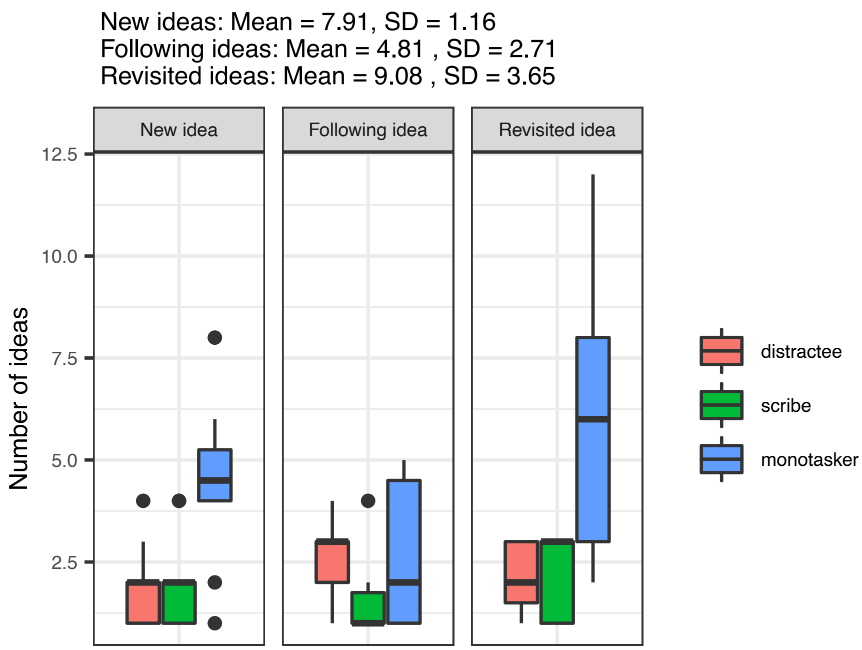

Figure 5: The contribution of each group member role to the three types of ideas generated: (1) New idea, (2) Following idea, and (3) Revisited idea. The facets are arranged in order from the group that has the highest number of causes generated (Group 4 with 35 causes) to the group with the lowest number of ideas (Group 6 with 10 ideas). The number of new ideas is also indicative of the unique ideas generated per group.

The contribution of each group member role was then examined in the context of these three idea types. Figure 5 shows the contribution across all groups in a stacked bar chart. Groups tend to revisit ideas (mean = 9.1, SD = 3.7) more often than generate new ideas (mean = 7.9, SD = 1.2) or following ideas (mean = 4.8, SD = 2.7).

4.3 Distractee performance

The distractee participated in the group task while also interleaving the phone task as part of their participation in the study. We report results on the distractee performance for both the divergent (task 3) and convergent task (task 4).

4.3.1 Phone task performance. Distractees completed an average of 7.1 questions in their phone task during the real world divergent task and 11.8 questions during the real world convergent task with similar mean response times and accuracy, see Table 1.

The survey question that was completed at the end of the convergent task showed that about 50% of distractees (7/12) reported that their performance in the convergent task reduced because of the phone task while 33.3% (4/12) reported that the phone task

