Replication Study: Investigating Media Choice and Types of Media Collections of Distributed Workers Using Digital Traces

JOSCHKA A. HÜLLMANN, University of Münster, Germany
SIMONE KREBBER, University of Münster, Germany

Work is increasingly conducted in a distributed manner, enabled by a heterogeneous and growing set of digital communication and collaboration tools, which we call a media collection. This array of new digital collaboration tools with a trend towards multi-purpose integrated systems raises existing research on media collection choice into question. Based on a unique quantitative sample of digital traces, describing activity logs of tool use over a period of eight months, we replicate existing research. First, we identify the frequently used media collections and their purpose. Then, we test the association between two social factors, the assigned supervisor and coworkers, as well as physical location, with a worker's choice of a media collection. Our findings corroborate existing results that see information sharing as the dominant communication purpose in the identified media collections. Our data shows that the supervisor is strongly associated with an employee's choice of media collection, whereas coworkers and physical location are of little relevance in distributed work.

1 INTRODUCTION

The COVID-19 pandemic reinforces the ongoing societal change towards more spatially and temporally distributed knowledge work [1, 2]. Distributed workers rely on communication tools that enable them to collaborate at a physical distance and asynchronous times [3]. Usually, these workers do not rely on a single tool to fulfil their tasks, but instead use a subset of tools routinely [4, 5]. We call this subset of tools their media collection, which the workers choose from a heterogeneous set of tools available in their organization, i.e., the organization's media landscape. Making this choice for a media collection depends on social factors such as supervisor and coworkers, which are theorized as antecedents of media choice.

Using multiple tools allows for flexibility but comes with drawbacks: Keeping track of when, how, and what tool to choose is challenging [3, 6–9]. Such flexibility and the availability of a wide variety of tools and a multiplicity of task structures require coordination for the alignment of media choice among distributed workers. This often leaves workers with a coordination overhead that negatively affects performance because discoordination and non-alignment lead to non-effective tool use and collaboration [6]. Exploring the antecedents of the selection process enables management to explain why workers choose a particular subset of tools and deal with their workers' coordination efforts [3, 5].

Previous studies focus on the media choice between traditional media and digital tools, for example, face-to-face, phone, chat, email, or paper documents [10]. The use of digital tools gains momentum due to more distributed workers [1] and results in a growing market of collaboration software, diversifying the tools available in an organization [11, 12]. Moving the focus from traditional media towards digital tools demands rethinking previous research. We find a dearth of empirical classifications on the frequently used media collections as opposed to studies on the characteristics of single collaboration tools [13, 14]. Contrary to previous studies that consider both traditional media and digital tools [15], our work focuses exclusively on digital tools in distributed work. In this context, the routine use of such tools generates digital traces [16, 17]. Digital traces

are activity and interaction data retrieved from extensive and routine technology use, which allow us to observe spatially distributed teams that primarily perform computer-based work and whose work activities are otherwise difficult to observe [16, 18–20]. In our study, the traces include the logs of work activities from workers of a distributed global service provider. Based on these digital traces, we replicate the identification of frequently used media collections and investigate the antecedents of media choice. Our research questions are:

RQ1: Which types of media collections (by communication purpose) are frequently used in distributed work?

RQ2: To what extent do supervisors, coworkers, and physical location influence the choice of media collections in distributed work?

Our goals are (1) to identify the media collections which are in frequent use among our sample and (2) to corroborate the existing theory of media choice in a distributed work setting. We address both questions by replicating and testing existing hypotheses using a unique sample of digital traces.

2 THEORETICAL FOUNDATIONS

2.1 Definitions

Before we outline the theoretical foundations relevant to our study, we briefly describe the central terms of our study in Table 1.

Table 1. Overview of Terms Employed Throughout the Study

Term	Meaning
Media landscape	A heterogeneous set of tools available in an organization.
Media collection	An individual set of communication and collaboration tools, which are routinely used by a worker.
Media choice	The selection process that explains why workers choose a particular subset of tools, with social factors such as supervisor and coworkers being theorized as antecedents.
Digital tools	Tools that run on computers or mobile devices and enable communication and collaboration over the internet protocol (e.g., Microsoft 365).
Digital traces	Activity and interaction data retrieved from extensive and routine technology use.

