Toward a Measure of Collective Digital Capacity: An Exploratory Analysis

Tawanna R. Dillahunt
tdillahu@umich.edu
University of Michigan, School of Information
Ann Arbor, Michigan, USA

Kerby Shedden
kshedden@umich.edu
University of Michigan, Department of Statistics
Ann Arbor, Michigan, USA

Mila Ekaterina Filipof
mfilipof@umich.edu
University of Michigan, School of Information
Ann Arbor, Michigan, USA

Soyoung Lee
soyolee@umich.edu
University of Michigan, School of Information
Ann Arbor, Michigan, USA

Mustafa Naseem
mnaseem@umich.edu
University of Michigan, School of Information
Ann Arbor, Michigan, USA

Kentaro Toyama
toyama@umich.edu
University of Michigan, School of Information
Ann Arbor, Michigan, USA

Julie Hui
juliehui@umich.edu
University of Michigan, School of Information
Ann Arbor, Michigan, USA

ABSTRACT

Digital training initiatives must shift toward critical cultural and social practices that encourage full participation in community affairs. However, no measure exists to account for digital capacity at the community level. Thus, we present this late-breaking work to begin designing and validating a measure of community digital capacity and report the results of an exploratory factor analysis. The analysis is based on 553 respondents across the United States to estimate an initial three-factor structure of (1) social digital capacity, (2) individual digital capacity, and (3) infrastructure. Such questions address limitations with existing theories that do not show digital inequities in the context of underlying systemic and structural challenges posed by one's social position. Our preliminary results suggest a potential measure for researchers and practitioners to understand whether people can access shared digital resources and activities with acceptable scientific guarantees, including favorable Akaike and Bayesian information criteria.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in collaborative and social computing**; **Collaborative and social computing**; • **Social and professional topics**;

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3650811

KEYWORDS

Digital literacy, Digital divide, Measure, Exploratory Factor Analysis, Community digital capacity, Social, Public Housing

ACM Reference Format:
Tawanna R. Dillahunt, Kerby Shedden, Mila Ekaterina Filipof, Soyoung Lee, Mustafa Naseem, Kentaro Toyama, and Julie Hui. 2024. Toward a Measure of Collective Digital Capacity: An Exploratory Analysis. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3613905.3650811

1 INTRODUCTION

Digital literacy is related to the capacity and ability to use digital resources to manage access, integrate, analyze, and synthesize digital information [19], and is considered a 'gate' skill often required by employers. Digital literacy is defined and measured in various ways depending on the scope and context. Earlier work by Martin and Grudziecki divides digital literacy into three levels: digital usage, competence, and transformation [21]. While Eshet-Alkalai separates digital literacy into six digital skills (real-time, digital reproduction, photo-visual, branching, digital information, and socio-emotional)[14], Phuapan et al. divide it into six indicators: managing, accessing, integrating, creating, evaluating, and communicating information to operate in a knowledge society [25]. Perspectives in other geographic contexts factor in as well, and bridging digital literacy has been extensively researched in low-resource contexts, particularly in the Global South [22, 23]. Digital literacy frameworks aim to empower individuals with technical skills, life skills, and access to digital services [23]. These frameworks share common components across regional contexts, such as information literacy, communication skills, and a focus on digital citizenship and safety [23]; however, it is important to tailor digital literacy initiatives to meet each region's specific needs and challenges [23]. For instance, while South Africa's National e-Skills Plan of Action prioritizes entrepreneurship and job skills to foster economic development and growth, Nedungadi et al. argues that frameworks in India should be tailored to the unique challenges that vulnerable populations in rural areas experience and provide essential digital literacy skills and awareness to participate in the digital world, emphasizing the importance of improved well-being [23]. Thus, many assessment instruments have emerged due to the broad definition of digital literacy [20], taking various factors into account—each measurement is likely to be partial to the researchers' direction. Lukitasari et al. call out that measurements do not reflect all characteristics of digital literacy and express concerns with instruments not aligning with local contexts because current digital literacy measurements are typically adopted and re-translated without proper construct and content validity tests [20].

Social networks could shape our digital literacy skills and relationship with technology in this work and context. In general, access to peer or social networks provides access to new information, resources [7, 10, 26], and even digital infrastructure (resources and accessibility). From the perspective of digital literacy, access to social networks could provide access to computer-related hardware or software, guidance, advice, and skills transfer [11, 12, 17]. Exposure to others with varying beliefs about technology and digital literacy could also contribute to a more nuanced perspective of technology. In other words, our networks could influence, motivate, or demotivate whether we learn, adopt, or resist new technologies. People in our networks could help teach and provide digital support when needed, and vice versa, influencing our level of individual digital capacity. Researchers in Human-Computer Interaction identified non-technical requirements such as social capital, social networks, and incubation from organizations as necessary to activate digital engagement in lower-income communities [13]. These scholars also found that community collectives comprising resource-connecting organizations foster informal opportunities for digital engagement and troubleshooting [16].

