Influences on the Scaling of Digital Learning Resources

Barbara Means and Vanessa Peters

SRI Education

Paper presented in the symposium "Scaling Up Digital Media Innovations" at the annual meeting of the American Educational Research Association, Washington, D.C., April 2016. The previously published empirical work described in this paper was funded by the Bill & Melinda Gates Foundation, but the views described herein are those of the authors and do not necessarily represent the policy or position of the foundation.

Influences on the Scaling of Digital Learning Resources

By enabling rapid dissemination and bypassing system hierarchies, the Internet has disrupted the prevailing linear relationship between learning product maturity and scale. Many technology developers embrace the strategy of getting a "minimum viable product" into the hands of many users quickly and then improving the product through user feedback.

In the education field, one of the best-known examples of a rapidly spreading digital learning resource is the Khan Academy, which grew from a few YouTube videos to having a million users in less than four years (Murphy et al., 2014). Backed by significant funding from donations, the Khan Academy now offers more than 5,000 online courses across the disciplines. Such examples have motivated the reconsideration of the nature of scaling of education interventions (Coburn et al., this symposium).

In the past, we have argued that the Internet provides an alternative to top-down selection of educational approaches and resources that can fuel bottom-up innovation (Means et al., 2004). But this paper will explore some of the limitations to that capability within the context of formal education systems. Clearly, there is a growing supply of digital learning platforms, instructional materials, and applications. But the question is how does rapid spread relate to product effectiveness and deeper impacts on education systems? In this paper we will argue that it is no coincidence that the prime examples of scaling technology-supported innovations such as Khan Academy and the youth programs described by Santo and Penuel and their respective colleagues in this symposium occurred outside of school time.

In a logical world, one would expect effective technology-based interventions to scale faster than ineffective ones. Proponents of the "minimum viable product" strategy argue that rapid scale generates feedback that is used to improve the product quickly and continually. With hundreds of thousands or even millions of users, the argument goes, there is a wealth of user-generated data to guide product improvements (see Münch et al., 2013; Scott, 2012). And indeed, we have observed the advantages of this approach in terms of improving product usability and functions (Murphy et al., 2014). But when it comes to educational impact, the research literature on both technology-supported and other educational interventions gives grounds for caution.

We are unlikely to see consistent positive impacts of a significant magnitude for any intervention that does not address core teaching and learning processes (Cohen & Ball, 1999). At the same time, any intervention that attempts to change those processes is more difficult to implement because it requires much more from the various actors in the education system (Coburn, 2003). For these reasons, one could make a case that potentially effective technology-based interventions will be harder to scale than those that affect only a few surface-level aspects of teaching and learning.

Defining a Technology Innovation

Understanding both learning impacts and the spread of digital learning innovations requires clarity about what we mean when we talk about a technology-based or a technology-supported innovation. Although a piece of hardware or new software often is the most tangible component of these innovations, it is how these resources are actually used along with other, often nondigital, resources in instruction that matters (Cohen, Raudenbush, & Ball, 2003). Educational effectiveness is the product of the broader instructional activity system, of which software resources are just one component. The role of the software in instruction (whether as supplemental practice or the central conveyor of content, whether confined to a short-term unit or used throughout the year, whether and how much student performance with the software is considered for grading, and so on) and the way in which the teacher introduces it, provides time for its use, coaches students as they work with the software, and draws connections between technology-based and non-technology activities are all part of the student's experience and can all affect learning and achievement outcomes.