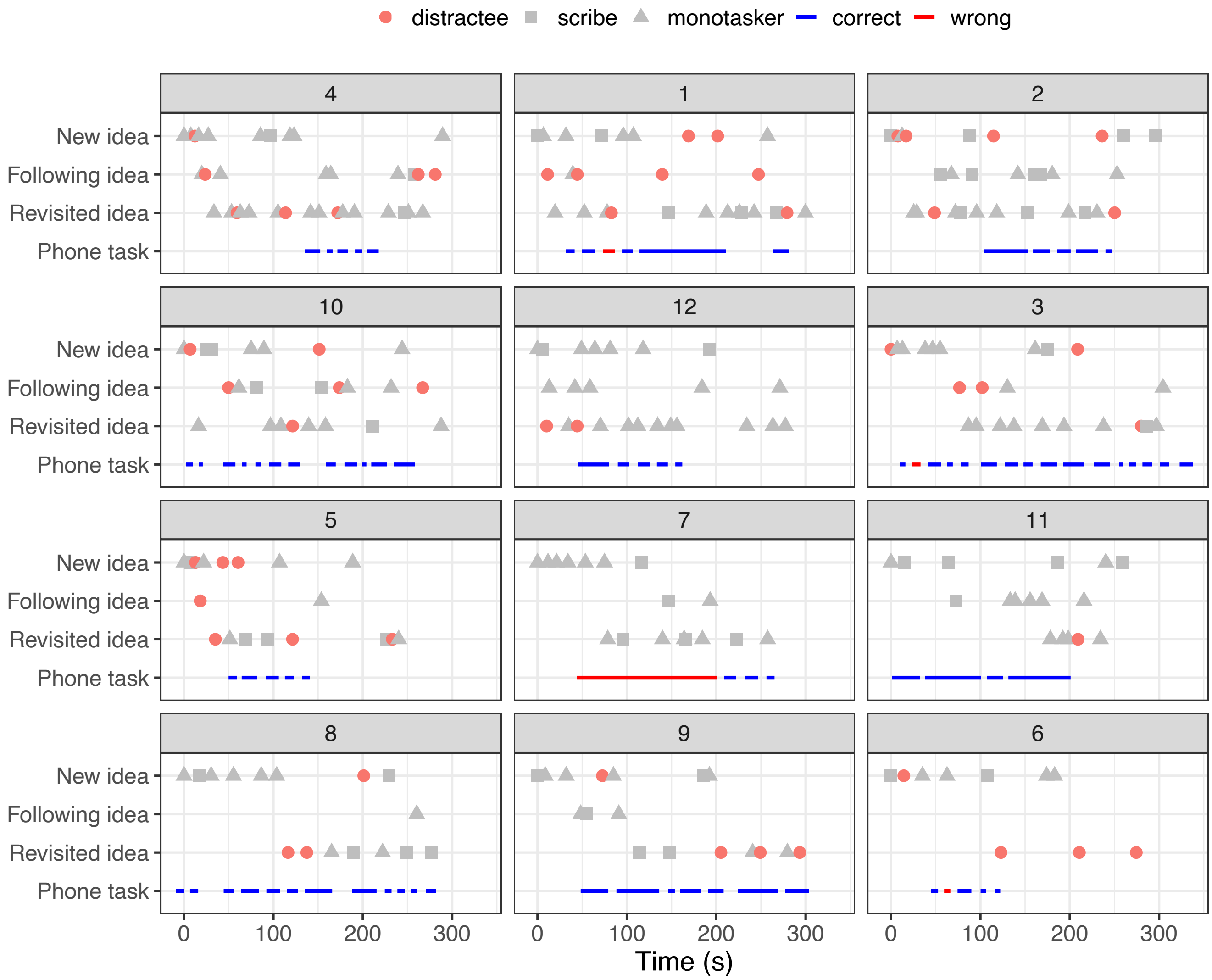

Figure 6: Type of idea contribution in the divergent task from all group members with distractee's contribution for each idea type highlighted. The phone task segments show when the participant was engaged in the phone task and the correctness of their response. The groups are ordered according to divergent task performance which is the same as Figure 2(a),2(b).

Table 1: Distractee performance on phone task questions for divergent and convergent task

	Response count Mean	Response count SD	Response time (s) Mean	Response time (s) SD	Accuracy Mean	Accuracy SD
Divergent task	7.08	3.82	21.49	13.44	94.34	9.33
Convergent task	11.75	8.17	20.91	6.83	96.95	5.15

neither improved nor reduced their performance in the convergent task. The survey question was only asked at the end of the convergent task to avoid any influence the response might have on the distractee performance for the convergent task (task 4).

The other three members of the group were asked if they could identify the distractee for the convergent task (task 4). 69% (25/36) participants chose to identify the distractee while 25% (9/36) responded they did not know which group member was the distractee. Two participants (a monotasker and a scribe) incorrectly identified themselves as the distractee for the task. Of the 25 participants who opted to identify the distractee, 16 participants (64%) correctly identified the distractee for the convergent task (task 4).