Despite the potential for our social networks to influence our digital literacy, to our knowledge, there are no explicit measures today taking such factors into account. What is especially concerning is how the COVID-19 pandemic further isolated individuals and exacerbated challenges around limited Internet access, and we have few measures, let alone interventions, that consider the impact of one's social network on digital capacity. Similar concepts include Intellectual Capital through Networks, which considers how tools, social networks, and media, with the appropriate techniques through these resources, can be used to integrate new insights into personal understandings and build collective intelligence [31]. However, existing measures do not take this factor into account.

Based on prior human-computer interaction research, our proposed concept of community considers shared resources and activities and social resources, such as the number of people a person could call upon to ask a basic technology question, to offset any limitations in an individual digital capacity. Our proposed concept of community digital capacity spans beyond an individual. In contrast, digital literacy, or digital capacity as a measure, solely refers to an individual's ability or capacity to use specific technologies. Building on prior research, this community perspective on digital literacy could provide a more accurate understanding of the digital capabilities and gaps of under-resourced populations more likely to rely on social support to bridge gaps in basic resources.

Thus, this late-breaking work develops such a measure and assesses whether this measure is meaningful. Building on insights in the field of human-computer interaction, past research on the digital divide, and results of an empirical study consisting of interviews and surveys, we conducted an exploratory factor analysis based on 553 respondents to estimate an initial three-factor structure of (1) social digital capacity, (2) individual digital capacity, and (3) infrastructure. Here, we tentatively conceptualize social digital capacity as the ability of individuals in the community to obtain help concerning digitally-mediated tasks; individual digital capacity as the magnitude and distribution of individual digital literacy within the community; and infrastructure as the physical and digital infrastructure available to the community. We propose a measure for researchers and practitioners to capture whether people can access shared digital resources and activities with acceptable scientific guarantees, including favorable reliability coefficients, using Cronbach's alpha.

2 METHODOLOGY

This section outlines the details of our survey questionnaire, our dissemination process to collect our data, and our factor analysis. We proposed four factors based on the prior research: resources and accessibility, individual digital capacity, social digital capacity, infrastructure, and digital assets/currency. Because we collaborated with a non-profit housing organization in Michigan, we included a question asking whether respondents were housing community residents. We compensated the public housing residents $10 for providing valid and complete survey responses. We drew most of our survey questions from existing scales (i.e., scales of individual capacity [5] and social networks [8, 9]).

Questions capturing individual digital literacy assessed their ability to perform a range of tasks subjectively ranging from easy to difficult (e.g., using a keyboard, sending a text message, recording a video, to more difficult tasks like using a spreadsheet) and their social digital capacity such as whether they could borrow devices from others if they did not own them, whether or not they knew someone who could teach them tasks that they needed to learn; and included questions about infrastructure and internet accessibility (e.g., questions of the availability and reliability of the internet and access outside of the home). To assess for infrastructure, survey questions also sought to capture general ownership of devices like smartphones and computers and their frequency of usage [27]. Measurement-related questions ended with digital-asset-related questions assessing social media use, amount of friends or followers, online credentialing, comfort posting a request for help or resources online, and the most funds raised with Internet support at one time. We concluded the survey with demographic information.

The university researchers worked with the community partner to clarify, add, and eliminate unnecessary questions. The goal was to keep the survey completion time to 10-15 minutes. We conducted an iterative series of cognitive interviews to identify problems with our survey questions [3] and correct them. Cognitive interviews ensured questions were understandable (content validity) and that they measured what we intended (face validity). Cognitive interview questions ranged from interviewee feelings about the overall length of the survey, how they decided to answer questions the way they did, to requests about what they thought questions were asking in their own words. We sought clarification about frequency—"Can you explain your understanding (about how many times) of 'Multiple times a week' but not 'Everyday' ?" After piloting our initial survey with four family members of one of the researchers, we followed these think-aloud interviews with respondents from the public housing community (i.e., those from our community of focus). Pilot survey respondents were unpaid; however, public housing respondents were paid $20 for each interview. Interviews spanned from 27 minutes to 1 hour and 40 minutes. On average, each interview lasted 85 minutes or 1 hour and 25 minutes (1.15 hours). We coded responses for "question comprehension, recall, and judgment" after receiving guidance from a survey expert. Our pilot cognitive interview results helped us identify questions that: (1) were difficult to understand, (2) required clarification (comprehension), and (3) were candidates for removal. For example, we learned that the question, "Who is the most powerful person you have connected with online," was unclear and saw uncertainty about what constituted something occurring "online." For example, "Does online mean Zoom?" "Do you mean connected through the Internet? What about messenger?" The initial think-alouds revealed that respondents faced challenges recalling questions inquiring about the number of Facebook friends or social media followers they had on specific sites, and we dropped these types of questions. Respondents suggested providing parenthetical examples for many questions that were unclear. Questions about digital assets/currency were eventually dropped and pilot participants identified typos and suggestions for making our survey more consistent.