Defining What Scaling an Innovation Means

A related issue is clarifying what we mean when we talk about scaling a technology-based innovation. For most digital learning technology developers, investors, and the general public, scaling means the number of users. However, Coburn's (2003) reconceptualization of scaling suggests that the number of users merely scratches the surface of what it means for an intervention to scale. In a similar vein, if one accepts our inclusive definition of what is being scaled as a technology-supported innovation rather than a technology product per se, scaling becomes more complex because of all of the necessary components of the innovation that need to be defined and implemented. Typically these include the roles and practices for instructors and students, the amount and distribution of time to be spent using the technology, training for the instructors and any others who will implement the innovation, aligned measures of learning, and ongoing pedagogical and technical support systems. In practice, many technology vendors do not attempt to develop a full theory of action for interventions incorporating their technology, deferring to educators on aspects of the intervention other than hardware and bandwidth requirements and basic training in the nuts and bolts of using their product. Any comprehensive description of how the product should be used and supported would reduce the number of schools interested in using it, a reduction that marketing departments are reluctant to accept.

But even if one does adopt the more limited definition of scale as the number of students, instructors, or schools using the digital learning product, the data provided by developers and vendors is often highly misleading. For a web-based resource, the easiest numbers to report, which are also the most expansive, are the number of downloads or unique user identifications. Products seeking massive scale want to make accessing their product as easy as possible, and initially may not even want to ask people coming to their web site to register for fear of losing potential users. Even when registrations are required, vast numbers of people typically register just to take a look at a learning product and never really engage with it. The discrepancy between the huge number of MOOC enrollments and the much smaller number of people who actually complete one of these courses, or for that matter, even complete a single assessment or assignment is well known (Jordan, 2014). In our experience, other kinds of learning platforms and applications have similar gaps. In a study of the popular Duolingo website (www.duolingo.com) for second-language learning (Vesselinov & Grego, 2012), for example, only 11 percent of the users who had received $20 for taking an online survey and who volunteered to be in the study actually spent the stipulated 30 or more hours on the Duolingo website. Being able to report higher user numbers enhances a product's ability to attract investment and new adopters so there is little incentive to publicize more nuanced figures on substantive product use.

The Tension Between Transformation and Scaling Goals

Business studies of product growth rates (Kutcher, Nottebohm, & Sprague, 2014; Rodgers, 1962) emphasize the importance of rapid growth and identify key product features that make it happen: perceived advantage relative to the available alternatives, compatibility with existing norms, ease of trying out the product, and visibility of the results of using the product. In other words, products that scale rapidly fill an unmet need of many individuals at low cost relative to value and with high convenience. It is for this reason that learning technology companies feel the pressure to offer at least some basic version of their product for a low price. Keeping the purchase price low for a product that costs a lot to develop is risky financially, of course, but it can be made to work if there is a sufficiently large number of buyers. Certainly, this strategy of gaining market share first and pursuing sustainable profit margins second has been tried by some notable learning technology companies (e.g., Coursera).

The compatibility with existing norms and ease of trying out criteria, on the other hand, are fundamentally at odds with the goal of transforming teaching and learning within an education system. Something is convenient if it is readily at hand and is easy to use, requiring little or no learning to start using it. Arguably, many of this century's most successful personal computing devices owe their popularity to this ease of use factor. For casual users, the instructional manuals that used to accompany computers and software packages are largely a thing of the past. A convenient product does not require new learning on the part of users. This attribute is in direct conflict with the nature of potentially transformative education innovations. You cannot fundamentally change education without changing teaching and learning of core content, and fundamental changes require adaptations and new learning on the part of teachers, students, and school leaders (Cobb & Jackson, 2012). By definition, fundamental change of teaching and learning means incompatibility with existing norms and difficulty, not ease, of trying out the new approach.

Thus, aspiring learning technology innovators face a dilemma: Should they forego dramatic change and design for school systems and classrooms as they are today or design for a transformative vision of what schools and classrooms could be? The former strategy is conducive to rapid adoption but not to major impacts; the latter strategy could have profound effects, but only if it manages to be adopted widely enough to affect whole systems and not just the classrooms of a few exceptional teachers.