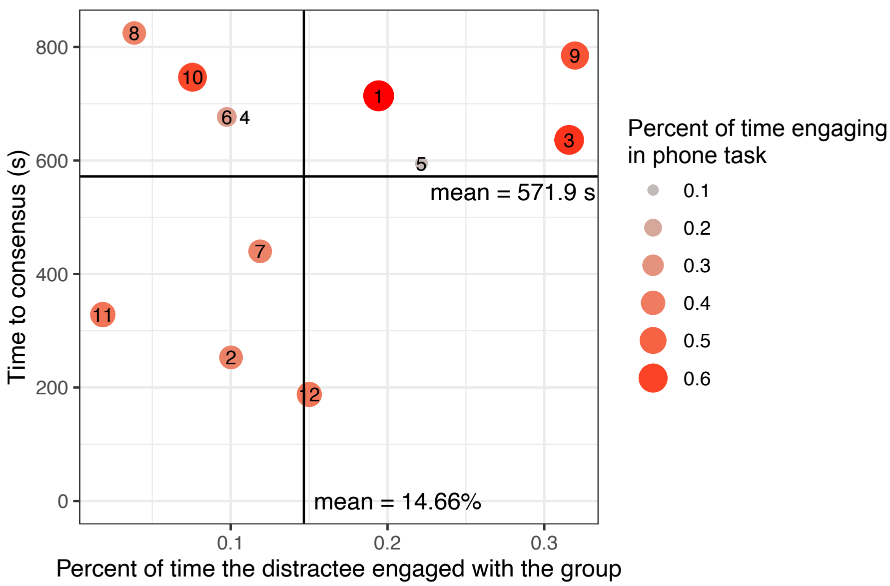

Figure 7: Relation between the time to consensus in the convergent task, the distractee's engagement in the phone task, and the distractee engagement with the group.

4.3.2 *How the distractee's engagement with the phone task impacted their participation in the group task.*

We analyzed the distractee's level of engagement in the phone task and their participation in the main tasks(divergent and convergent tasks). For the divergent task, the number of ideas the distractee contributed was correlated with the level of engagement in the phone task (see Figure 6). Our findings showed that 75% of distractees contributed more answers when not actively engaged in the distraction task. 70.8% of the new ideas and 81.3% of the following ideas contributed by distractees occurred when the distractee was not actively engaged in the phone task.

In terms of the convergent task, the relationship between the time to conclusion and the distractee's engagement in both the main task and the phone task is illustrated in Figure 7. No strong correlation was observed between time to conclusion and the other two measures.

5 DISCUSSION

5.1 Distractee impact on group performance on divergent and convergent thinking tasks

For the divergent thinking tasks, group creativity is assessed based on the number of items the groups generate. Groups that generated more items were those where the distractee was more engaged with the group. Considering the uniqueness of ideas generated, distractees suggested fewer new ideas (17.9%) relative to the monotasker (56.8%) and the scribe (25.3%). The suggestions of the following and revisited idea appear to boost the number of items that the group generates as seen in the top six groups (see Figure 6).

Results indicate that distractee engagement with the group is less than the monotasker and scribe as shown in the number of items contributed and the number of words spoken by the distractee in the divergent task (see Fig6). However, no significant effect was observed.

From the results of the convergent thinking task, although there was no significant relationship between the time to consensus and involvement of the distractee, an increased level of engagement from the distractee did not seem to correlate with the time to reach consensus (Figure 7). On the other hand, the results suggest that when the distractee was more engaged in the phone task, the group took longer to reach consensus.