Without providing too many details, we modified our survey accordingly, using the framework of common errors identified in cognitive interviews [1]: comprehension, retrieval, estimation/judgment, and reporting. For example, we clarified survey instructions (e.g., moving direction text such as "check all that apply" to the beginning of questions versus the need and integrated the definition of a smartphone into the question text. We added "flip/feature" phones to help describe what smartphones were not and clarified and changed words that could be interpreted in multiple ways (i.e., we changed the most "powerful" person to the most influential person). Our analysis resulted in 15 individual digital capacity items (3 Likert scale), 10 social digital capacity items (True/False), and 4 infrastructure items. (True/False). 4 of the 10 social digital capacity items were combined into 2 items because of dependency between the items. For example, the responses of Q6-1 were re-coded by combining the two categories of True/False so that high numbers indicate high capacity (See Table 3 in Appendix A).

2.1 Data collection methods and Analysis

We used multiple offline and online approaches to collect data in collaboration with our community partner, yielding 721 responses. Offline approaches included in-person attempts to obtain responses at offline community events and door-to-door. Our community partner also disseminated a link to the online version of the survey via SMS text message. We later learned the survey was advertised to online Facebook groups and spread to community members outside of the city our community partner was located. The average time to complete the survey for respondents completing more than 80% of the survey was 3.59 hours. However, some respondents left the survey and returned at a later time to complete their survey. Thus, this is not an accurate account of the average time to completion. Most respondents took less than 30 minutes to complete the survey.

We excluded any survey time to completion less than 100 seconds ("speeders"); respondents who entered the same answer multiple times ("straight lines"); those who completed less than 80% of the survey ("slackers"); and survey bots [4]. We could identify survey bots from invalid open-ended responses and determined locations based on self-reported zip codes. Using Tableau, the third author plotted zip codes onto a map of the United States. From there, she filtered respondents from our community partners based on the specific zip code of the state and community development. We removed respondents with fewer than 20 of our 29 items completed. We kept duplicate IP addresses for respondents from the public housing community as we entered in-person surveys from the same laptop, and there was a computer lab for participants to complete surveys. There were 553 valid survey responses. We present and discuss our early results of all valid survey responses (See Table 1 for respondent demographic characteristics). One hundred nine (N=109, 19%) respondents were from Michigan, the remaining were non-Michiganders (about 40% hailing from California), and only 56 valid responses were from the partnering community.

We turned to factor analysis, a multivariate statistical procedure considered the approach of choice for interpreting self-reporting survey results [6]. It is a method of data reduction to identify underlying latent variables and is divided into Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). We used EFA, a statistical technique commonly used in education, information systems, social science, and psychology, to determine the scale's dimensionality [28]. EFA allowed us to, as the name of the analysis suggests, explore the main variables to create a model identifying a small set of latent factors represented by a set of items [24]. EFA, in contrast to CFA, makes no assumptions about the factor structure [30]. Thus, EFA is most appropriate for scale development and applied when there is minimal theoretical grounding for prescribing the number and pattern of common factors [15, 18]. We conducted a factor analysis for all valid data using 28 items. The items were coded numerically and analyzed as quantitative variables in this analysis. We coded items such that greater numerical scores correspond to greater skill level, social support, or greater access to infrastructure. Factor loadings and scores were estimated using maximum likelihood analysis and were rotated to satisfy condition IC3 of Bai and Li[2]. We used Akaike and Bayesian information criteria (AIC and BIC) to assess model fit. While more analyses are needed, we next present the initial results of our late-breaking work.