In addition to the overview of relevant terms, Fig. 1 provides a visualization of how the three concepts media landscape, media collection, and media choice relate to each other. The media landscape is the set of tools available in an organization, while the media collection is a subset of this larger set of tools actively chosen by an individual through the process of media choice. The two research questions of our paper are visualized in Fig. 1. The prevalence of media collections is addressed by the first research question, while the media choice process is subject to the second one.
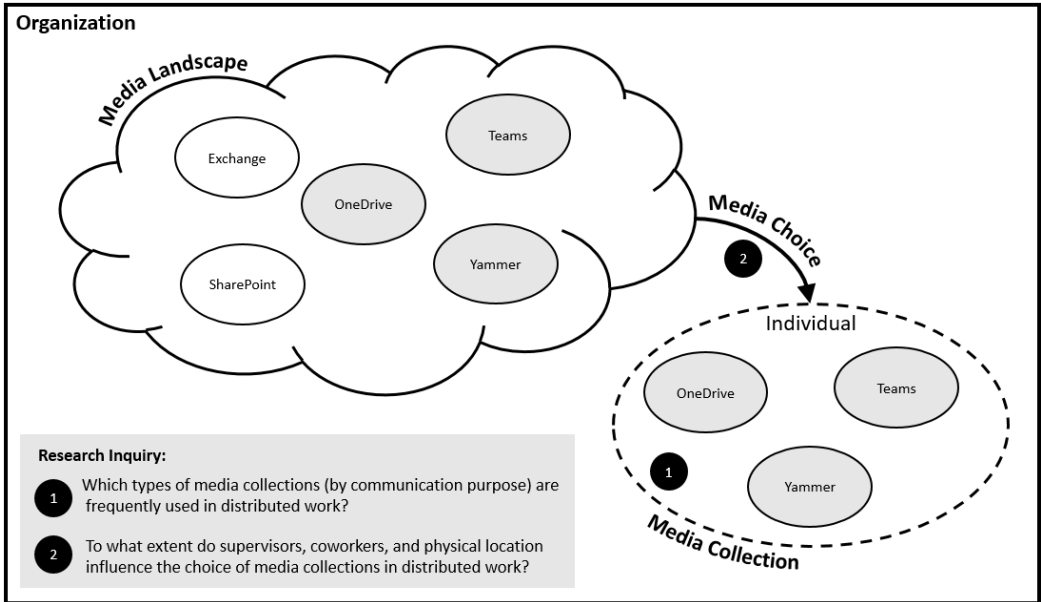
Fig. 1 Relationship between media landscape, media choice, and media collection.

2.2 Media Collections

Distributed workers require a set of communication and collaboration tools, i.e., a media collection for collaborating effectively [3]. The concept of a media collection is embedded in and derived from the framework of communication media repertoires by Watson-Manheim and Bélanger [7]. In addition to the routinely used communication media, the original framework seeks to explain the communication purpose and the behavioural patterns of usage, whereas we focus on depicting the media collection and the antecedents of its choice.

Watson-Manheim and Bélanger [7] group different media collections according to their communication purpose, for example, coordination, information sharing, or relationship development. Lee et al. [4] characterize media collections by their size (as in several tools included) and how the tools are used (sequentially or concurrently).

Contrary to this research on media collections, which consist of multiple tools, extensive research describes the use of single tools [4]. Lee et al. [4] argue that research on single tools informs our understanding of media collections. Previous studies on single tools (e.g., groupware, or e-collaboration systems) have resulted in multiple classifications. Single tools are either synchronous or asynchronous and require colocation of workers or are spatially flexible [14, 21, 22]. However, most tools nowadays provide asynchronous communication features such as text, voice, and video messaging—even if they primarily aim at real-time communication, for example, Skype. Thus, by design, digital tools do not require the colocation of the users as they enable distributed work. Another classification scheme considers the type of communication and the features of tools [15, 22]. However, modern tools advanced from single-functionality towards multi-purpose integrated systems [14], which renders classification by features difficult. Other studies abstract from the technical features to the primary communication purpose of the tools, e.g., communication, information sharing, and collaboration [23, 24].

In our analysis, we first identify the frequently used media collections in the organization at hand. As explicated above, distinguishing media collections by technical features, synchronicity, or colocation is infeasible for integrated systems and distributed work. Hence, we try to replicate the results of Watson-Manheim and Bélanger [7] and focus on the primary communication purpose (either information sharing, relationship development, or both) for distinguishing the identified media collections.

2.3 Antecedents of Media Choice

After having identified the media collections prevalently used, this study addresses the antecedents of choosing a media collection. Research on media choice matured over time with the emergence of two perspectives: rational choice models and collective choice models. While the rational choice approaches presume objective factors of choice, the collective choice models argue that media use is contingent and decisions are subject to social influence and perceptions of media appropriateness [25]. We follow the collective choice models, which posit that the social factors, actions, and behaviours within workgroups affect a worker's attitude towards communication technologies and affect the choice of tools [26]. Examples of social factors include norms for media use and the imitation of coworkers' media usage [25], perceived media richness (as opposed to objective media richness), and media experience [27], task experience, and situational factors [28], as well as organizational factors, such as job role or position in the hierarchy [5]. Further factors include the perceived communication risk, for example, impaired reception or faulty understanding [4], structuring conditions and perceived consequences of media use [7], and the availability and capability of media in conjunction with the strategic goals of the workers [3]. These previous studies are based on surveys, interviews, and observations. We aim to replicate and corroborate their results using the digital traces of digital tools. Since previous studies [29, 30] find that the attitudes of coworkers and supervisors are a critical social influence for media choice, we focus on these two hypotheses in distributed work.

Hypothesis Development

Supervisors exert influence via verbal statements through which workers adopt the supervisors' perceptions of media choice [27]. As part of such verbal statements, supervisors may push and promote particular tools, for example, if they favour one tool [27]. Conversely, workers may choose to imitate the supervisor's media choice to ease communication [25]. Consequently, we hypothesize:

H1a: The assigned supervisor is positively associated with the choice of media collection in distributed work.