In remote meetings, it is fairly difficult to tell when an individual is distracted. The survey results show that even though in most cases, it is difficult to tell if an individual is multitasking, some behaviors and actions can be an indicator. Group members were able to guess who the distractee was based on observations of behavioral difference relative to other participants and behavioral difference between tasks. For behavioral differences between tasks, the participant's behavior was compared to a previous task. For example, one participant mentions "He was more quiet that usual. And he didn't appear as involved as he was in previous tasks." For behavioral differences relative to other participants, comparisons were made in terms of engagements between members in the group during the task. For example, a participant mentioned "she wasn't contributing as much as the other people". Of note is that engagement was assessed either using the participants gaze towards or away from the camera or based on the verbal contribution to the discussion. For example,a participant stated "Looked down a couple times so I assumed he was using his phone" and another participant also mentioned "They seemed less vocally active".

These indicators of distraction can influence the design of remote meetings such that strategies to facilitate the meeting can be adopted when distraction indicators are identified.

Our results show that group creativity is affected by distractees engagement in non-meeting tasks. Reengaging these distracted group members could enhance the group creativity. Identifying and understanding these distraction indicators can help identify strategies that can re-engage group members.

5.2 Implications for future work from home

A study conducted by Chaitali and Shimul [16] suggests that multitasking can increase creativity through activation and cognitive flexibility acting in tandem. Often the multitasking activities that are examined within meeting settings are related more to the work and the ability to be more productive [30]. Our study focused on non-meeting related distractions and the ability to contribute to new ideas.

Our data provides preliminary evidence that despite being engaged in a non-meeting task, the distractees were able to contribute to the group activities. While in some cases their engagement with the other task may have undermined group performance (i.e., time to reach consensus), in other cases, the distractee was able to strategize and plan around how they could interleave both tasks to achieve acceptable performance on both. These results align with findings from Mark et al. [20] who showed how users compensated for the negative effects of interruptions by increasing their task speed once they returned from an interruption.

While studies have shown that those that are distracted by non-meeting related activities can have worse performance, this study showed that those engaged in meeting-related activities (e.g., the scribe) also suffered worse performance in terms of conversation pace and interactions. Hence, future studies should consider ways that can enhance multi-tasker performance.

These findings have broader implications for the future of remote work, where people can work from various locations. This could include environments where attention is limited (e.g., in a shared work space with other family members), when a secondary task is necessary (e.g., interleaving care-giving responsibilities during a meeting), or when a meeting participant have unplanned periods of focused and unfocused times. The data from this study was examined in terms of uniqueness, accuracy, and percent contribution. Another important consideration is the diversity and similarity of causes that were generated. This may be impacted by how attentive or distracted the individuals or group may be. Our findings suggest that divided attention participants are still able to contribute and adopt strategies to maximize their participation within the constraints of their situation. Future work should explore ways that companies and technology can better support such participants. For example, a system could potentially 'nudge' them at opportune moments to encourage participation, provide support for catching up with content missed due to the secondary interactions, and providing awareness of the limited attention to other meeting participants [15].

6 LIMITATION

This paper presents the analysis of 12 groups. These groups were chosen on the basis of team size to ensure that there were 4 team members in each group. While this limits the sample size of the current study, the remaining groups with different team sizes will be explored in future studies to examine the effect of team size on team performance.

7 CONCLUSIONS

Our research provided some new insights on the impact of distracted participants on group performance and creativity during remote meetings. Many companies are adopting hybrid meeting models, which include workers in various locations, such as home, office, factory, car, and sporting events. As working environments continue to evolve, this begins the examination of distractions in other settings. For example, many studies show that distractions undermine driving performance. However, as we move forward with semi or fully autonomous vehicles, the human operator may feel the need to prioritize productivity over safety; particularly if have inappropriate trust in the vehicle automation over time. Future studies can therefore extend the research of remote meetings to different worker settings—including the car—so we can better assess how well people can safely switch between tasks. One way to facilitate task switching is through feedback to all group members on the attention state of participants. For instance, whether the driver is attending to a driving maneuver, such a crossing an intersection with multiple pedestrians, or has cruise control engaged on the freeway. Our findings show that divided attention participants are still able to contribute in a meaningful way. Hence, by developing solutions that support multi-tasking, remote meetings tools can be designed to enhance the overall creativity of a group.