3 ASPECTS OF COMMUNITY DIGITAL CAPACITY

Initial results of our proposed measure of community digital capacity are promising. Whereas most people talk about a single

		All Valid (n=553)	Public Housing Community (n=56)
Age	10s	16 (3%)	1 (2%)
	20s	247 (45%)	5 (9%)
	30s	175 (32%)	10 (18%)
	40s	53 (10%)	15 (27%)
	50s	25 (5%)	17 (30%)
	60s		4 (7%)
	70s		1 (2%)
		516 (93%)	**53 (95%)**
Gender	Male	295 (53%)	16 (29%)
	Female	215 (39%)	35 (63%)
	Non-binary		1 (2%)
	Self-describe		1 (2%)
		510 (92%)	**53 (95%)**
Race/Ethnicity	White	301 (54%)	16 (29%)
	Black	89 (16%)	40 (71%)
	Hispanic	60 (11%)	
	American Indian	48 (9%)	1 (2%)
	Native Hawaiian/Pacific Islander	16 (3%)	2 (4%)
	Asian	14 (3%)	
		528 (95%)	***59 (105%)**

Table 1: Respondent Demographics, *Participants could enter more than one race.

component of digital literacy, skills assessment, we begin to uncover the multidimensional aspects of digital literacy. We use the goodness of fit, internal consistency, and uni/multidimensionality to describe aspects of the proposed community digital capacity measure. A single item, *"Most of my time was spent in places that had Internet access,"* did not load with the other infrastructure items. Thus, we excluded this item from the proceeding exploratory factor analysis, which was informed by the results of the exploratory analysis. We also combined know_fewprob and know_allprob_given inconsistent responses. If a respondent knew anyone who could address any problem, we combined it as 1 instead of differentiating between knowing someone who could solve any problem versus knowing someone who could solve all problems. Thus, our final survey item count was 28 instead of 29.

We used the Akaika information criterion (AIC) and Bayesian information criterion (BIC) to evaluate the goodness of model fit. Based on AIC, the number of factors was estimated to be 8, and based on BIC, the analogous number of factors was estimated to be 4. To simplify interpretation, we focused on the two-factor fit. In all cohorts, all items loaded concordantly with the dominant factor (individual digital capacity), while both positive and negative loadings occurred for the second factor (social digital capacity).

We report Cronbach's Alpha statistics as a measurement of internal consistency. There were a total of 553 observations across 28 items (a=.91), the ability subscale consisted of 15 items (a=.93), the social subscale consisted of 9 items (a=.65), the infrastructure subscale consisted of 4 items (a=.46). The internal consistency for ability is high while the social and infrastructure items have moderate internal consistency.

In terms of dimensionality, although all items loaded positively on the dominant factor (i.e., individual digital capacity), the strengths of these loadings varied (See Table 2). In the full cohort of all valid responses, the (a_)bility/skill items had an r-squared exceeding 0.4 with the dominant factor, except for "spreadsheet," which had an r-squared of 0.37. The social variables had r-squared less than 0.1 with the dominant factor, and the infrastructure items had r-squared ranging from 0.07 to 0.21 with the dominant factor. Therefore, we interpreted the dominant factor as capturing overall capacity encompassing greater skills, more social support, and better access to infrastructure, but primarily reflecting abilities and skills.

The items a_spreadsheet, know_fewprob[1], know_anytask,know_provide, know_comeover, and know_place had correlations greater than 0.3, with the second factor in the cohort containing all valid responses. The items a_text, a_keyboard, and a_search had negative correlations of less than -0.2 with the second factor (i.e., social digital capacity) in this cohort. Therefore, we interpreted the second factor as indicating social support and skills associated with workplace computing. People with less social support and/or workplace computing skills were more comfortable texting and conducting internet searches.

We graphically displayed factor analysis results as a plot whose axes correspond to loadings of each item's first and second estimated factors (See Figure 1). Three colors were used to distinguish the three domains of items (orange: social digital capacity; green: individual digital capacity; and blue: infrastructure).

The EFA results of our public housing cohort, which are not shown for space limitations, align with the larger sample. To evaluate the measure's validity, we assessed how well our demographic results aligned with prior research. There were no sex differences in the factor 1 scores. However, increasing age and Black race were associated with lower scores, while white race, greater wealth, and

[1]Know_fewprob was collapsed with know_anyprob to indicate that "I know someone who can solve a few or almost all of the problems I have with my computer, smartphone, or flip phone/feature phone."

