Pre-presentation version

to be presented at the 57th Session of the International Statistics Institute, Durban, South Africa, 17-21 June 2009.

Please check **http://isi.cbs.nl/iamamember/index.htm** or www.jonathandonner.com/DonnerToyama_ISI2009prepub.pdf for updates concerning the formal proceedings

Persistent themes in ICT4D Research: priorities for inter-methodological exchange

Donner, Jonathan
Microsoft Research India, Technology for Emerging Markets Group
196-36 2nd Main, Sadashivnagar
Bangalore 560080 India
E-mail:jdonner@microsoft.com

Toyama, Kentaro
Microsoft Research India, Technology for Emerging Markets Group
196-36 2nd Main, Sadashivnagar
Bangalore 560080 India
E-mail:kentoy@microsoft.com

"Information and communication technologies for development" (ICT4D) is the name given to a range of activity which considers how electronic technologies can be used towards socio-economic development of developing communities worldwide. Typical technologies considered in ICT4D include PCs, the Internet, and mobile phones, and the domains of application are as broad as global development itself: public health, education, agriculture, microfinance, sanitation, political activism, and so on. As a field of research, the area is diverse, with the involvement of disciplines ranging from econometric analyses (what is the impact of Internet use on GDP?), to ethnography (how do users of mobile phones in poor communities differ from wealthy users?), to engineering (how to design PCs for use by multiple students simultaneously?), to experimental intervention (what is the effect of participatory video on agriculture extension?).

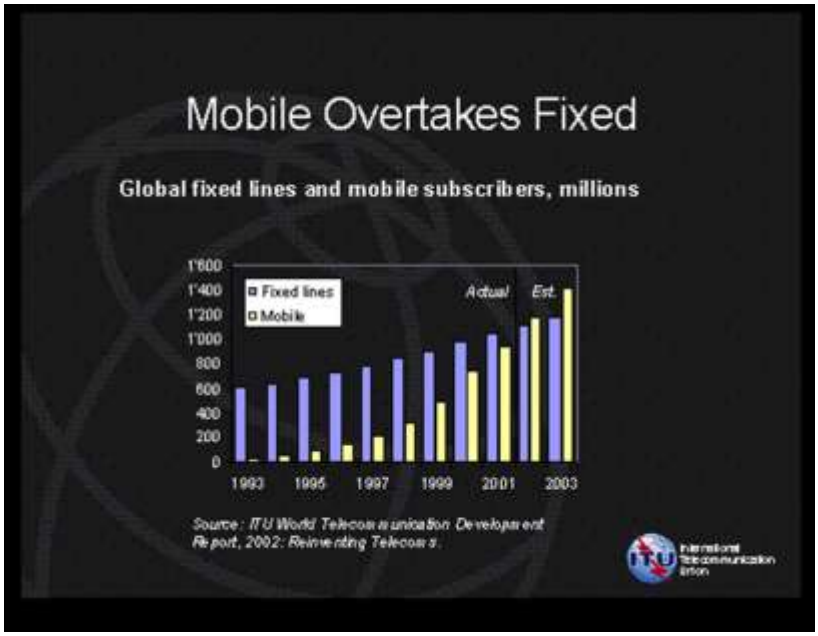
Given the unified goal of supporting global development, it is imperative for the progress of the field that such research efforts are complementary and cumulative, rather than siloed or oppositional. This means neither that some methods are 'better', nor that every research project should be mixed-method. However, efforts to bridge methodological and theoretical gaps are of particular value to the ICT4D field. In this paper, we highlight how certain persistent themes in ICT4D research are amenable to 'bridging' exercises. The themes are: defining users, closing divides, and

establishing impact. For each theme, we present a call for more actionable statistics and data, and suggest parameters for improved interdisciplinary dialogue. We close by considering new ways to leverage technology in these bridging exercises.

This is a paper written by two ICT4D researchers working in an interdisciplinary setting at Microsoft Research India. We find these themes cropping up in our individual projects and in moments when ICT4D interacts with broader technical, policy, and popular discussions.