H1b: The assigned supervisor's media collection choice is positively associated with the choice of media collection in distributed work.

Coworkers. In addition to the supervisor, the coworkers shape attitudes towards tools through everyday talk, discussing benefits and drawbacks of tools as well as sharing knowledge on how tools are used. More specifically, the coworkers influence how tasks are perceived and what is the appropriate media choice to solve a task [27]. They establish social structures and norms on media use in the organization through their routine tool use [25]. Coworkers co-learn about the tools and

influence each other's perceptions of a tool [25], as they must use the same, or at least compatible tools, to communicate. Because of these reasons, we hypothesize:

H2: The coworkers are positively associated with the choice of media collection in distributed work.

Physical location is a confounding variable, which constrains how people meet and communicate and is therefore included as a control variable. As the social influence disseminates via social encounters such as water-cooler chats, ad-hoc meetings, and random encounters, previous research has considered the physical location as a relevant factor for the choice of media collection [29–31]. However, in our study, we look at distributed work across multiple sites and exclusively at digital tools. We do not consider any traditional media such as face-to-face or phone. As a result, in our setting, we hypothesize that the physical location has only a negligible effect on the worker's media choice because physical dispersion is less important if the choice is only between digital tools. Instead, we hypothesize:

H3: The assigned location is negligibly associated with the choice of media collection in distributed work.
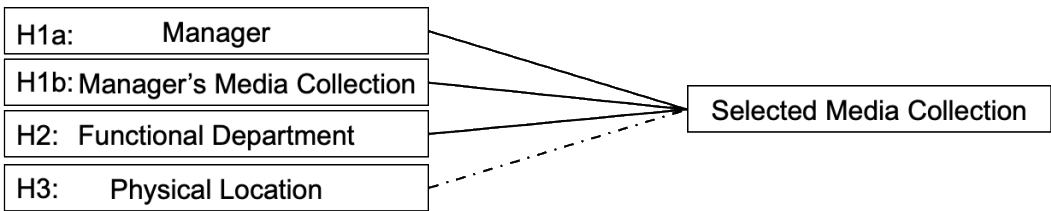

Fig. 2. Hypothesized associations. The solid line depicts a hypothesized association. The dashed line depicts the hypothesized negligible association.

3 METHODS

Our sample consists of Microsoft 365 data of an organizational unit from a global systems integrator and managed service provider with 30,000 employees. The selected organizational unit operates across 18 locations in one European country. Contrary to survey and laboratory studies (e.g., [27, 32]), we analyse longitudinal and anonymized usage data from Exchange, OneDrive, SharePoint, Teams, and Yammer contained in the Microsoft 365 Suite. The sample consists of 813 knowledge workers and contains usage data that is aggregated per month and covers the timespan from June 2018 until January 2019. The usage frequency of each tool is given as the sum of actions performed per tool and per month. For example, downloading or uploading a file to OneDrive, sending an email or chat message, accessing files in SharePoint, and other actions are counted for each tool. We filtered obsolete organizational subunits with zero members from the data set. The supervisor of each worker and the physical location are included in the analysis as factor variables. An excerpt of the data and the underlying data structure is illustrated in Table 2.

Table 2. Excerpt from the Data Set Showing Exemplary Usage Frequencies of a Worker

ID	Exchange	One Drive	Share Point	Yammer	Teams	Org. Subunit	Supervisor ID	Location
891	10,583	32	69	47	0	Operations	1111	Location1
892	10,670	1,524	48	49	0	Operations	891	Location1
893	10,086	18	12	19	0	Operations	891	Location4
894	8,387	988	41	69	0	Operations	891	Location1

Pre-processing. As with any statistical model, we have underlying assumptions that manifest themselves in parameter configuration for our pre-processing and data cleansing. Because choosing fixed values for these parameters would be arbitrary, we test multiple parameter configurations for our models—as is recommended practice [33, 34]. The most central assumption is the minimum usage frequency threshold that determines whether a worker is actively using a particular tool. Only if workers are actively using a tool, it is included in their media collection. For defining active usage, we test thresholds between 1 and 200 activities per month, all leading to similar results. Besides the usage frequency threshold, we filter organizational subunits depending on the minimum number of members constituting an organizational subunit. We test the values 0, 5, 10 for minimum members of a unit. In our subsequent analyses, we test all parameter combinations (N = 60) and report the mean and box plots for the calculated test statistics.

Visual and cluster analysis. For identifying the frequently used media collections, a visual analysis is performed and then corroborated with Ward's hierarchical clustering (minimum within-cluster variance criterion) [35, 36].

In the visual analysis, we assess which tools are used together frequently. We determine the primary communication purpose of a media collection through its included tools. Based on the majority of purposes of the included tools in the media collection, we derive the purpose of the media collection itself [4]. Based on Schwade and Schubert [37], we consider OneDrive and SharePoint as information sharing tools, whereas Yammer and Teams are relationship development (or "social") tools. Exchange is considered an essential tool for communication and is deemed as the core tool because all workers use it.

The cluster analysis is based on a table with binary values that indicate for each tool whether an employee is an active user or not (active=1; not active=0). As stated above, the cluster analysis is repeated for different parameter combinations. As hierarchical clustering requires to choose a fixed number of clusters a priori, we determine the number of clusters through manual tuning from the visual analysis.

