Human-Human Collaboration: What can we learn from Human Group Collaboration to improve Human-Robot Collaboration?

Eike Schneiders
eike.schneiders@nottingham.ac.uk
School of Computer Science,
University of Nottingham
Nottingham, United Kingdom

Christopher Fourie
Julie Shah
ckfourie@mit.edu
julie_a_shah@csail.mit.edu
Dept. of Aeronautics and
Astronautics, MIT
Cambridge, MA, USA

Malte Jung
mfj28@cornell.edu
Department of Information Science,
Cornell University
Ithaca, NY, USA

(a) Iteration 24 (b) Iteration 44 (c) Iteration 64 (d) Iteration 84

Figure 1: Entrainment leads to consistency in point-of-assembly as illustrated here over the course of 60 iterations.

ABSTRACT

Entrainment during collaboration positively affects trust, willingness to collaborate, and likeability towards collaborators. In this workshop paper, we present a mixed-method lab study to investigate characteristics of successful entrainment leading to group-based temporal synchronisation. Inspired by industrial work, we developed a fast-paced, short-cycle repetitive task. Using motion tracking, we investigated entrainment in both dyadic and triadic task completion. Initial findings are related to different leader-follower patterns, the importance of the point-of-assembly, and the value of sensory information. Based on these findings we hope to inform robotic behaviour for improved human-robot collaboration.

CCS CONCEPTS

• **Human-centered computing → Laboratory experiments**; **User studies**; *Empirical studies in collaborative and social computing*.

KEYWORDS

Entrainment in dyads and triads, temporal synchronisation, collaboration in groups, non-dyadic human-robot interaction

ACM Reference Format:
Eike Schneiders, Christopher Fourie, Julie Shah, and Malte Jung. 2018. Human-Human Collaboration: What can we learn from Human Group Collaboration to improve Human-Robot Collaboration?. In *HRI '23: WYSD - ACM/IEEE International Conference on Human-Robot Interaction, March 13–16, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 4 pages.

1 INTRODUCTION

The industrial sector remains one of the fastest growing application areas for collaborative robotics [11]. Given the complexity of manufacturing processes, many tasks require collaboration between multiple actors, including humans and robots. Yet, most studies within the field of Human-Robot Interaction (HRI) emphasise the investigation of dyadic interaction and collaboration [13], i.e., the investigation of the interaction between *one* human and *one* robot. In complex group configurations, efficient collaboration around tasks becomes even more relevant. Prior research has highlighted the importance of temporal synchronisation between collaborators (i.e., rhythms between interaction partners) for efficient collaboration [5, 9, 15]. To achieve this temporal synchronisation, it is vital that collaborators *entrain* with one another, thereby achieving better coordination. As Cross et al. [1] state: *"Entrainment refers to temporally coupled or synchronised systems, and it is the process of things moving in time together"*. In other words, entrainment refers to the process of falling into temporal synchronisation. Achieving temporal synchronisation during collaboration through the process of entrainment leads to a multitude of benefits: a stronger feeling of togetherness and connection [2], greater likeability between collaborators [2, 3], as well as willingness to cooperate [8, 17]. The occurrence of entrainment, and its effects on human collaborators, has been investigated in a variety of contexts (e.g., [7, 14, 16]).

To utilise the strengths of both humans and robots during collaboration, an understanding of how these can collaborate efficiently is needed. To investigate this, we study human entrainment in dyads and triads to understand the entrainment process between humans with the goal of informing future research on human-robot collaboration. This is led by the two research questions:

(1) *How do humans achieve collaborative rhythms?*
(2) *How can lessons from human-human entrainment be transferred to improve human-robot collaboration?*

In this paper, we present our completed study, its design, and initial results in order to spark reflection on potential further steps as well as improvements. The study design was developed in order to get insights into how human dyads and triads entrain during the completion of a fast-paced, short-cycle repetitive task. We investigated this using a mixed-method lab study with 50 participants.