"User" "Divide" and "Impact" are each basic terms, part of everyday life (Chaffee, 1996; Hempel, Dewey, & Neurath, 1952). Thus the temptation is to allow their common meanings to shape what should be our more careful and critical applications of the terms within ICT4D. These temptations not new—discussions of the inherent challenges of theorizing about sociotechnical phenomena are as old as social science itself. However, we argue that since the community concerned with ICT4D is *particularly and necessarily* interdisciplinary, the (mis)application of these basic and overarching themes both reveal and construct tensions within the field. This brief paper will not revisit methodological battles over "qual" vs. "quant", nor epistemological ones contrasting positivists vs. social constructionists; rather, it identifies points of commonality and mutual exchange on the path towards more useful, actionable ICT4D research.

Figure 1 – ITU Report: "Mobile Overtakes Fixed"



Defining users

We can start with the deceptively simple exercise of counting "users" of a technology. As the building blocks of assessments of the diffusion of a technology, such enumerations are often at the core of popular and policy pronouncements about the relative success or failure of a technology. For example, the International Telecommunication Union (ITU) issued a report in 2003 entitled "mobile overtakes fixed" (ITU, 2003). At the center of the report was a graph where the total number of subscriptions to mobile telephones crossed and (and was projected to surpass) the total number of subscriptions to landline telephones. The report captured an important moment and concept in the ICT4D arena.

The use of a 'baseline" of subscriptions which can be reviewed each year and contrasted across regions is probably the best overall indicator available to policy communities; it is illuminating, clear, and valuable. The ITU counts subscriptions, rather than users, because it can reliably gather these counts, year after year, from most of the countries on Earth. Unfortunately, it is hardly a secret that the numbers underpinning that fantastic graph are based on a series of aggregations and compromises, the net result of which is a great deal of ambiguity about what is actually being represented. Counting the easily countable—likely a number that most closely resembles the number of active mobile phone accounts or SIM cards— misses some of the more interesting figures that could be measured. As others (James & Versteeg, 2007; Sutherland, 2008) have pointed out, there are instances in which counting subscriptions can under or over-estimate the interactions of individuals with the mobile telephony system. In fact, the ITU statistics from 2007, the most recent year available, already indicate 50 countries with more subscriptions than people—penetration rates of over 100%. Counting subscriptions is not the same as counting people who have a subscription, nor handsets, active handsets, people with a handset, nor people who can make or receive a mobile phone call even without their own personal mobile account. (ITU, 2008).

Yet this choice to focus on subscriptions places telephone numbers, rather than telephone *users*, at the center of the policy discussion. Qualitative work reveals instances of SIM switching (Fjuk, Furberg, Geirbo, & Helmersen, 2008), phone sharing (Steenson & Donner, 2009), and the maintenance of multiple subscriptions, each of which hint at other metrics which may be equally interesting and useful to the policy and practitioner communities than the number of subscriptions. Bridging exercises, in which these insights are converted into better metrics and quantitative analyses, remain relatively scarce. For global estimates, these exercises could inform whether it is worthwhile to invest in assessments of other ICT metrics, such as the proportion of the population who has made a telephone call (ever or last week, or yesterday), or who currently has an active telephone number, fixed or mobile, at which they can be reached. At finer degrees of resolution, when examining technology use in smaller communities, workplaces, or households, there is even more latitude to deviate from the "counting subscriptions" model. Again, qualitative inquiry can help inform which indicators are most useful, and most representative of the factors we care about.

There are similar challenges when counting internet users, visitors to cybercafés and telecenters, and PC users. On the one hand, continued efforts to gather better survey data is welcome, particularly from regions were population sampling remains expensive and difficult. More creatively, quantitative research can continue to inform the development of indices which break down the use/nonuse dichotomy, replacing it with indices of intensity of use, ease and cost of access, and relative reliance on one channel vs. another. Each might paint a more nuanced and revealing picture of technology utilization use around the world. Similarly, use of some functions of a technology might be relevant to development than other functions. For example, the key predictor of productive telecenter use might be the possession of an email address, rather than proximity to or number of hours spent at a community telecenter. How might we estimate the proportion of the world with an active email address?