Need to Involve the Education System in Technology-based Innovations

For developers of technology-supported innovations who have chosen the latter goal of transformative impact on classroom learning, engaging with multiple levels of the education system is imperative (Fullan & Donnelly, 2013). If teachers require time to learn about the deeper principles underlying the innovation and to acquire a repertoire of new practices, that time needs to be identified and paid for somewhere within the work week—a decision that must be made at higher levels of the education system. Similarly, if teachers are going to assume a different kind of role within their classrooms to provide students with more agency in their learning, school administrators cannot undermine that role by evaluating teachers based on criteria developed for teacher-led instruction. If the innovation calls for concentrated periods of activity that take more than an hour, the school schedule needs to accommodate adequate blocks of time. If the technology-based innovation requires high-bandwidth access to web-based systems, the school system needs to provide that level of access reliably and to troubleshoot any technology incompatibilities. In short, technology-based innovations in education require larger accommodations on the part of multiple players in an organization than does an individual's decision to purchase a smartphone or start using a fitness app; theories and models that describe scaling of direct-to-consumer technology products may be a poor fit for the education sector.

Data Sources and Methods

Wave I of the "Next Generation Learning Challenges" (NGLC) hosted by the nonprofit organization EDUCAUSE provided an unusual opportunity to examine empirically the relationship between digital learning technologies' scale and their educational impact. The grant opportunity announcement released in fall 2010 sought "sustainable, broad-scale, technology-enabled" products and services capable of "transforming education and improving student outcomes." After receiving more than 600 initial proposals, EDUCAUSE invited 50 respondents to submit full proposals and funded 29 projects starting in spring 2011 and extending through 2013. The recipient organizations were diverse, but predominantly included colleges or college systems. Exceptions were several for-profit learning technology companies and several consortia of multiple education systems. Concurrent with this grant term, SRI Education undertook evaluation activities to determine the extent to which the funded projects achieved their expected results in terms of scaling and student outcomes.

SRI classified the projects within the NGLC portfolio into five general categories:

- *Whole-course models* (for example, an instructional approach and curriculum for developmental math)
- *Supplemental resources* (for example, supports for using technology to promote inquiry-based science learning units)
- *Supports for course redesign* (such as resources and training for faculty who wish to redesign their courses to use blended models of instruction)
- *Learning analytics and early warning systems* (for example, a system that uses performance data from a learning management system to flag students who may be at risk of failing the course)
- *Supports for peer learning* (for example, programs for student-led study groups)

We also coded each digital learning resource using a set of product design and implementation model features (see Means et al., 2014 for a description of the full set of coding categories).

The NGLC projects varied considerably in their scaling strategies. Some of these projects aimed to get to students and instructors by getting institutional commitments that would lead to top-down uniform adoption. Others sought to go directly to either faculty or students, using a "retail" scaling model.

The scale achieved by each of the digital learning products during the NGLC grant period of roughly two years was obtained from project reports submitted to the funder. In addition, 22 of the projects provided student outcome measures for students experiencing their intervention and for students taking the same or a very similar course without the digital learning intervention. SRI conducted meta-analyses on these measures of effectiveness using effect size estimates for the binary outcome of earning a course credit and for whatever continuous learning outcome measure (typically either grade or assessment score) a project could provide for treatment and comparison students. Moderator variable analyses identified factors associated with more positive learning outcomes. Correlational analyses explored the relationships between product type and product features and the scale achieved (the number of users attained within the grant period). Details of the analysis and findings not reported here can be found in Means et al. (2014).

Results

Scaling

Table 1 shows the number of student users reported by each of the 28 projects with user data for the grant period. The range in scale attained by the NGLC products was dramatic: from a low of 181 students to a high of 130,000. The latter figure, which approximated the kind of dramatic growth investors in the ed tech sector like to see, was for Open Study, an online environment designed to promote student-to-student peer support for students taking the same or similar college courses. The number was an estimate because the only information the provider had was the number of unique user names and the percentage of users who were college students according to an online survey. This product's scaling strategy involved using social media to go directly to students; their colleges and instructors typically were not involved. This project was an example of the retail scaling strategy for a product that did not demand any fundamental changes in the way the course was taught.