Hypotheses testing. We operationalize the coworkers' influence through the assigned organizational subunit of the worker. The media collections are given as distinct sets of tools. Both variables as well as supervisor and physical location are given as IDs. Because all variables are of nominal scale, and the factor levels reach up to 119, a multinomial regression would show very different factor loadings and not yield helpful results. Instead, we show the association between the factor variables and test the stochastic independence using Pearson's chi-squared test as well as Fisher's exact test (with Monte Carlo simulations for the p-values based on Patefield [38]). Accordingly, posthoc correlation analysis to estimate the effect size is performed with Cramer's V, which is suited for nominal measurements [39, 40]. Our interpretation of effect sizes follows Cohen's remarks on cross-tabulation [41, p.224, cited via 42, p.41].

4 RESULTS

After pre-processing, the analysis is conducted with 813 workers. Table 3 shows that Exchange is the tool used by all workers in the sample. The median count of the total sent emails over the eight months is 12,363. The next most used tools are SharePoint, OneDrive, and Yammer with median activity between 113 and 263. Teams is not in use by most workers. The 813 workers work in 10 organizational subunits, 18 different locations, and have 119 supervisors (see Table 3). Based on the visual analysis results, we identify eight relevant media collections (Fig. 3).

Table 3. Descriptive Statistics and Factor Levels

Descriptive Statistics							Factor Levels	
Tool/ Statistic	N	Mean	St. Dev.	Min.	Median	Max.	Variable	Levels
Exchange	813	16,049	14,635	495	12,363	128,878	Employee	N=813
OneDrive	813	5,627	33,116	0	114	707,030	Org. Subunit	N=10
SharePoint	813	808	1,852	0	263	19,636	Supervisor	N=119
Yammer	813	454	794	0	113	6,155	Location	N=18
Teams	813	19	183	0	0	4,987	Media Collection	N=8

Table 4 depicts the tools included in the identified media collections. The importance of Exchange is emphasized as 262 out of 813 workers use only Exchange. It is part of every frequently used media collection that we identified. Besides Exchange, SharePoint is another popular tool in the media landscape and part of four media collections. Another observation is the recurring absence of Microsoft Teams, which is not extensively used, and, thus, not part of the media collections, except for the "All" collection. Table 4 shows three media collections with an information-sharing focus, three collections with both information sharing and relationship development focus, but no media collection with only a relationship development focus.

Table 4. Media Collections in frequent use. N=60 Depicts the number of parameter configurations tested. Values given in the table are the average active users of each collection

Media Collection	Purpose	N	Mean	St. Dev.	Min.	Median	Max.
Exchange	Core	60	268.517	154.500	64	262	498
Exchange, OneDrive	Information Sharing	60	91.050	53.181	13	106	148
Exchange, SharePoint	Information Sharing	60	86.733	37.136	33	96	134
Exchange, SharePoint, OneDrive	Information Sharing	60	75.567	27.111	31	88	106
Exchange, SharePoint, OneDrive, Yammer	Both	60	121.917	108.223	6	87	316
Exchange, SharePoint, Yammer	Both	60	58.617	42.346	6	46	116
All	Both	48	18.167	15.833	3	12	46
Others	–	60	75.400	23.290	35	76	108

Visual and cluster analysis. Fig. 3 depicts the distribution of the media collections for varying parameter configurations. It is visible that Teams sees only little use, as higher thresholds for the minimum usage activity lead to an exclusion of Teams from many workers' media collections.

Higher thresholds for the minimum usage activity lead to an increase in the "Only Exchange" collection because email has the most regular activity (Table 2), which means that it is included even for high thresholds of minimum usage activity (Fig. 3). From the visual analysis, we conclude that eight media collections are relevant. Fig. 3 shows that the maximum number of active users for each of the media collections exceeds 81, which is 10% of all workers. For five of the media collections, the mean number exceeds 81 as well. The threshold of 81 was determined through manual tuning.