Closing Divides

A byproduct of a further refinement of metrics for use might be a re-framing of "non-use", and particularly of the popular concept of the "digital divide". The digital divide is perhaps the most powerful *popular* concept in the ICT4D arena. Its roots are in a report contrasting Internet users and nonusers in the United States in the mid 1990's (NTIA, 1995), and it has theoretical antecedents as far back as the knowledge gap (Tichenor, Donohue, & Olien, 1970). With time, however, the digital divide has become a shorthand reference for almost any situation where a one set of people are using or accessing less of a given technology than some other people would like them to be. In almost any situation where the penetration of users, as per the county exercise as described above, is less than 100%, policymakers, activists, and other interested members of the community will step in to identify a digital divide, whether the divide involves cell phones, Internet, broadband, or something else.

At core, the phenomenon which the digital divide seeks to describe is a quantitative state of affairs—X% with, Y% without—but it is the seductive strength of the metaphor, of the *divide*, and of the various initiatives required to span it, which can narrow the breadth and quality of an inquiry or a project, rather than improve it. The digital divide concept is prone to misapplication in two critical ways: first, it offers an easy conflation of cause and effect, pointing out an inequitable distribution and use of one resource without necessarily acknowledging the tapestry of other such imbalances. Frequently, the concept is associated with the notion that the divide leads to, or reinforces, inequities of various kinds, although many argue that the digital divide is a symptom, not a cause of the same inequities (Economist, 2005). Second, it tends to privilege technologies' use, rather than their usefulness. As measured, the divides often count users, not what the users do or gain from the technology. These misapplications are not universal; it is *possible* to structure and nuanced and productive inquiry around the digital divide, in the same way that it is possible to use the ITU's subscription statistics in powerful ways. However, nuance and reflection is not always the strong suit of policy, nor activism, nor innovation. If unchallenged, pursuit of the "digital divide" logic results in technologies looking for users.

In our own lab's work evaluating the value of telecenters and cybercafés in poor communities, we repeatedly found the digital divide logic vexing (Kuriyan & Toyama, 2007). Well-intentioned NGOs or microentrepreneurs had opened telecenters in order to bring access to PCs and the Internet to communities. But had the 'divides' been closed? If a school library gets a computer and opens its doors to the community, does that community have "access"? What if only men feel comfortable in the facility? Or people with certain skills or language abilities? And, what if, even for actual users, the ability to convert real-time information into real-world value were constrained due to limits of finances, education, or physical infrastructure? Qualitative inquiry is the key to unpacking the complexities of technology use in the rich contexts of the home, the community, the workplace. It is the key to identifying the interactions between skills, attitudes, literacies, dependencies, and desires which complicate popular conceptions of a digital divide waiting to be eliminated or bridged (Warschauer, 2003). Indeed, our own studies of telecenters suggests that at least in poorer environments, even when there are large numbers of telecenters theoretically providing access to thousands of potential users each, most telecenters are short on footfall, and even lighter on real impact (Kuriyan & Toyama, 2007). Qualitative evidence turns the causal connotations of the digital divide on its head: It's not that limited access to the Internet constrains the least educated from learning more; it's that the least educated are unable to take advantage of the resources of the Internet due to limitations of language, literacy, and formal education. The "democratizing" power of the Internet doesn't allow the poor to fight for their rights online; rather, an e-mail account can do much more for the wealthy and powerful than it can for the poor and politically unconnected. Telecenters do not bring medical care into poor villages through telemedicine; instead, the same economics that inhibits doctors from working in rural areas means that impoverished villagers have a harder time getting the attention of medical systems with wealthier clients elsewhere. Much other qualitative research echoes these findings (Sey & Fellows, 2009), and it would be valuable to see larger-scale studies confirm them.

Of course, as a concept separated from its causal implications, the "digital divide" retains value, if only as a shorthand to indicate one of the many differences between the rich and the poor, and there is a role for quantitative research to refine it. Careful application of advanced models can separate preconditions from outcomes; they can isolate pre-existing factors of socioeconomic status, explore unfolding dynamics in networked environments, and can generally improve the adoption models available to policymakers. That said, there may be more to be gained by steering quantitative research resources away from further enumerations of levels of use and breadth of divides, and instead towards better understandings of usefulness and hard metrics. As the next section describes, an understanding of the dynamics by which ICT's can be used to improve people's lives remains the *sine qua non* of the field.