Table 2: Correlations and squared correlations between each item and each factor

	1st Factor		1st Factor (Squared)		2nd Factor (Squared)	
Item	1_load	2_load	1_r	1_r2	2_r	2_r2
1: a_text	0.708889	-0.356811	0.687137	0.472158	-0.103940	0.010804
2: a_keyboard	0.665078	-0.303902	0.657287	0.432026	-0.087418	0.007642
3: a_download	0.726077	0.034107	0.742840	0.551812	0.247502	0.061257
4: a_watch	0.700837	-0.161062	0.702607	0.493657	0.060464	0.003656
5: a_record	0.710481	-0.031935	0.738866	0.545923	0.190138	0.036153
6: a_openread	0.682043	0.025328	0.699668	0.489536	0.231990	0.053819
7: a_email	0.726119	0.213753	0.764162	0.583943	0.445869	0.198799
8: a_search	0.771713	-0.259884	0.754125	0.568704	0.004460	0.000020
9: a_calendar	0.673799	-0.029583	0.704866	0.496836	0.168758	0.028479
10: a_videocall	0.600930	0.257791	0.642536	0.412852	0.456747	0.208618
11: a_payment	0.753032	0.071231	0.775279	0.601058	0.300586	0.090352
12: a_purchase	0.730588	0.027321	0.751107	0.564162	0.250775	0.062888
13: a_socialmedia	0.732492	-0.151031	0.734185	0.539027	0.088751	0.007877
14: a_prvsetting	0.748831	0.026885	0.773219	0.597867	0.268797	0.072252
15: a_spreadsheet	0.557721	0.418179	0.614147	0.377177	0.645879	0.417160
16: k_once	0.025694	0.080692	0.030676	0.000941	0.109933	0.012085
17: k_daily	0.093517	0.265110	0.116361	0.013540	0.389349	0.151592
18 & 19: k_prob	0.240636	0.351688	0.277514	0.077014	0.554706	0.307698
20: k_teach	0.260704	0.190158	0.280762	0.078827	0.309566	0.095831
21: k_anytask	0.268734	0.317602	0.304299	0.092598	0.501295	0.251297
22: k_provide	0.180520	0.331671	0.208100	0.043306	0.452607	0.204853
23: k_loan	0.183030	0.117918	0.197530	0.039018	0.227486	0.051750
24: k_comeover	0.247730	0.302929	0.278160	0.077373	0.463146	0.214504
25: k_place	0.302501	0.230488	0.329765	0.108745	0.377248	0.142316
26: i_address	0.325735	0.143723	0.338174	0.114361	0.241129	0.058143
27: i_reliable	0.439494	0.110383	0.463320	0.214665	0.266270	0.070900
28: i_place	0.221843	0.503868	0.263357	0.069357	0.639134	0.408492
29: i_outside	0.437947	0.117030	0.459099	0.210772	0.276433	0.076415

greater education were associated with higher factor 1 scores. In multiple regression analysis, male sex, younger age, larger households, and greater education were associated with greater factor 1 scores. Wealth was no longer significantly associated with the factor 1 scores after controlling for other factors. Male sex, white race, larger households, and greater wealth were significantly associated with scores on the second factor. No variables remained associated with the factor 2 scores in the multiple regression analysis.

4 CONCLUSION

This late-breaking work proposes a novel approach to understanding and addressing digital inequality by introducing a new scale for measuring community digital capacity. We identified 28 items that measure the individual or ability (15 items), social (9 items), and infrastructure (4 items) expressions of digital capacity. The scale explores the crucial role of social and communal support in enhancing digital literacy, especially in the absence of strong individual skills, which is a relatively underexplored area in HCI. This work is important because it enriches digital inequality studies and offers practical insights for technology implementation and targeted training in underserved communities, with a broader goal of refining existing assessment methods and promoting inclusivity.

The internal consistency for individual digital capacity was high, and the validity met the goodness of fit criteria. In addition, aligning with prior research, we found that certain demographic factors (i.e., age, race, wealth, and education) were associated with differences in digital capacity scores, highlighting underlying social inequalities [29]. However, there are limitations considering the preliminary nature of our work. The social and infrastructure items only had moderate internal consistency, which suggests a potential weakness in the instrument and the need for future work. Thus, confirmatory factor analysis and assessment of construct validity are needed to ensure the robustness and generalizability of our proposed instrument. Along these lines, we emphasize that measuring community digital capacity in areas with low digital literacy is difficult. Time for in-person paper-based surveys must be considered in addition to providing options to complete surveys on provided devices with a stable internet connection. We acknowledge that expanding our research to include diverse urban and rural communities within and beyond the United States can provide more generalizable findings and insights. For example, the proposed instrument should also be implemented across other public housing communities in Michigan, other cities, and states across the United States. While our work focuses on a specific population within the United States, bridging digital literacy is also an extensive issue worldwide and