ACKNOWLEDGMENTS

This work was in part supported by NSF grants CMMI- (University of New Hampshire: 1840085; Wellesley College: 1840031; University of Washington: 1839666; University of Wisconsin-Madison: 1839484) Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSF.

REFERENCES

[1] Global Workplace Analytics. 2019. Number of meetings attended per week in the United States in 2019, by on-site and remote work. [Graph] in Statista. Retrieved July 27, 2021 from https://www.statista.com/statistics/1123269/number-meetings-attended-per-week-us-on-site-remote-work/

[2] Global Workplace Analytics. 2021. Number of daily active users (DAU) of Microsoft Teams worldwide as of April 2021. [Graph] in Statista. Retrieved Sep 03, 2021 from https://www.statista.com/statistics/1033742/worldwide-microsoft-teams-daily-and-monthly-users/

[3] Candid. 2021. Philanthropy classification system. https://taxonomy.candid.org/

[4] Hancheng Cao, Chia-Jung Lee, Shamsi Iqbal, Mary Czerwinski, Priscilla NY Wong, Sean Rintel, Brent Hecht, Jaime Teevan, and Longqi Yang. 2021. Large scale analysis of multitasking behavior during remote meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[5] Mary Czerwinski and Eric Horvitz. 2002. An investigation of memory for daily computing events. In *People and computers XVI-memorable yet invisible*. Springer, 229–245.

[6] Mary Czerwinski, Eric Horvitz, and Susan Wilhite. 2004. A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 175–182.

[7] Ian J Deary. 2000. *Looking down on human intelligence: From psychometrics to the brain*. Oxford University Press.

[8] Bob Evans. 2020. The Zoom Revolution: 10 Eye-Popping Stats from Tech's New Superstar. Cloud Wars. Retrieved Sep 03, 2021 from https://cloudwars.co/covid-19/zoom-quarter-10-eye-popping-stats-from-techs-new-superstar/

[9] Liz Fosslien and Mollie West Duffy. 2020. How to combat zoom fatigue. *Harvard Business Review* 29 (2020).

[10] Michael Gibbs, Friederike Mengel, and Christoph Siemroth. 2021. Work from Home & Productivity: Evidence from Personnel & Analytics Data on IT Professionals. *University of Chicago, Becker Friedman Institute for Economics Working Paper* 2021-56 (2021).

[11] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[12] Helene Hembrooke and Geri Gay. 2003. The laptop and the lecture: The effects of multitasking in learning environments. *Journal of computing in higher education* 15, 1 (2003), 46–64.

[13] Shamsi T Iqbal, Jonathan Grudin, and Eric Horvitz. 2011. Peripheral computing during presentations: perspectives on costs and preferences. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 891–894.

[14] Shamsi T Iqbal and Eric Horvitz. 2007. Disruption and recovery of computing tasks: field study, analysis, and directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 677–686.

[15] C. Janssen, Shamsi T. Iqbal, and Y. Ju. 2014. Sharing a driver's context with a caller via continuous audio cues to increase awareness about driver state. *Journal of experimental psychology. Applied* 20 3 (2014), 270–84.

[16] Chaitali Kapadia and Shimul Melwani. 2021. More tasks, more ideas: The positive spillover effects of multitasking on subsequent creativity. *Journal of Applied Psychology* 106, 4 (2021), 542.

[17] Katherine A Karl, Joy V Peluchette, and Navid Aghakhani. 2021. Virtual Work Meetings During the COVID-19 Pandemic: The Good, Bad, and Ugly. *Small Group Research* (2021), 10464964211015286.

[18] Charles Mackay. 1984. *Extraordinary Popular Delusions and the Madness of Crowds: Vol. 1-3*. e-artnow.

[19] T Malone and R Laubacher. 2009. Harnessing crowds: Mapping the genome of collective intelligence. *Retrieved from http://cci. ...* (2009).