Demonstrating Impact

The tiny "4" in ICT4D (the "for" in Information and Communications for Development") points to the implicit assumptions around which the majority of the field has coalesced—that there is a

positive impact associated with the deployment and use of technologies in resource constrained settings. However, there is work to be done, in order to convert as much of under-examined assumptions underlying that directionality with a clear understanding of what works, what does not, and why. We require, as many fields do, causal theories which describe the ways in which certain combinations of technologies can be deployed in ways that help people, we don't need one size fits all models, but we need theories of the middle range (Merton, 1949) which can help guide action, by NGOs, by policymakers, by individuals.

This is no small feat. Early efforts in communication research, for example, were criticized for drawing on linear, reductionist "transmission models", presuming strong, observable effects in what has turned out to be a complex, noisy world full of effects at cross-purposes. Indeed, some would argue, the effects paradigm itself misses the point, and that questions of how people instead use technologies to make sense of (and create) meaning in a social world are more important (Carey, 1988). While the complementary voices within ICT4D, are important reminders about context and contestation, about power and perspective, the "4" in ICT4D—the search for impact—remains central. We convene panels on "best practices", debate policies for their usefulness, and always, it seems, are looking for technologies or solutions that can deliver "impact" at large scales and low costs.

These debates are important. Thus, the field needs data, and theories to make sense of the data, in order to justify, design, and guide policy and technology interventions. In our own work, we find that peeling back the layers of the onion to understand impact and root causes to be a perennial challenge. For example, when we were looking at the impact of mobile-banking services on households in Kenya, we found that m-banking users were transferring money to relatives more frequently, but in smaller amounts. Evaluating the impact of the new m-banking service on households required us to understand how households seek to smooth consumption and expenditures over times, and how the same amount of money delivered in small but can (at times) before valuable than the same amount of money coming infrequently. Through a mix of qualitative and quantitative inquiry, we settled on a definition of impact which focused on the variability of inflows and outflows, rather than exclusively on their magnitude (Ratan, 2008).

Assessing impacts of ICT programmes or interventions can be a particularly challenging. Unlike the search for medical cure for a terminal disease, where discrete treatments can be evaluated by their ability to save discrete lines, ICT interventions frequently involve the introduction of bundles of capabilities and behaviors into complex social or organizational milieu. Perhaps unsurprisingly, these interventions yield a range of benefits, some tangible some intangible, some proximate, some distant, some expected, some unexpected (Kaplan, 2001).

On the inputs side (e.g., interventions and independent variables), practitioners and advocates often introduce a technology (along with training protocols, subsidies, official pronouncements, etc) without a clear model of precisely how that technology will bring about desired change. Substituting

a PC for a clipboard in rural clinic may reduce the error rate in capturing medical records, but it is likely to also influence organizational budgets, staff burden (time of input), morale, etc. Meanwhile, on the outputs side (e.g., impacts and dependant variables), the goals of ICT4D interventions are also often ill defined. Indeed, closing the digital divide is sometimes a goal onto itself. Telecenter projects frequently seek to provide a "bouquet of services" in microfinance, agriculture, healthcare, and education, all the while profiting the telecenter operator through income or increases in prestige and self-confidence. Post-hoc and rolling assessments of impact are common.

Qualitative inquiries help us refine our models of casual change, identifying what changes, and why, and which changes matter enough to be considered "impact". Bridging exercises work these qualitative insights back into statistical tests for replication and refinement. In these cases, the discipline imposed by the development and tracking of quantitative metrics is helpful, because it provides some counter-balance to the qualitative temperament to keep looking for additional impacts.

Randomized trials and rigorous experiments are one way to demonstrate impact, but some of the greatest contributions from the quantitative branch of the ICT4D field in the next few years will undoubtedly be with messy and multivariate data, from field experiments and quasi-experiments, from structural equation models and other ways to make the most of poor and incomplete data. Theories of the middle range (and a strong conversation between qualitative and quantitatively minded) will be the glue to hold this endeavor together.