Having conducted the cluster analysis for different minimum usage thresholds, we compare the results to the identified media collections from the visual analysis and find that they are largely consistent. We provide a textual summary, as the full results with illustrations exceed the available space. Three up to six media collections are perfectly replicated by the clustering algorithm at every minimum usage threshold level. The collections "Exchange, SharePoint" and "Exchange, OneDrive" are uniquely identified by one cluster across all tested minimum usage threshold levels. The collections "Only Exchange" and "Exchange, SharePoint, OneDrive" correspond to one unique cluster at four of five minimum usage threshold levels. Even when a cluster cannot be perfectly mapped to one single media collection, there is still a significant overlap. This occurs, when one media collection is split into two clusters, two media collections are combined in one cluster, or one cluster contains only one single misclassified observation. Across the tested parameter configurations, the cluster results match the identified media collections for more than 700 workers. Consequently, we conclude that the results of the cluster results support the robustness of our analysis regarding the identification of media collections.
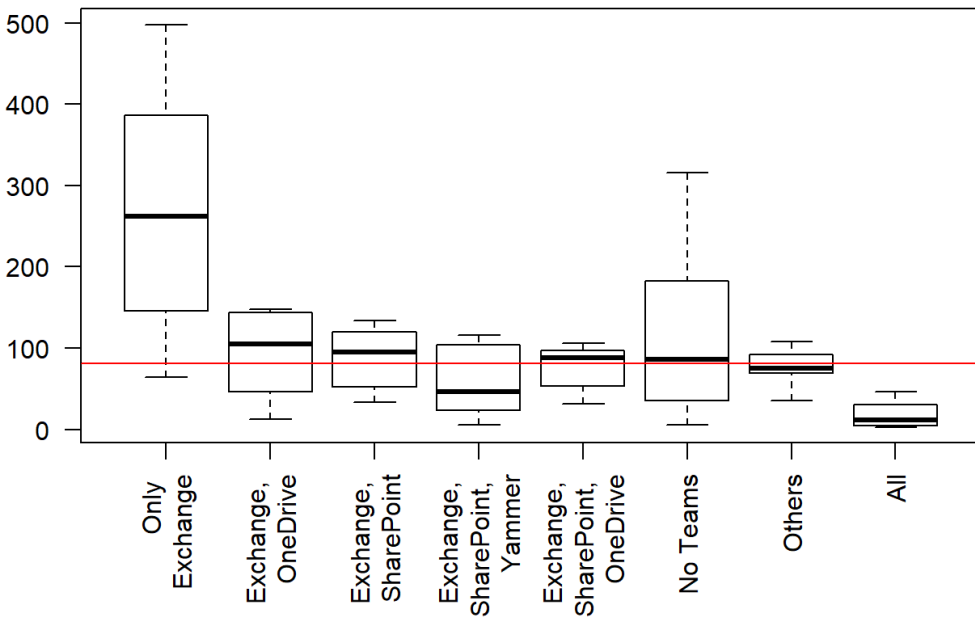

Fig. 3. Box plot of media collection distribution. The red line is at 81, which equals 10% of all employees.

Hypothesis testing. Table 5 and Fig. 4 show the Cramer's V correlations between the covariates and the selected media collection. The variance stems from the 60 different parameter combinations tested. According to Cohen [41, 42], a correlation of 0.5 is high, 0.3 is medium, and 0.1 is small[1]. Hence, our results show a high correlation between the supervisor and the choice of media collection, but only a small correlation between the supervisor's own choice of media collection and the worker's choice. Consequently, the hypothesis on the supervisor's social influence holds (H1a), whereas its particularization in the imitation hypothesis (H1b) does not hold. As coworkers and physical location show a small correlation with the media collection choice, the hypothesis about the coworkers' social influence does not hold (H2), whereas the negligible association with the physical association does hold (H3). None of the covariate pairs are independent across the tested parameter combinations according to Pearson's Chi-Squared test and Fisher's exact test. This, however, is expected given a sample size of N = 813 (cf. [43]).

Table 5. Cramer's V Correlation. None of the Variable Pairs are Independent. Pearson's Chi-Squared Test and Fisher's Exact Test Show Associations at the p < 0.01 Level for all Pairs

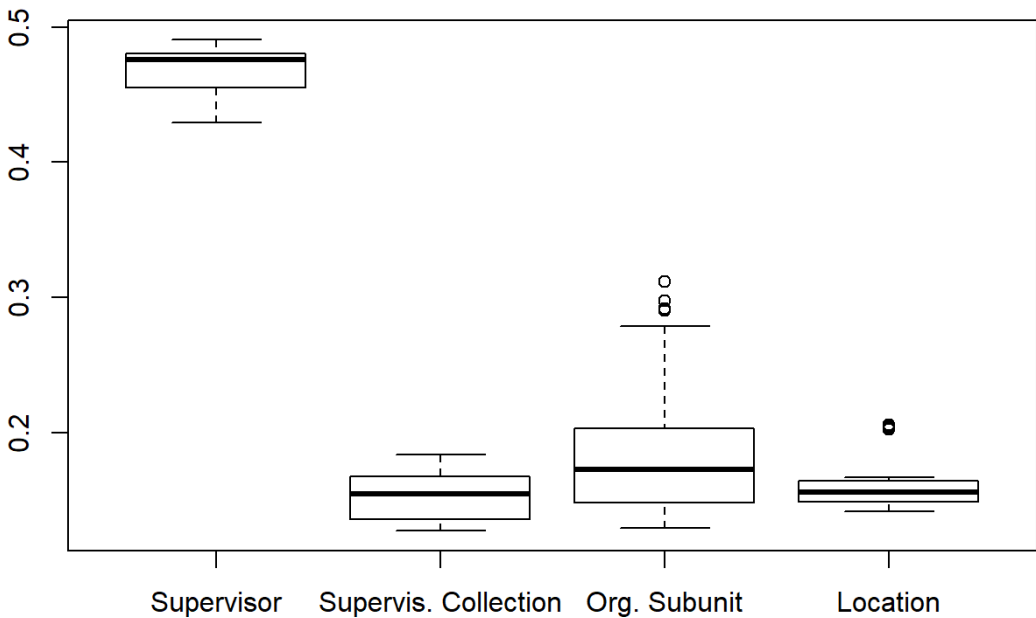| Covariate	N	Mean	St. Dev.	Min.	Median	Max.
Supervisor	60	0.467	0.019	0.430	0.476	0.490
Supervisor's Collection	60	0.151	0.016	0.127	0.154	0.184
Organizational Subunit	60	0.183	0.047	0.129	0.173	0.311
Location	60	0.164	0.021	0.142	0.156	0.205

Fig. 4. Box plot of correlations.