2 RELATED WORK

A recent paper by Roy and Edan [10] presents a lab-based study to investigate how pairs of human shelf fillers entrain with each other. This study was inspired by the working context observed at supermarkets. They conducted a mixed-method study utilising field studies to better understand supermarket employees' tasks and re-create those in a controlled lab environment. Using both subjective data collection methods, such as a verbal description of the experience during task completion or post-session questionnaires, as well as quantitative performance metrics, including 'the number of bottles for each 10-second interval' or 'level of coordination amongst team members', Roy and Edan identified several factors related to dyadic entrainment in short-cycle repetitive tasks. They observed different behaviour, as shelving bottles included two clear roles (giver and receiver). For instance, while the giver uses visual cues to move the bottles towards the receiver, the receiver does not often look away from the shelf, simply assuming that the point-of-handover will be the same for every shelving cycle. Furthermore, the authors could observe a clear improvement over time in both the consistency of working speed and coordination, which indicates a better synchronisation achieved through entrainment. While Roy and Edan's study investigated dyadic entrainment in one particular setting (bottle shelving), these lessons might be transferable to other tasks and group formations.

Rinott and Tractinsky [8] present a literature review on interpersonal motor synchronisation (IMS)-based literature. They propose a framework mapping out different types of Joint-Action—which does not require any type of temporal synchronisation as long as all actors work towards the same goal. As part of this framework, the authors classify 'Synchronisation' as the result of 'Entrainment': *"Entrainment is a process of reaching the same rhythm and is a required part of synchronising with someone else."* [8, Section 2.1]. Furthermore, based on existing literature, they synthesise eleven dimensions (e.g., temporality, information exchange, or the number of participants) relevant to the study of IMS. Specifically, an important aspect during the investigation of IMS is the temporal aspect. IMS can be investigated using external entrainment, such as a metronome (e.g., [2]) providing a rhythm, or using 'Mutual entrainment', in which participants entrain with each other. Alternative names for the same phenomena have been used, including 'Social entrainment' and 'Mutual social entrainment' [6].

3 STUDY DESIGN

This section describes the study as well as the utilised data collection and analysis methods. The conducted study is inspired by previous work on human-human entrainment to inform human-robot entrainment (e.g., [10, 12]).

3.1 Participants

We recruited 50 participants (38 female, 12 male, average age: 22.96, SD: 4.06) with varying educational backgrounds, including health, engineering, and art, to mention a few. Participants were recruited using social media postings, flyers on campus, as well as through a dedicated web page. The study was approved by Cornell Universities ethics review board for studies involving human participants, and all participants were compensated with a $50 Amazon gift card for their participation. We matched participants randomly into ten triads (N = 30, G1 - G10) and ten dyads (N = 20, G11 - G20). Additional participants were recruited for pilot testing.

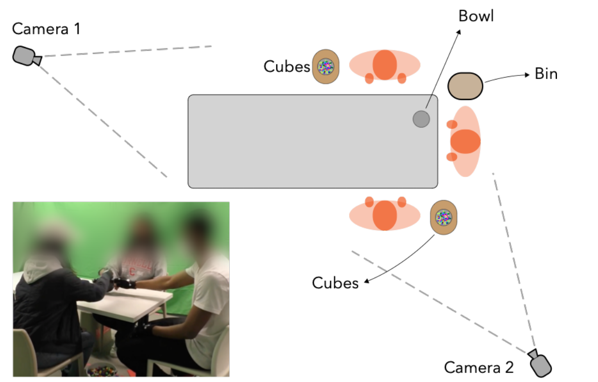

Figure 2: *Triadic setting* - Cubers are placed along the long sides of the table with a bowl of cubes placed next to them. The bowler is placed at the end of the table with the bowl and the collection bin next to them. The cubers/bowler could choose to place the cubes/bin on their right or left side. The lower left corner shows the cropped view of Camera 2 (G9).

3.2 Tasks

3.2.1 Development and Pilot testing. To identify suitable tasks in which entrainment has a high potential to naturally occur, we brainstormed potential ideas resulting in a total of eight different tasks. All tasks were inspired by typical industrial tasks such as pick and placing or assembly. After experimenting with these eight tasks, we narrowed the selection down to two viable candidates, which were piloted both in the dyadic as well as the triadic configuration. Based on our observations as well as the feedback provided by the pilot participants, we selected a single task for extensive analysis as described below.

3.2.2 The Task: Pick-and-Place. Inspired by previous studies [4, 10] we designed one task, inspired by industrial pick-and-placing, to investigate entrainment. The goal of the task was to move 1 × 1 × 1 cm plastic cubes from a bowl to a collection bucket through collaboration. To accomplish this task, we defined two distinct roles, the 'bowler' and the 'cuber(s)'. In the *dyadic setting*, these two roles would be placed across from one another with a table in between them. It was the cuber's task to pick up *one* cube at a time and drop it in the bowl. When precisely one cube was in the bowl, it was the bowler's responsibility to move the bowl over the collection bin—placed to the right or left of the bowler—and drop the cube from the bowl into the collection bin. Following the emptying of the bowl, the task would repeat.