Discussion

We have discussed the continued importance to ICT4D of bridging exercises, which combine (or at least sequence) qualitative and quantitative inquiry. Frankly, most qualitative work demands time in the field, with users or would-be users, understanding the contexts in which technologies are used, and how these contexts and demands may differ from how technologies are intended to be used (Heeks, 2002) . There are neither short cuts nor substitutes for listening, observing, testing, and exploring.

That said, it is a curious property of the study of ICTs that the artifacts themselves often produce useful data, which can be analyzed to yield insights about the ICT or pilot in question. Quantitative versions of these data streams are relatively well known. In our own work, we have used mobile handset call logs to provide evidence of who has called whom (Donner, 2006), and PC based server logs to track who is using a pilot system, and for what purposes (Veeraraghavan, Yasodhar, & Toyama, 2007). We were surprised in the latter case (a mobile phone based system to check agricultural information) to see people unconnected to the pilot project accessing the system. Word of mouth had brought news of the utility of the system to a proximate population. This outcome illustrates challenges of counting users, and discerning between users from non-users.

The use of less structured, qualitative data emerging from these same systems is less common, and worth discussing here. In other work in our lab, researchers have used the user-generated

outputs of ICTs in qualitative ways, to inform our assessments of the utility and impacts of the technologies. In India, a project called Kelsa+ explores how users can self-teach and self-direct their exploration of PCs and the internet, on a PC used in free time at their place of employment (Ratan et al., 2009). During the project the researchers logged the evolution of applications over time (a quantitative task) but also noticed and particular events from a qualitative lens. For example, on a number of occasions, users replaced the screen saver page on the Kelsa+ machine in our lab with pictures of their own faces, taken with the webcam. There is no way to anticipate a coding scheme that would account for that, but that moment, captured on the device under study, revealed a great deal about the symbolic and practical significance of the PC to these new users. In South Africa, we are exploring how minimal training and a strong community can support semi-literate and inexperienced users as they venture on to the internet for the first time via their cell phones (Gitau, Donner, & Marsden, 2009). In each case, we can look at both the patterns and anomalies coming from use. For example, the email that arrived to the trainer, from one of the women trainees, unsolicited and well outside the parameters of the training sessions, is a strong a marker of appropriation and the success of the training efforts.

Conclusion

This brief paper has revisited some core themes in the field of ICT4D: counting *users*, closing *divides*, and assessing *impact*. We argue that further counting and dividing may do less to push the field of ICT4D forward than further attempts to improve the quality of impact assessments, often under condition of poor and absent data. With every bit of new evidence matching the use of technologies to improvements in livelihoods, well-being, happiness and health, we can do more to guide action and marshal resources. We can even use qualitative and quantitative data from the ICTs themselves to help gather this evidence.

We close with some examples of important recent work isolating and refining impact assessments. At the micro level, Jensen's (2007) study capturing the impact of the introduction and use of mobile phones on the price of fish (it reduces the variance in prices paid, and reduces waste) is both elegant and rigorous. At the macro level, Waverman, Meschi, and Fuss (2005) associate higher levels of mobile use with faster per-capita GDP growth, using simultaneous equations to first control for the impact of GNP growth on the use of mobiles. There are numerous needs for similar studies at micro and macro levels about PCs, internet subscriptions, TVs and so on. If we listen first and use qualitative-quantitative bridging exercises to move beyond the 'common' aggregates of users and divides, we can better assess the impact of telecenters on communities, of phones on business, of laptops on children, and so on. In this effort, we look to the members of ISI and beyond.

REFERENCES

Carey, J. W. (1988). *Communication as Culture: Essays on Media and Society*. New York: Routledge.

Chaffee, S. H. (1996). Thinking about theory. In M. B. Salwen & D. W. Stacks (Eds.), *An integrated approach to communication theory and research* (pp. 15-32). Mahwah, NJ: Lawrence Erlbaum Associates.

Donner, J. (2006). The use of mobile phones by microentrepreneurs in Kigali, Rwanda: Changes to social and business networks *Information Technologies and International Development, 3*(2), 3-19.