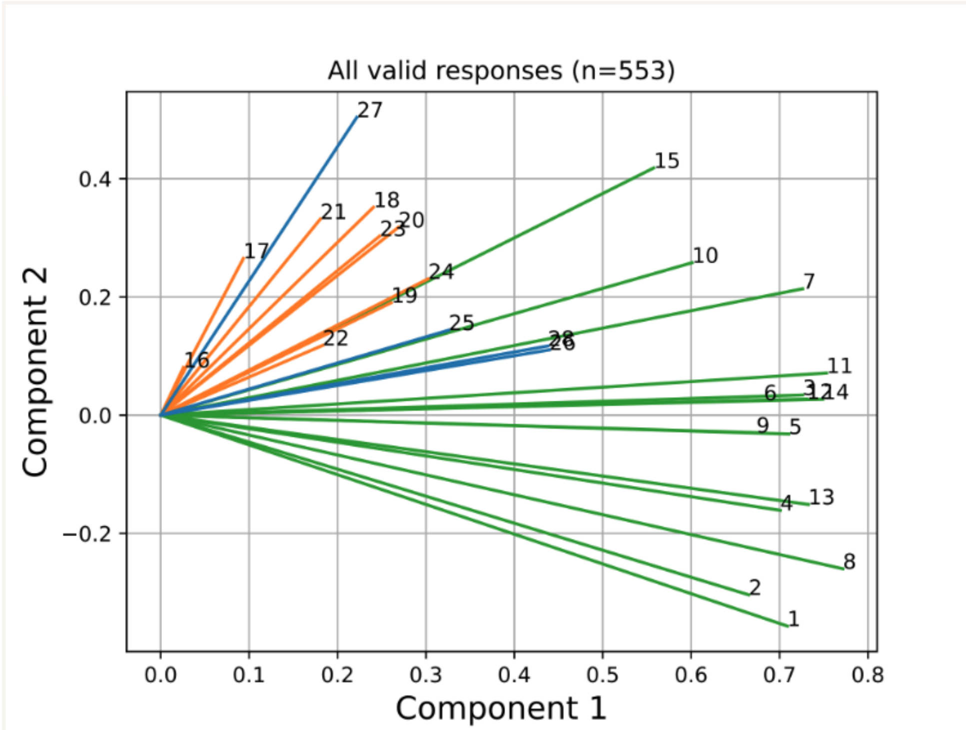

Figure 1: Exploratory Factor Analysis (EFA) Diagram with three components reflected in three colors: Green lines represent individual digital capacity or Component 1; Orange represent social digital capacity, or Component 2; and Blue lines represent infrastructure or Component 3.

particularly in the Global South (e.g., [22, 23]. We must look for opportunities in the future to connect efforts across borders. Future efforts should also be made to assess community digital capacity in rural communities. This will enable us to get more comprehensive results on community digital capacity and employ confirmatory factor analysis to assess specific hypotheses about the measure's psychometric properties. Future work must assess whether the measure's structure is similar across these demographic groups and locations. Determining the level of community digital capacity will reveal to what extent future funding efforts to bridge the digital divide should be focused.

ACKNOWLEDGMENTS

This work was supported by the NSF award CNS-2125012 and the first author completed this work as an MIT MLK Visiting Scholar. We thank Alexandra K. Murphy and our ISR colleagues Lisa Holland, Raphael Nishimura, and Wen Chang for providing early insights on this work. We thank our CEDER colleagues Inah Ko and Kolby Gadd for their early analysis and Shruti Sannon for reviewing an early draft and providing feedback on this work. We thank our anonymous reviewers for their supportive feedback. Finally, we thank ChatGPT for helping us to create accessible tables and figures.

REFERENCES

[1] Carol E Adair, Anna C Holland, Michelle L Patterson, Kate S Mason, Paula N Goering, Stephen W Hwang, and At Home/Chez Soi Project Team. 2012. Cognitive interviewing methods for questionnaire pre-testing in homeless persons with mental disorders. *Journal of Urban Health* 89 (2012), 36–52.

[2] Jushan Bai and Kunpeng Li. 2012. Statistical analysis of factor models of high dimension. (2012).

[3] Paul C Beatty and Gordon B Willis. 2007. Research synthesis: The practice of cognitive interviewing. *Public opinion quarterly* 71, 2 (2007), 287–311.

[4] Jacob Belliveau and Igor Yakovenko. 2022. Evaluating and improving the quality of survey data from panel and crowd-sourced samples: A practical guide for psychological research. *Experimental and Clinical Psychopharmacology* 30, 4 (2022), 400.

[5] Walter R Boot, Neil Charness, Sara J Czaja, Joseph Sharit, Wendy A Rogers, Arthur D Fisk, Tracy Mitzner, Chin Chin Lee, and Sankaran Nair. 2015. Computer proficiency questionnaire: assessing low and high computer proficient seniors. *The Gerontologist* 55, 3 (2015), 404–411.