---

[1] We acknowledge that the effect size depends on the theoretical effect under study. However, to the best of our knowledge, we are the first to quantify this particular effect using digital traces, so no values for comparison are available.

5 DISCUSSION

5.1 Media Collections

Our results show that Exchange is prevalent in all media collections, supporting Watson-Manheim and Bélanger [7], who show that email is the most frequently used media tool in their study and relevant for all communication purposes. We find superior use of the information sharing collections compared to the relationship development collections. This corroborates the result by Watson-Manheim and Bélanger [7], who find that relationship development is a weaker theme compared to information sharing. Explicit relationship development in teams via digital communication and collaboration tools is less common than assumed, requiring future research in subsequent studies.

Lee et al. [4] find that smaller media collections have more users than larger collections. We also find a tendency towards smaller media collections with "Only Exchange" having the highest active user rate, although the results are not as clear as in Lee et al. [4]. Despite a heterogeneous media landscape (i.e., many different tools being available), the identified media collections clearly show Exchange and SharePoint as the most frequently used tools. Small media collections being favoured implies that the explicit management and coordination of media collection choice is not as critical as assumed because there seems to be no coordination overhead. Nevertheless, we only looked at Microsoft 365, and samples with a larger media landscape may yield varying results.

Lee et al. [4] and Watson-Manheim and Bélanger [7] made theoretical propositions based on a literature review and an interview study. We replicate their results with a unique quantitative sample of digital traces, thus, adding non-perceptual evidence to their theoretical propositions about media collections in the context of distributed work.

5.2 Antecedents of Media Choice

Social factors are theorized to influence workers' media choice. Behaviours of coworkers and supervisors, as well as espoused attitudes surrounding tool use, shape the workers' perceptions of media appropriateness and media collection choice [25].

According to Treviño et al. [29, 30], the supervisors' media behaviours and attitudes influence the media choice of individual workers in distributed work settings through verbal statements such as part of conversations, meetings, and collaborative work. The supervisors may also promote specific tools [27]. Our results corroborate the association between the assigned supervisor and a worker's media choice. Yet, our results do not substantiate the hypothesis that workers may imitate the media choice of their supervisor for joint communication purposes [25]. The supervisor has different tasks than the subordinates and thus may require a different media collection. Except for the imitation aspect, our analysis is agnostic to the specific behaviours of a supervisor that influence the subordinates' media choice.

Coworkers establish norms and values surrounding media collections through routine and joint use of tools. They shape the perceived task characteristics and media richness, which Schmitz and Fulk [27] find to influence the perceptions of appropriate media choice. Our data suggests that the coworkers do not have a clear association with the media choice of distributed workers. Again, our data does not provide rich insights into the behaviours of individual coworkers. Perhaps, the perceived task characteristics and attitudes towards media appropriateness are not sufficiently

homogenous within the organizational subunits to paint a clear picture. Intra-job differences and a potential lack of widely shared norms on tool use between coworkers may lead to different media collection choices [30].

Multiple studies find a relationship between physical location and media choice [29–31]. The physical location presents a constraint towards communication, as a high distance prevents face-to-face communication and encourages the use of digital tools. In our case, the workers are spatially distributed across 18 locations and we hypothesize that the physical location should have a negligible effect. Our analysis shows that the hypothesis of a negligible effect between physical location and media choice holds. A posthoc analysis shows that the media collection distributions across the locations all follow the general distribution of the organization, i.e., physical location does not explain any variance.

Summarizing, we find evidence in digital traces for the hypotheses that the supervisor has a strong effect and that physical location has a negligible effect on media choice. We replicate and substantiate previous survey-based studies on media choice in the context of distributed work and digital tools. The hypotheses regarding coworkers and the imitation of the supervisor are not supported.

6   LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Our analysis of supervisor and coworker influence is agnostic to their specific behaviours, actions, and attitudes. It does not disentangle dyadic influences and shared norms on the perceptions of task and media appropriateness (e.g., [29, 30]) because we only correlate the nominal supervisor and the organizational subunit with the selected media collection. Our results suggest that the influence of coworkers is not as homogenous as expected, i.e., not all coworkers share the same media collection. Potential causes may be a lack of shared norms on tool use or that our analysis misses intra job differences in the same way as other studies do (e.g., [30]). For an inquiry into the dyadic social influence of coworkers, previous studies built the ego-network of workers and administered surveys (e.g., [27]) or engaged in laboratory and experimental settings (e.g., [32]). The advent of enterprise social software provides digital traces on the dyadic relationships between coworkers. Beyond mere activity data on Yammer or email use, we encourage social network analysis based on dyadic relationship traces to investigate the coworkers' influence on media choice. Our study provides a cross-sectional view of the topic of media collections and media choice. With the ongoing development of digital collaboration tools, the media landscape is under continuous change, and longitudinal research designs may further elucidate the phenomenon.