Economist. (2005, 12 March). The real digital divide. *The Economist, 374,* 11.

Fjuk, A., Furberg, A., Geirbo, H. C., & Helmersen, P. (2008). New artifacts – new practices: Putting mobile literacies into focus. *Nordic Journal of Digital Literacy (Digital Kompetanse), 3*, 21-38.

Gitau, S., Donner, J., & Marsden, G. (2009). *Circumventing the gender/technology/economic disparity: A study of first-time mobile internet use among low-income women in South Africa*. Paper presented at the Digitally Empowering Communities: Learning from Development Informatics Practice (IDIA2009).

Heeks, R. (2002). Information Systems and Developing Countries: Failure, Success, and Local Improvisations. *The Information Society, 18*(2), 101-112.

Hempel, C. G., Dewey, J., & Neurath, O. (1952). *Fundamentals of concept formation in empirical science*. Chicago: University of Chicago Press.

ITU. (2003). Mobile overtakes fixed: implications for policy and regulation. Retrieved November 1, 2007, from http://www.itu.int/osg/spu/ni/mobileovertakes/Resources/Mobileovertakes_Paper.pdf

ITU. (2008). Online statistics, ICTeye. Retrieved 25 September, 2008, from http://www.itu.int/ITU-D/icteye/Indicators/Indicators.aspx

James, J., & Versteeg, M. (2007). Mobile phones in Africa: how much do we really know? *Social Indicators Research, 84*(1), 117-126.

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *Quarterly Journal of Economics, 122*(3), 879-924.

Kaplan, B. (2001). Evaluating informatics applications--some alternative approaches: theory, social interactionism, and call for methodological pluralism. *International Journal of Medical Informatics, 64*(1), 39-56.

Kuriyan, R., & Toyama, K. (2007). *Review of Research on Rural PC Kiosks*. Bangalore: Microsoft Research India.

Merton, R. K. (1949). *Social Theory and Social Structure*. Glencoe, IL: Free Press.

NTIA. (1995). *Falling through the net: A survey of the "have-nots" in rural and urban America*. Washington, DC: United States Department of Commerce.

Ratan, A. L. (2008). *Using technology to deliver financial services to low-income households: a preliminary study of Equity Bank and M-PESA customers in Kenya*. Bangalore: Microsoft Research.

Ratan, A. L., Satpathy, S., Zia, L., Toyama, K., Blagsvedt, S., Pawar, U. S., et al. (2009). *Kelsa+: Digital Literacy for Low-Income Office Workers.* Paper presented at the 3rd IEEE/ACM International Conference on Information and Communication Technologies and Development, Doha.

Sey, A., & Fellows, M. (2009). *Literature Review on the Impact of Public Access to Information and Communication Technologies.* Seattle: University of Washington Center for Information & Society.

Steenson, M., & Donner, J. (2009). Beyond the personal and private: Modes of mobile phone sharing in urban India. In S. W. Campbell & R. Ling (Eds.), *The Reconstruction of Space and Time: Mobile Communication Practices* (pp. 231-250). Piscataway, NJ: Transaction Books.

Sutherland, E. (2008). *Counting mobile phones, SIM cards & customers*. Wits, South Africa: LINK Centre, University of the Witwatersrand.

Tichenor, P. J., Donohue, G. A., & Olien, C. N. (1970). Mass media flow and differential growth in knowledge. *Public Opinion Quarterly, 34*, 159-170.

Veeraraghavan, R., Yasodhar, N., & Toyama, K. (2007). *Warana Unwired: Mobile Phones replacing PCs in a rural sugarcane cooperative.* Paper presented at the 2nd IEEE/ACM International Conference on Information and Communication Technologies and Development (ICTD2007), Bangalore, India.

Warschauer, M. (2003). *Technology and Social Inclusion: Rethinking the Digital Divide*. Cambridge, MA: MIT Press.

Waverman, L., Meschi, M., & Fuss, M. (2005). The Impact of Telecoms on Economic Growth in Developing Nations. Retrieved November 1, 2007, from http://www.vodafone.com/etc/medialib/attachments/cr_downloads.Par.78351.File.tmp/GPP_SIM_paper_3.pdf