The second highest scale achieved was for the University of Hawaii's STAR system, designed to assist students with academic planning and to generate automated messages to a student's advisor when the student's course selections threaten to take him or her off track for graduation. University of Hawaii's STAR project illustrates a very different pathway toward scale. An initiative of the University of Hawaii System, which includes all of the state's 2- and 4-year campuses, STAR was implemented top down by requiring all students to login to the system in order to register for courses. Again, the number of users was an estimate, in this case based on enrollment numbers for the university system. This project was an example of a top-down institutional implementation achieving scale within its own organization.

None of the other NGC projects broke the 10,000-user mark during the period of their grant, but a half dozen did attract 5,000 – 9,000 users. Qualitative analysis of project features associated with the scale achieved identified three features associated with rapid scaling:

- promise of cost savings,
- absence of a requirement for face-to-face instructor training, and
- minimal requirements for changing instructional processes or organizational structures.

When these relationships were tested quantitatively, estimated product implementation costs/cost reduction was the only feature with a statistically significant correlation ($p < .05$) with achieved scale.

Table 1. *Digital learning innovation scale and effectiveness, by project*

Project Institution	Users	Effect Size	Project Institution	Users	Effect Size
Abilene Christian, *MEIBL*	375	-.095	Marist College, *Open Academic Analytics Initiative*	2,288	+.119*
Arizona State U, *simSchool*	7,508	+.006	Missouri, NCAT gateway course redesign	8,522	-.008
Bryn Mawr, Blended STEM courses	729	+.053	OhioLINK, *Scaffold to the Stars*	1,482	+.084
Carnegie Learning, *Fluency Games*	421	-.006	*Open Study*	130,000	na
Community College of District of Columbia, *Learner Web*	297	na	Open University, *Bridge to Success*	4,185	+.268*
Central Florida University, *Blended Learning Toolkit*	5,799	na	SUNY, *Catch Up and Complete*	443	-.046
Central Piedmont, *Online Student Profile Learning System*	7,436	-.077***	U of California Office of the President, *Next Generation Online Instruction*	515	na
Cerritos College, *Kaleidoscope*	3,980	+.103*	U of Hawaii, *STAR*	23,509	na
Chattanooga, *Do the Math!*	9,465	+.326***	U of Massachusetts, *Wayang Outpost*	127	+.107
Community Colleges of Chicago, *Math on Demand*	648	+.320***	U of Michigan, *E^2Coach*	3,993	+.092**
CSU Northridge, *Hybrid Labs*	6,303	+.587***	U of Wisconsin, *Salon*	1,070	+.174
Eckerd College, *IPAL*	531	na	U of Wisconsin, *U-Pace*	181	+.621***
Indiana UPUI, *Cyber Peer-Led Team Learning*	252	-.383	Wake Forest, *BioBook*	504	-.112
Iowa CCOC, *eAnalytics*	1,344	-.080	WICHE, *NANSLO*	879	-.024

na = No effectiveness data available

Table 2 shows the mean and median scale achieved by each type of NGLC technology-based innovation. Because the Open Study and University of Hawaii are positive outliers, the medians are better indicators of the central tendency for these categories. The data suggest that the learning analytics/early alert projects—which all used an institutional scaling strategy—had the highest adoption rates within the grant term. The second highest adoption rate was found for the whole-course model projects, some of which used an institutional scaling strategy and some of which worked directly with individual faculty members.

We expected peer learning supports, tools for redesigning a course to use blended learning, and supplemental learning resources to attain larger scales than they did during the grant period because they do not require any surrender of faculty autonomy or require institutional change. With the very dramatic exception of Open Study, however, the particular technology-based innovations in these categories within the NGLC portfolio did not reach an impressive scale.