In the *triadic setting*, the bowler would sit at the end of the table, and the *two cubers* would sit to the right and left of the bowler, respectively (see Figure 2). Other than the inclusion of an additional cuber, the task remained the same. For both the dyadic and triadic task completion, the division of roles and the point-of-assembly (i.e., the location where collaborators' actions meet), were decided amongst participants.

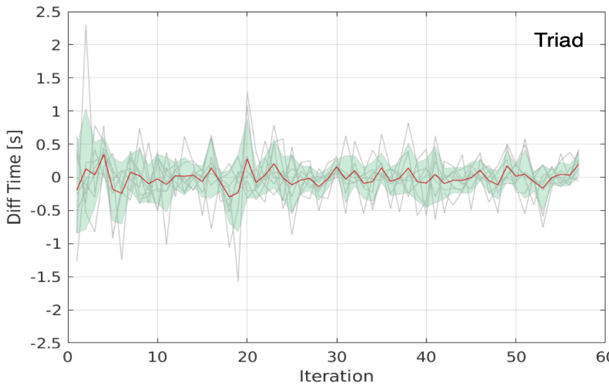

Figure 3: Motion tracking - triadic collaboration

3.3 Experimental Procedure

Upon entering the lab area, participants received written information about the study, including its duration and overall structure. This was further elaborated on verbally, and participants had the chance to ask questions before reading and signing the informed consent form. We first collected demographic data (i.e., age, gender, and handedness) using a Qualtrics questionnaire. Following this, we introduced each pair or triad to the task, after which participants could ask questions again and start completing the task for an indefinite period. After four minutes, we would stop the participants for the post-task semi-structured interview (average interview time was 11min and 40sec). The interview was structured around five separate topics: i) general experience during task completion, ii) experience and preference towards the chosen role, iii) experiences and strategy for negotiation of point-of-assembly, iv) trust towards other participants, and v) perceived performance. The post-task interviews were conducted in dyads or triads to foster insightful conversations about the experience.

3.4 Data Collection and Analysis

For this study, we utilised a mixed-method approach combining multiple qualitative methods with quantitative measurements. Specifically, we used an OptiTrack camera setup with 18 cameras for motion tracking of participants' hands and the position of the bowl. Additionally, we used two video cameras (see Figure 2) to ensure audio-video recordings of participants for later video analysis.

As we collected multiple data streams, we used various data analysis methods. We applied thematic analysis to the semi-structured post-session interviews to identify themes relevant to the investigation of entrainment. Using the motion tracking data, we detected the completion of each iteration of the task, quantifying changes in task completion over time. This approach allowed us, for example, to identify the number of task iterations completed during each 10-second interval [8]. Furthermore, we used motion-tracking to identify a decrease in temporal variation between task iterations as indicative of successful temporal entrainment (see Figure 3). Lastly, we used the video recordings, analysed using the video annotation software ELAN, to contextualise participants' reported experiences. For the purposes of this workshop paper we focused on the qualitative data.

4 FINDINGS

This section will briefly highlight three of the preliminary findings related to the 1) emergence of different leader-follower patterns, 2) the consistency of the point-of-assembly, and the 3) importance of sensory information.

Leader and Follower. The data analysis revealed three distinct types of leader-follower patterns: static, flexible, and absent. During most dyadic collaborations, we observed a pattern emerge focusing on a *static* leader. Participants in the dyadic configurations perceived the cubers task as more difficult, making this task the bottleneck. Given that this task was being perceived as slower, the bowler would typically be ready before the cuber would had picked the next cube. Therefore, the cuber was often perceived as the leader during the dyadic collaboration (7/10 dyads). As expressed by G15: *'I like to believe that I set the speed [...] I don't recall waiting for the bowl very often'.* Contrasting these dyads, G14 and G16 described the *absence* of a distinguishable pattern. They described that each dyad member was seemingly independent of one another, completing their task. The dyads found a rhythm that led them to complete their sub-task at the same speed without needing a dedicated leader.

Contrasting the dyads, the triads were less explicit about the presence of a leader. Given that each person interacted with two other collaborators, the presence of two leaders was a possibility (e.g., G1). Here, the bowler followed both cuber's behaviour, who were leading the interaction and providing the rhythm to which the bowler adjusted their pace.