In our study, we look at an idiosyncratic sample that describes a distributed organizational unit from a global services provider. Although Watson-Manheim and Bélanger [7] show that media collection types persist across two organizations, the identification of media collections and frequently used tools are specific to the task structures of the organization. Hence, we expect the nature of the task to influence a worker's choice of a media collection and our results may not generalize to other contexts with different task structures.

Beyond our hypotheses and analysis, previous research theorizes other factors to be relevant for media choice that are unavailable in the digital traces of Microsoft 365. For example, individual roles, strategies, and experiences as well as personal preferences are theorized to affect media collection choice. Given the limitations of the available data, we only analyse the influence of the

supervisor, coworkers, and physical location. Future work can extend the granularity of the digital traces and complement the analysis with interviews or surveys for more differentiated insights. The organizational unit under study has a strong focus on Microsoft 365. Nevertheless, external communication and collaboration tools that are not part of the Microsoft 365 suite may be in use by the workers, although such tools are not approved by the organization ("Shadow IT"). These tools are out of the scope of this study and require further data collection and research. For providing recommendations and best practices on the explicit management and coordination of media use, follow-up research should link performance data to the identified media collections. Further elucidating the link between media collections, media choice, and performance will expose levers for managerial interventions geared towards media synchronization and coordination. Since our approach is based on digital traces, it can be automatized and adapted into dashboards. However, caution is required because the analysed activities in the data set do not necessarily consume the same amount of time, for example, crafting an email may take longer than downloading a file.

In conclusion, we explore the media collections used by workers of a distributed organizational unit of a global service provider, in light of a heterogeneous and integrated media landscape of digital tools. In this empirical setting, we identify the frequently used media collections and address the antecedents of media choice in distributed work settings based on analysing a unique quantitative sample of digital traces. We contribute evidence to media collection research and replicate that information sharing is the primary purpose of media collections in our sample. Our analysis partially corroborates the existing theory on collective media choice, showing that the supervisor is associated with an individual worker's media choice, whereas the association with the physical location is negligible. Despite the limitations of our data, digital trace research shows prospects for subsequent inquiries. Through more fine-grained traces, researchers can substantiate and extend existing research on media collections and media choice in the future.

ACKNOWLEDGMENTS

REFERENCES

1. Wageman, R., Gardner, H., Mortensen, M.: The changing ecology of teams: New directions for teams research. J. Organ. Behav. 33, 301–315 (2012).
2. Mattern, J., Lansmann, S., Hüllmann, J.A.: It's not that bad! Perceived Stress of Knowledge Workers During Enforced Working From Home due to COVID-19. In: Proceedings of the 16th International Conference on Wirtschaftsinformatik (WI) (2021).
3. Bélanger, F., Watson-Manheim, M.B.: Virtual Teams and Multiple Media: Structuring Media Use to Attain Strategic Goals. Gr. Decis. Negot. 15, 299–321 (2006).
4. Lee, C.S., Watson-Manheim, M.B., Ramaprasad, A.: Exploring the Relationship Between Communication Risk Perception and Communication Portfolio. IEEE Trans. Prof. Commun. 50, 130–146 (2007).
5. Stephens, K.K.: The Successive Use of Information and Communication Technologies at Work. Commun. Theory. 17, 486–507 (2007).
6. Chudoba, K.M., Wynn, E., Lu, M., Watson-Manheim, M.B.: How virtual are we? Measuring virtuality and understanding its impact in a global organization. Inf. Syst. J. 15, 279–306 (2005).