Table 2. *Digital learning innovation scale, by innovation type*

Innovation Type	**Number of Cases**	**Number of Students**			
		Mean	*Median*	*Minimum*	*Maximum*
Learning analytics / Early alerts	5	7,714	3,993	1,344	23,509
Peer learning supports	3	43,774	1,070	252	130,000
Supplemental resources	7	1,478	504	127	7,508
Supports for course redesign	7	3,067	1,482	443	8,522
Whole-course model	6	3,513	2,417	181	9,465
Overall	**28**	**7,957**	**1,207**	**127**	**130,000**

Impact on Student Outcomes

The effectiveness of the NGLC technology interventions was estimated by SRI using data provided by 22 of the 28 projects for both an intervention group and a comparison group (generally other sections of the same course during the same or prior academic terms). Student outcome measures varied: some projects used common learning assessments while others used course grades or successful course completion (i.e., receiving credit for the course).

The majority of NGLC technology interventions had no significant effect on student outcomes, indicated by an effect size that was not significantly different from zero. Eight interventions did have a significantly positive effect, and one had a significantly negative effect. A test of the relationship between intervention type and effectiveness revealed that, on average, interventions that involved whole-course redesign were more effective than the other types of technology-supported innovations. (See Figure 1.) In an instructor survey, these projects also were rated by a higher percentage of instructors as requiring a *major change* in their instruction (39% of respondents participating in whole-course innovations compared to 2% of those implementing supplemental resources, for example).

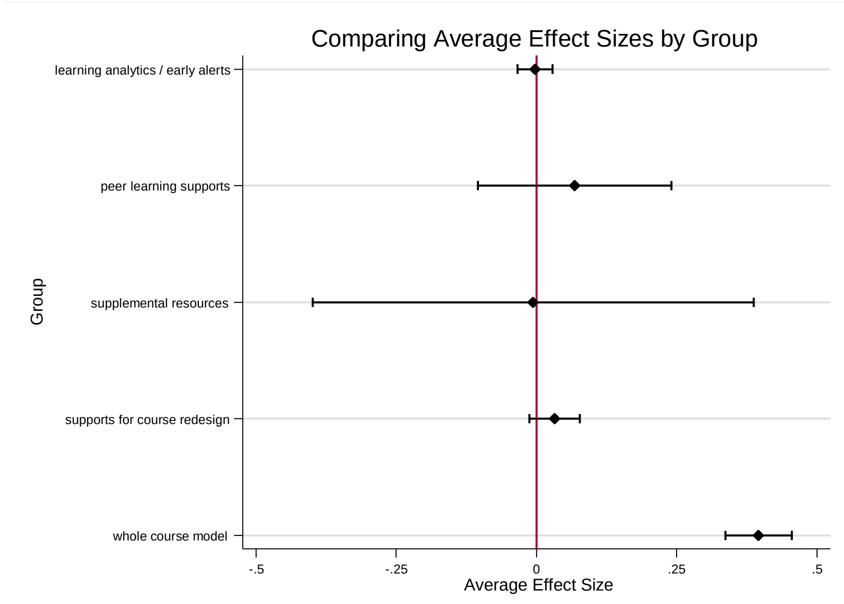

Figure 1. Digital learning innovation effectiveness, by innovation type

The innovation with the largest effect size for student outcomes was the U-Pace introductory psychology course developed at the University of Wisconsin-Milwaukee. This course involves online lessons and interspersed quizzes delivered within a self-paced mastery learning model that incorporates personalized feedback and motivational supports. This innovation's scaling strategy was to work with individual interested faculty members. In addition to the online learning modules, the project created templates for faculty feedback to students, a U-Pace instructional manual, a U-Pace online training module, and a learning management system (LMS) checklist for faculty who wanted to adopt the model. The project team also conducted a national training for prospective U-Pace instructors at a conference sponsored by the American Psychological Association.