> *'I feel I [the bowler] ... am the follower for their collective behaviours, because I was trying to match the tempo. ... I personally see the dropping of the cube as an audio cue on what the rhythm is.'* - G1

The last pattern observed was the presence of a flexible—or varying—leader during the collaboration. This was, for instance, expressed by one of the cubers in G8 who stated that: *'I don't think there is a leader role, but I kind of follow her because ... she puts the cube in first, and then I will follow.'*

Point-of-Assembly (PoA). We define the PoA as the point in which the bowl is placed/held and the cube(s) are dropped in. We observed different strategies employed to increase individual and/or group efficiency. Two different strategies for negotiating of the PoA emerged, namely: i) 'optimisation for group efficiency' and ii) 'optimisation of own task'.

For strategy i), groups had the tendency to prioritise adjustment of the individual task, not focusing on ease of task completion, but to decrease the collaborative group effort. Several groups (e.g., G2, G4, G8, G12, G16) expressed this behaviour. The bowler in G4, for instance, used this strategy to facilitate the two cubers: *'They were putting the cubes down at the same time. So I just need to place it in the middle.'.* A similar approach was described by the bowler in G2, who optimised the placement of the bowl to increase overall group efficiency. Even though this increased the range of motion needed by the bowler, it reduced the task complexity for the two cubers.

A different approach was observed in which several participants of both roles reported on the *optimisation of their own task*—not necessarily to the benefit of the other collaborators— leading to a change in behaviour of the other participants. While both strategies led to an increase in overall efficiency, the difference lies in the intention. Contrasting the adjustment of their own behaviour to facilitate collaborators aims at improving group efficiency, the second approach achieves this only as a potential side effect resulting from the individual optimisation. Nevertheless, the overall efficiency increased as it resulted in an optimisation of the group's bottleneck. Similar observations were made in the triadic setting. In G5, for instance, the bowler describes adjusting the position of the bowl to reduce the distance to the bin, thereby decreasing the need for movement for their own task.

Lastly, regardless of the strategy utilised by the groups, multiple groups expressed (e.g., G1, G3, G5, G7, G11, G12) that the efficiency of task completion depends on the consistency. As a higher consistency led to greater predictability in motion this led to less downtime in task completion.

In addition to participant descriptions, the consistent PoA selection was evident from the video analysis. After, e.g., G4 had found a position that worked for them—the bowl slightly closer to the side of the bin—it was clear to see that an effort was put into keeping the PoA as consistent as possible. Figure 1 illustrates the bowl placement for G4 over the course of 60 iterations ($24 \rightarrow 44 \rightarrow 64 \rightarrow 84$).

Importance of Sensory Information. An additional characteristic identified through the study is related to the dependence on sensory information. While the task required visual and tactile information, the importance of auditory information was observed during the video analysis and described by participants.

While the use of visual and tactile information was expected, most groups (e.g., G1, G5, G8, G9, G12, G17) further relied on audio cues produced by the task. The benefit of the auditory signalling was expressed in multiple directions, both to and from the bowlers' task. As it could at times be difficult to see if the cubes were dropped by the cuber(s), bowlers reported using sound to confirm that the cubes were in the bowl. Auditory cues further provided information about when the bowl could be moved towards the bin. Audio cues were relevant in both the dyadic and the triadic condition.

´I would look, but I also make sure that I heard something.' - G8

During the video analysis, we could observe that bowlers stopped the interaction when they missed the bin. This was due to the absence of audio cue when the cubes did not hit the bin, but landed on the carpet, not producing the distinct sound the bowler was listening for. Just as the bowler used audio cues to know when to proceed to the next step of their task, so did the cuber(s). In the cuber(s) case the signal used was caused by the bowler dropping cubes in the bin next to them, thereby signalling the end of the iteration. This sound of falling cubes would signal the cuber(s) that the bowl is about to return to the PoA to collect the next cube.

These findings highlight the value of multi-modal signalling during task completion. The use of, e.g., acoustics, frees other senses for the preparation of the next iteration, while simultaneously providing feedback about the progression of the iteration.

5 WORKSHOP DISCUSSION

With participation in the WYSD workshop at HRI'23 we invite criticism and reflection on the described study. We hope this will provide valuable insights allowing us to make the most of the collected data. Furthermore, we hope to have stimulating discussions on how the study could have been improved to inform further steps in the project.