7. Watson-Manheim, M.B., Bélanger, F.: Communication media repertoires: Dealing with the multiplicity of media choices. MIS Q. 31, 267–293 (2007).
8. Karr-Wisniewski, P., Lu, Y.: When more is too much: Operationalizing technology overload and exploring its impact on knowledge worker productivity. Comput. Human Behav. 26, 1061–1072 (2010).
9. O'Leary, M.B., Wilson, J.M., Metiu, A.: Beyond Being There: The Symbolic Role of Communication and Identification in Perceptions of Proximity to Geographically Dispersed Colleagues. MIS Q. 38, 1219–1244 (2014).
10. Woerner, S.L., Orlikowski, W.J., Yates, J.: The Media Toolbox: Combining Media in Organizational Communication. In: Proceedings of the Academy of Management Conference (2004).
11. Statista: Collaboration software market revenues from 2015 to 2023, https://www.statista.com/statistics/590412/worldwide-collaboration-software-market/.
12. Gartner: Gartner Says Worldwide Social Software and Collaboration Revenue to Nearly Double by 2023, https://www.gartner.com/en/newsroom/press-releases/09-24-2019-gartner-says-worldwide-social-software-and-collaboration-revenue-to-nearly-double-by-2023.
13. Markus, M.L.: Technology-Shaping Effects of E-Collaboration Technologies. Int. J. e-Collaboration. 1, 1–23 (2005).
14. Riemer, K.: E-Collaboration Systems: Identification of System Classes using Cluster Analysis. Interational J. e-Collaboration. 5, 1–24 (2009).
15. Riemer, K., Steinfield, C., Vogel, D.: ECollaboration: On the nature and emergence of communication and collaboration technologies. Electron. Mark. 19, 181–188 (2009).
16. Hüllmann, J.A.: The Construction of Meaning through Digital Traces. In: Proceedings of the Pre-ICIS 2019, International Workshop on The Changing Nature of Work (2019).
17. Hüllmann, J.A.: Smarter Work? Algorithmic Management in the Workplace Using Digital Traces. University of Münster (PhD Thesis) (2021).
18. Østerlund, C., Crowston, K., Jackson, C.: Building an Apparatus: Refractive, Reflective & Diffractive Readings of Trace Data. J. Assoc. Inf. Syst. 1–43 (2020).
19. Hüllmann, J.A., Krebber, S.: Identifying Temporal Rhythms using Email Traces. In: Proceedings of the America's Conference of Information Systems (AMCIS). , Salt Lake City, Utah, USA (2020).
20. Hüllmann, J.A., Kroll, T.: The Impact of User Behaviours on the Socialisation Process in Enterprise Social Networks. In: Proceedings of the Australasian Conference on Information Systems (ACIS). , Sydney, Australia (2018).
21. Johansen, R.: Groupware: Future directions and wild cards. J. Organ. Comput. 1, 219–227 (1991).
22. Fouss, J.D., Chang, K.H.: Classifying Groupware. In: Proceedings of the ACM Southeast Regional Conference. pp. 117–124 (2000).
23. Poltrock, S., Grudin, J.: CSCW, groupware and workflow. In: Proceedings of the Conference on Human factors in computing systems. p. 120. ACM Press, New York, New York, USA (1999).
24. Grudin, J., Poltrock, S.: Taxonomy and Theory in Computer Supported Cooperative Work. Oxford University Press, New York, New York, USA (2012).
25. Fulk, J., Schmitz, J., Steinfield, C.: A social influence model of technology use. In: Fulk, J. and Steinfield, C. (eds.) Organizations and communication technology. pp. 117–142. SAGE Publications Ltd, Newbury Park, CA (1990).
26. Fulk, J.: Social Construction of Communication Technology. Acad. Manag. J. 36, 921–950

(1993).

27. Schmitz, J., Fulk, J.: Organizational Colleagues, Media Richness, and Electronic Mail: A Test of the Social Influence Model of Technology Use. Communic. Res. 18, 487–523 (1991).

28. Stephens, K.K., Davis, J.: The Social Influences on Electronic Multitasking in Organizational Meetings. Manag. Commun. Q. 23, 63–83 (2009).

29. Webster, J., Treviño, L.K.: Rational and Social Theories as Complementary Explanations of Communication Media Choices: Two Policy-Capturing Studies. Acad. Manag. J. 38, 1544–1572 (1995).

30. Treviño, L.K., Webster, J., Stein, E.W.: Making Connections: Complementary Influences on Communication Media Choices, Attitudes, and Use. Organ. Sci. 11, 163–182 (2000).

31. van den Hooff, B.: Situational Influences on the Use of Communication Technologies: A Meta-Analysis and Exploratory Study. J. Bus. Commun. 42, 4–27 (2005).

32. Mennecke, B.E., Valacich, J.S., Wheeler, B.C.: The Effects of Media and Task on User Performance: A Test of the Task-Media Fit Hypothesis. Gr. Decis. Negot. 9, 507–529 (2000).

33. Schwab, A., Abrahamson, E., Starbuck, W.H., Fidler, F.: PERSPECTIVE—Researchers Should Make Thoughtful Assessments Instead of Null-Hypothesis Significance Tests. Organ. Sci. 22, 1105–1120 (2011).

34. Mertens, W., Recker, J.: New Guidelines for Null Hypothesis Significance Testing in Hypothetico-Deductive IS Research. J. Assoc. Inf. Syst. 21, 1072–1102 (2020).

35. Murtagh, F., Legendre, P.: Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? J. Classif. 31, 274–295 (2014).

36. Ward, J.H.: Hierarchical Grouping to Optimize an Objective Function. J. Am. Stat. Assoc. 58, 236–244 (1963).

37. Schwade, F., Schubert, P.: Social Collaboration Analytics for Enterprise Collaboration Systems: Providing Business Intelligence on Collaboration Activities. In: Proceedings of the 50th Hawaii International Conference on System Sciences. pp. 401–410 (2017).

38. Patefield, W.M.: Algorithm AS 159 : An Efficient Method of Generating Random R × C Tables with Given Row and Column Totals. J. R. Stat. Soc. Appl. Stat. 30, 91–97 (1981).

39. Cramér, H.: Mathematical Methods of Statistics. Princeton University Press (1946).

40. Sheskin, D.J.: Handbook of Parametric and Nonparametric Statistical Procedures. Chapman & Hall/CRC, Boca Raton (2000).

41. Cohen, J.: Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates, New York, New York, USA (1988).

42. Ellis, P.D.: The Essential Guide to Effect Sizes. Cambridge University Press, Cambridge (2010).

43. Mertens, W., Recker, J.: New Guidelines for Null Hypothesis Significance Testing in Hypothetico-Deductive IS research. J. Assoc. Inf. Syst. 21, 1072–1102 (2020).