Another innovation with reasonably strong evidence of impact was the Hybrid Labs model for the initial college mathematics course developed at California State University, Northridge. This course model calls for a combination of face-to-face lecture, facilitated group work in a supplemental course hour with a tutor or teaching assistant, and independent online homework with a mastery learning program for introductory mathematics combined with a common set of assessment and remediation activities. Because this model requires changes to course structure, assessment, and staffing practices, its adoption occurred at the departmental level. The project's scaling strategy involved working through supportive academic leaders at additional campuses in the California State University system.

The features associated with positive impacts in the NGLC data set included support for changes in pedagogy and the comprehensiveness of the intervention (e.g., inclusion of professional development and a full set of course curriculum resources). These features were quite different from those associated with widespread scaling described above.

Relationship between Effectiveness and Scale

The NGLC data provide some modest support for the hypothesis that effectiveness and scale are inversely related in that the average effect size for interventions scaling to fewer than five campuses was 0.14 compared with 0.06 for interventions scaled to more campuses during the grant term. But neither the correlation between number of students involved and effectiveness nor that between the number of instructors involved and effectiveness was significant.

To examine the relationship between effectiveness and scale in another way, we compared impact estimates for the institution where the innovation originated to those for other campuses to which the innovation was spread. Here we found a major difference: For the 12 effect estimates provided for a home campus, the average effect size was a statistically significant +.168 ($p < .05$) while the 29 estimates from expansion campuses averaged an effect size of -.004, a clear lack of impact at scale.

Discussion

The NGLC projects with evidence of positive impacts on student outcomes involved something more than supplemental resources for student or instructor use. These projects had a deeper vision for a transformed course experience and a theory of action that extended beyond the adoption of digital tools. Several, for example, called for conversion to a mastery learning system in which students' time on task would be allowed to vary as needed so that every student could attain the course learning objectives. Such changes require significant modifications of instructor practices and department policies, considerably raising the stakes for adoption. Such conditions raise challenges for scaling (cf. Blumenfeld et al., 2000) and make it difficult to obtain large numbers of users through a "retail" scaling strategy.

Prevailing wisdom is that because of the complexity and difficulty of getting buy-in and support from education institutions, technology-based learning innovations that can circumvent institutions are more likely to achieve broad scale. Certainly, this is the current thinking in the venture capital world, which tends to steer digital learning startups toward the consumer rather than the school market. The experience of the NGLC initiative suggests that this is an oversimplification.

Our expectation that the NGLC projects requiring less adjustment on the part of instructors and institutions would scale more rapidly was not confirmed by the data. Clearly other factors were at play, including the scaling strategy used by the grantee organization and the capacity of the project team to pursue innovation development and dissemination simultaneously (cf. Coburn, 2003). Reasonable levels of scale can be achieved by technology-supported interventions that require significant changes in teaching and learning but only if they have support from institutional leadership as well as other key actors in the system.

In hindsight, we should probably not have expected to find an inverse relationship between impact and scale for the set of NGLC projects. After all, achieving "Internet scale" for an educational innovation is—and is likely to remain—the exception rather than the rule. Even among non-education technology-based companies, only something like 1 in 10 achieves the dramatic growth rate that investors look for (Kutcher, Nottebohm, & Sprague, 2014). One could argue that Open Study was the only NGLC project achieving anything like Internet scale, and that project was not able to provide any data on the impact of its peer learning activities on students (and in fact, could not even identify which of its users were enrolled in college courses).

In considering the influences of technology on the scaling of education innovations, then, we need to remember that relatively few learning technologies achieve truly widespread use, and that those that do usually market themselves directly to users. In our experience, it is no small matter to move from broad use by individual learners to real assimilation within education systems. The publicity and satisfied users of such an innovation are likely to generate interest on the part of education institutions, but really incorporating the resource into core teaching and learning processes requires additional development of the innovation and organizational change on the part of education institutions.