[6] Fred B Bryant, Paul R Yarnold, and Edward A Michelson. 1999. Statistical methodology: VIII. Using confirmatory factor analysis (CFA) in emergency medicine research. *Academic emergency medicine* 6, 1 (1999), 54–66.

[7] Raj Chetty, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. 2022. Social capital I: measurement and associations with economic mobility. *Nature* 608, 7921 (2022), 108–121.

[8] Sheldon Cohen and Harry M Hoberman. 1983. Interpersonal support evaluation list. *Journal of Applied Social Psychology* (1983).

[9] Sheldon Cohen and Harry M Hoberman. 1983. Positive events and social supports as buffers of life change stress. *Journal of applied social psychology* 13, 2 (1983), 99–125.

[10] James S Coleman. 1988. Social capital in the creation of human capital. *American journal of sociology* 94 (1988), S95–S120.

[11] Brian Detlor, Heidi Julien, Tara La Rose, and Alexander Serenko. 2022. Community-led digital literacy training: Toward a conceptual framework. *Journal of the Association for Information Science and Technology* 73, 10 (2022), 1387–1400.

[12] Tawanna R Dillahunt, Vaishnav Kameswaran, Desiree McLain, Minnie Lester, Delores Orr, and Kentaro Toyama. 2018. Entrepreneurship and the socio-technical chasm in a lean economy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[13] Tawanna R Dillahunt, Vaishnav Kameswaran, Desiree McLain, Minnie Lester, Delores Orr, and Kentaro Toyama. 2018. Entrepreneurship and the socio-technical chasm in a lean economy. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* CHI '18 (2018), Paper 240, 14 pages.

[14] Y. Eshet-Alkalai. 2004. Digital literacy: a conceptual framework for survival skills in the digital era. *Journal of Educational Multimedia and Hypermedia* 13, 1 (2004), 93–106.

[15] James C Hayton, David G Allen, and Vida Scarpello. 2004. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods* 7, 2 (2004), 191–205.

[16] Julie Hui, Nefer Ra Barber, Wendy Casey, Suzanne Cleage, Danny C Dolley, Frances Worthy, Kentaro Toyama, and Tawanna R Dillahunt. 2020. Community collectives: Low-tech social support for digitally-engaged entrepreneurship. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), 1–15.

[17] Julie Hui, Kentaro Toyama, Joyojeet Pal, and Tawanna Dillahunt. 2018. Making a living my way: Necessity-driven entrepreneurship in resource-constrained communities. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–24.

[18] Amy E Hurley, Terri A Scandura, Chester A Schriesheim, Michael T Brannick, Anson Seers, Robert J Vandenberg, and Larry J Williams. 1997. Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of organizational behavior* (1997), 667–683.

[19] Siriwatchana Kaeophanuek, Jaitip Na-Songkhla, and Prachyanun Nilsook. 2018. How to enhance digital literacy skills among. *International Journal of Information and Education Technology* 8, 4 (2018), 292–297.

[20] Marheny Lukitasari, Wasilatul Murtafiah, Siti Ramdiah, Rusdi Hasan, and Akhmad Sukri. 2022. Constructing Digital Literacy Instrument and Its Effect on College Students' Learning Outcomes. *International Journal of Instruction* 15, 2 (2022), 171–188.

[21] Allan Martin and Jan Grudziecki. 2006. DigEuLit: Concepts and tools for digital literacy development. *Innovation in teaching and learning in information and computer sciences* 5, 4 (2006), 249–267.

[22] Devansh Mehta, Ramaravind Kommiya Mothilal, Alok Sharma, William Thies, and Amit Sharma. 2020. Using mobile airtime credits to incentivize learning, sharing and survey response: Experiences from the field. In *Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies*. 254–264.

[23] Prema P Nedungadi, Rajani Menon, Georg Gutjahr, Lynnea Erickson, and Raghu Raman. 2018. Towards an inclusive digital literacy framework for digital India. *Education+ Training* 60, 6 (2018), 516–528.

[24] Marjorie A Pett, Nancy R Lackey, and John J Sullivan. 2003. *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. sage.

[25] Piatip Phuapan, Chantana Viriyavejakul, and Paitoon Pimdee. 2016. An analysis of digital literacy skills among Thai university seniors. *International Journal of Emerging Technologies in Learning (Online)* 11, 3 (2016), 24.

[26] Robert D Putnam. 2000. *Bowling alone: The collapse and revival of American community*. Simon and schuster.