[20] Gloria Mark, Daniela Gudith, and Ulrich Klocke. 2008. The Cost of Interrupted Work: More Speed and Stress. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 107–110. https://doi.org/10.1145/1357054.1357072

[21] Gloria Mark, Andrew L. Kun, Sean Rintel, and Abigail Sellen. 2022. Introduction to this special issue: the future of remote work: responses to the pandemic. *Human–Computer Interaction* (2022). https://doi.org/10.1080/07370024.2022.2038170

[22] Jennifer Marlow, Eveline Van Everdingen, and Daniel Avrahami. 2016. Taking notes or playing games? Understanding multitasking in video communication. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 1726–1737.

[23] Joseph Edward McGrath. 1984. *Groups: Interaction and performance*. Vol. 14. Prentice-Hall Englewood Cliffs, NJ.

[24] Sarnoff A Mednick. 1968. The remote associates test. *The Journal of Creative Behavior* (1968).

[25] Felix Müller-Wienbergen, Oliver Müller, Stefan Seidel, and Jörg Becker. 2011. Leaving the beaten tracks in creative work–A design theory for systems that support convergent and divergent thinking. *Journal of the Association for Information Systems* 12, 11 (2011), 2.

[26] Joseph W. Newbold, Anna Rudnicka, David Cook, Marta E. Cecchinato, Sandy J.J. Gould, and Anna L. Cox. 2021. The new normals of work: a framework for understanding responses to disruptions created by new futures of work. *Human–Computer Interaction* (2021). https://doi.org/10.1080/07370024.2021.1982391

[27] Brid O'Conaill and David Frohlich. 1995. Timespace in the workplace: Dealing with interruptions. In *Conference companion on Human factors in computing systems*. 262–263.

[28] Gartner Research. 2020. COVID-19 Bulletin: Executive Pulse. Retrieved Sep 7, 2021 from https://www.gartner.com/en/documents/3982949/covid-19-bulletin-executive-pulse-3-

[29] Gartner Research. 2020. Covid-19 could cause permanent shift towards home working. Retrieved Sep 7, 2021 from https://www.gartner.com/en/documents/3982949/covid-19-bulletin-executive-pulse-3-

[30] Advait Sarkar, Sean Rintel, Damian Borowiec, Rachel Bergmann, Sharon Gillett, Danielle Bragg, Nancy Baym, and Abigail Sellen. 2021. *The Promise and Peril of Parallel Chat in Video Meetings for Work*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411763.3451793

[31] Stefan Seidel, Felix Müller-Wienbergen, and Jörg Becker. 2010. The concept of creativity in the information systems discipline: Past, present, and prospects. *Communications of the Association for Information Systems* 27, 1 (2010), 14.

[32] Abigail J Sellen, Giford Louie, JE Harris, and AJ Wilkins. 1997. What brings intentions to mind? An in situ study of prospective memory. *Memory* 5, 4 (1997), 483–507.

[33] Minhyang (Mia) Suh, Frank Bentley, and Danielle Lottridge. 2018. "It's Kind of Boring Looking at Just the Face": How Teens Multitask During Mobile Videochat. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 167 (Nov. 2018), 23 pages. https://doi.org/10.1145/3274436

[34] Thomaz Teodorovicz, Raffaella Sadun, Andrew L. Kun, and Orit Shaer. 2021. How does working from home during COVID-19 affect what managers do? Evidence from time-Use studies. *Human–Computer Interaction* (2021). https://doi.org/10.1080/07370024.2021.1987908

[35] W Scott Terry. 1988. Everyday forgetting: Data from a diary study. *Psychological reports* 62, 1 (1988), 299–303.

[36] Lena Waizenegger, Brad McKenna, Wenjie Cai, and Taino Bendz. 2020. An affordance perspective of team collaboration and enforced working from home during COVID-19. *European Journal of Information Systems* 29, 4 (2020), 429–442.

[37] David Weschsler. 1971. Concept of collective intelligence. *American Psychologist* 26, 10 (1971), 904.

[38] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *science* 330, 6004 (2010), 686–688.