We invite input to, e.g., the following questions:

(1) Can the data be used to answer additional open questions?
(2) What should we have done differently?
(3) Any, maybe obvious, next directions we have overlooked?
(4) How to identify the most impactful finding for the HRI community?

We initially utilised this study to provide design considerations for HRI. We invite alternate strategies for creating impactful contributions for the HRI community based on this study.

ACKNOWLEDGMENTS

This work is supported by the Engineering and Physical Sciences Research Council [grant number EP/V00784X/1].

REFERENCES

[1] Liam Cross, Martine Turgeon, and Gray Atherton. 2019. How Moving Together Binds Us Together: The Social Consequences of Interpersonal Entrainment and Group Processes. *Open Psychology* 1, 1 (2019), 273–302.
[2] Michael J. Hove and Jane L. Risen. 2009. It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition* 27, 6 (2009), 949–960.
[3] Jacques Launay, Roger T Dean, and Freya Bailes. 2014. Synchronising movements with the sounds of a virtual partner enhances partner likeability. *Cognitive processing* 15, 4 (2014), 491–501.
[4] Tamara Lorenz, Alexander Mörtl, Björn Vlaskamp, Anna Schubö, and Sandra Hirche. 2011. Synchronization in a goal-directed task: Human movement coordination with each other and robotic partners. In *2011 RO-MAN*. 198–203.
[5] Taiwoo Park, Uichin Lee, Bupjae Lee, Haechan Lee, Sanghun Son, Seokyoung Song, and Junehwa Song. 2013. ExerSync: Facilitating Interpersonal Synchrony in Social Exergames. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. ACM, New York, NY, USA, 409–422.
[6] Jessica Phillips-Silver, C. Athena Aktipis, and Gregory A. Bryant. 2010. The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement. *Music Perception* 28, 1 (09 2010), 3–14.
[7] Paul Reddish, Ronald Fischer, and Joseph Bulbulia. 2013. Let's Dance Together: Synchrony, Shared Intentionality and Cooperation. *PLOS ONE* 8, 8 (08 2013).
[8] Michal Rinott and Noam Tractinsky. 2021. Designing for interpersonal motor synchronization. *Human–Computer Interaction* 37, 1 (2021), 1–48.
[9] Raquel Breejon Robinson, Elizabeth Reid, James Collin Fey, Ansgar E. Depping, Katherine Isbister, and Regan L. Mandryk. 2020. Designing and Evaluating 'In the Same Boat', A Game of Embodied Synchronization for Enhancing Social Play. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. ACM, New York, NY, USA, 1–14.
[10] Someshwar Roy and Yael Edan. 2020. Investigating joint-action in short-cycle repetitive handover tasks: The role of giver versus receiver and its implications for human-robot collaborative system design. *International Journal of Social Robotics* 12, 5 (2020), 973–988.
[11] Lindsay Sanneman, Christopher Fourie, and Julie A. Shah. 2021. The State of Industrial Robotics: Emerging Technologies, Challenges, and Key Research Directions. *Foundations and Trends® in Robotics* 8, 3 (2021), 225–306.
[12] Eike Schneiders and Stanley Celestin. 2022. Non-Dyadic Entrainment for Industrial Tasks. In *Workshop on Joint Action, Adaptation, and Entrainment in Human-Robot Interaction.*
[13] Eike Schneiders, EunJeong Cheon, Jesper Kjeldskov, Matthias Rehm, and Mikael B. Skov. 2022. Non-Dyadic Interaction: A Literature Review of 15 Years of Human-Robot Interaction Conference Publications. *Transactions on Hum.-Robot Interact.* 11, 2, Article 13 (2022), 32 pages.
[14] Bronwyn Tarr, Jacques Launay, and Robin I.M. Dunbar. 2016. Silent disco: dancing in synchrony leads to elevated pain thresholds and social closeness. *Evolution and Human Behavior* 37, 5 (2016), 343–349.
[15] Bronwyn Tarr, Mel Slater, and Emma Cohen. 2018. Synchrony and social connection in immersive virtual reality. *Scientific reports* 8, 1 (2018), 1–8.
[16] Jorina von Zimmermann and Daniel C. Richardson. 2016. Verbal Synchrony and Action Dynamics in Large Groups. *Frontiers in Psychology* 7 (2016), 2034.
[17] Scott S. Wiltermuth and Chip Heath. 2009. Synchrony and Cooperation. *Psychological Science* 20, 1 (2009), 1–5. PMID: 19152536.