Early efforts to bring the Khan Academy into schools are a prime example of this process. That work uncovered a number of basic incompatibilities between the product designed for individual learners working at home and the existing circumstances and practices in schools and classrooms. To meet schools' needs, Khan Academy had to develop large amounts of additional content aligned to grade-level mathematics standards and to add features providing teachers with the option of directing students to work on certain skills (Murphy et al., 2014). The schools themselves had to make accommodations such as providing additional computing devices so every student could have one during math instruction and supporting teacher learning of new classroom organization and student support routines compatible with individualized online learning. This bidirectional, interactive process of mutual accommodation in effect resulted in development of a new innovation. One indicator of the extent of the mutation of the innovation was that the Khan Academy videos that are such a prominent part of home use were rarely used by students in the 20 schools studied by Murphy et al.

All of this suggests that deep implementation and reliably positive impacts within an education system depend on multiple levels of institutional buy-in and support. The fact that an innovation has a strong technology component and can be accessed online may overshadow the time-intensive, nitty gritty process of cultivating that support, but it will not obviate its necessity.

References

Blumenfeld, P., Fishman, B., Krajcik, J. S., Marx, R. W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling-up technology-embedded project-based science in urban schools. *Educational Psychologist*, *35*(3), 149–164.

Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective. *Journal of the Learning Sciences, 21*(4), 487-521.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3-12.

Cohen, D., & Ball, D. (1999, June). *Instruction, Capacity, and Improvement.* CPRE Research Report Series RR-43. Consortium for Policy Research in Education: University of Pennsylvania.

Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 1-24.

Fishman, B., Marx, R., Blumenfeld, P., Krajcik, J. S., & Soloway, E. (2004). Creating a framework for research on systemic technology innovations. *The Journal of the Learning Sciences*, *13*(1), 43–76. doi:10.1207/s15327809jls1301_3

Fullan, M., & Donnelly, K. (2013). *Alive in the swamp: Assessing digital innovations in education.* Nesta Innovation Lab and the New Schools Venture Fund.

Jordan, K. (2014). Initial trends in enrollment and completion in massive open online courses. *International Review of Research in Open and Distributed Education, 15*(1). 133-159.

Kutcher, E., Nottebohm, O., & Sprague, K. (2014). Grow fast or die slow. McKinsey & Company. Available at http://www.mckinsey.com/industries/high-tech/our-insights/grow-fast-or-die-slow.

Means, B., Roschelle, R., Penuel, W., Sabelli, N., & Haertel, G. (2004). Technology's contribution to teaching and policy: Efficiency, standardization, or transformation? In R. E. Floden (Ed.), *Review of research in education* (Vol. 27). Washington, DC: American Educational Research Association.

Means, B., Shear, L., Zheng, Y., & Deutscher, R. (2013). *Next generation learning challenges: Evaluation of Wave I.* Menlo Park, CA: SRI International.

Münch, J., Fagerholm, F., Johnson, P., Pirttilahti, J., Torkkel, J., & Jäarvinen, J. (2013). Creating minimum viable products in industry-academia collaborations. In B. Fitzgerald, K. Conboy, K. Power, R. Valerdi, L. Morgan, & K.-J. Stol (Eds.), *Lean Enterprise Software and Systems* (pp. 137–151). Berlin, Heidelberg: Springer Berlin Heidelberg.

Murphy, R., Gallagher, L., Krumm, A., Mislevy, J., & Hafter, A. (2014). *Research on the Use of Khan Academy in Schools.* Menlo Park, CA: SRI Education.

Rogers, E. (1962). *Diffusion of innovations.* New York: Free Press.

Scott, A. (2012, April). The Dangers of the Minimally Viable Product. *Harvard Business Review*. Available from https://hbr.org/2012/04/the-dangers-of-teh-minimal-via

Vesselinov, R., & Grego, J. (2012). Duolingo effectiveness study: Final report. Available at https://static.duolingo.com/s3/DuolingoReport_Final.pdf