[27] Glenna L Read, Harry Yaojun Yan, Philip B Anderson, Laura PB Partain, Zachary Vaughn, Antonina Semivolos, Yeweon Kim, and Amy L Gonzales. 2022. Making stability dependable: stable cellphone access leads to better health outcomes for those experiencing poverty. *Information, Communication & Society* 25, 14 (2022), 2122–2139.

[28] HAMED Taherdoost, SHAMSUL Sahibuddin, and NEDA Jalaliyoon. 2022. Exploratory factor analysis; concepts and theory. *Advances in applied and pure mathematics* 27 (2022), 375–382.

[29] Alexander JAM Van Deursen and Jan AGM Van Dijk. 2019. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New media & society* 21, 2 (2019), 354–375.

[30] Brett Williams, Andrys Onsman, and Ted Brown. 2010. Exploratory factor analysis: A five-step guide for novices. *Australasian journal of paramedicine* 8 (2010), 1–13.

[31] Mark Wilson, Kathleen Scalise, and Perman Gochyyev. 2015. Rethinking ICT literacy: From computer skills to social network settings. *Thinking Skills and Creativity* 18 (2015), 65–80.

A QUESTION MAPPING

The question mapping in Table 3 includes the list of 29 Items (15 individual digital capacity items, represented by variable *a_*, 4 infrastructure items, represented by *internet_*, and 10 social digital capacity items, represented by variable *know_*). "Original name" corresponds to the final survey questions and the # on Plot align with EFA results in Figure 1.

Original name	# on plot	Renamed Variable	Question
		a_	**The header: "The next set of statements asks about your ability to do different tasks with a computer, smartphone, or flip phone/ feature phone. Choose how easy or difficult it is for you to do each task," applies to all a_variables**
Q5-1_1	1	a_text	Send a text message
Q5-1_2	2	a_keyboard	Use a keyboard (including one on a touchscreen)
Q5-1_3	3	a_download	Download an app
Q5-1_4	4	a_watch	Watch videos (movies)
Q5-1_5	5	a_record	Record a video
Q5-1_6	6	a_openread	Open and read emails
Q5-1_7	7	a_email	Send the same email to multiple people at the same time
Q5-2_1	8	a_search	Search for information online (example: using Google)
Q5-2_2	9	a_calendar	Enter events and appointments into a digital calendar (example: on Google, Outlook, or Apple Calendar)
Q5-2_3	10	a_videocall	Join a video call (example: on Zoom, Google Meet, or FaceTime)
Q5-2_4	11	a_payment	Make a digital payment to a friend (example: using Cash App, Paypal, Venmo, Zelle, or your bank)
Q5-2_5	12	a_purchase	Purchase an item online
Q5-3_1	13	a_socialmedia	Add, follow, or accept new friends/connections on social media like Facebook, Instagram, or Twitter
Q5-3_2	14	a_prvsetting	Change my privacy settings on social media so only my friends can see my posts
Q5-3_3	15	a_spreadsheet	Use a spreadsheet to keep track of things (e.g., contact list, taxes, budget)
		know_	**If I have difficulty doing something on my computer, smartphone, or flip phone/feature phone:**
Q6-1_1	16	know_once	I know someone who is willing to help me once in a while
Q6-1_2	17	know_daily	I know someone who is willing to help me on a daily basis
Q6-2_1	18	know_fewprob	I know someone who can solve a few of the problems I have
Q6-2_2	19	know_allprob	I know someone who can solve almost all of the problems I have
Q6-3_1	20	know_teach	I know someone who would patiently teach me how to learn a new skill
Q6-3_2	21	know_anytask	I know someone who would do any task on my behalf
Q6-3_3	22	know_provide	I know someone who would provide me with a computer if needed
Q6-4_1	23	know_loan	I know someone who would loan me their device
Q6-4_2	24	know_comeover	I know someone who would let me use their device at their place
Q6-4_3	25	know_place	I know a place where I could affordably use the Internet
		internet_	**The next statements ask about Internet accessibility. Choose True or False for each statement:**
Q8_1	26	internet_address	Internet service is available at my address
Q8_2	27	internet_reliable	There is reliable Internet connection where I live
Q8_3	28	internet_place	Most of my time was spent in places with Internet access
Q8_4	29	internet_outside	I can access the Internet outside my home

Table 3: List of 29 Items (15 individual digital capacity items, represented by variable *a_*, 4 infrastructure items, represented by *internet_*, and 10 social digital capacity items, represented by variable *know_*). Note that we ultimately combined know_fewprob and know_allprob (Q18 & Q19). "Original name" corresponds to the final survey questions, and the # on Plot aligns with EFA results in Figures 1.