'Uberizing' Assessment: Why the Trust Economy Could Provide a Model for Schooling

By Jal Mehta on May 23, 2016 9:43 AM | 2 Comments

By Jennifer Poon, Innovation Lab Network Director at the Council of Chief State School Officers.

I've never used Airbnb because I don't trust strangers. Airbnb is the hot new service that allows me to rent someone else's home while they're gone. But how can I be sure that the sheets in this stranger's house are clean? That there are no bed bugs? That some creeper won't watch me sleep through the window at night?

No, I've always preferred large conglomerate hotel chains with their dependable, if not sterile, rooms; standard operating procedures; and star ratings provided by external reviewers.

And yet, just as my average nightly experience at the Hilton's of the world is on the decline (who left that soap in the shower and why does my room smell musty?), Airbnb has emerged as the **largest hotel chain in the world**, roughly 79 percent larger than the second place chain, Marriot International. And it has become the swear-by-it darling of my every friend who has tried it.

The success of Airbnb is representative of a larger shift in our society toward peer-to-peer sharing- one where anyone can rent a home, hitch a ride, run errands, or get recommendations from anyone else. And lately, this societal shift has me thinking a lot about the future of assessments, too. Like Uber and Airbnb, educational assessment is enveloped in a national debate about **authenticity, quality and trust**. And while it is wholly insufficient to compare rating a car service to the very complex and technical process of assessing student learning, I ask you to join me on a thought experiment reflective of this larger socio-anthropological shift in society.

At its core, "uberization" is about the redistribution of authority to first-hand users. It values authenticity. Take **Uber**, the car-sharing services that relies on users to rate their drivers on a one-to-five star scale. These ratings provide drivers with real-time feedback on their customer service while at the same time alerting passengers to the performance of individual drivers. The judgments are immediate, nuanced, and stem from users directly, making them more useful to drivers and passengers than a one-time taxi license provided by an agency that never once sat in the back seat of the driver's cab during rush hour.

Just as Uber and Airbnb place greater trust in authentic, user-driven judgments, there is **renewed debate** among education leaders about placing greater value on the expert judgement of teachers, especially when it comes to assessing student learning. Since No Child Left Behind, schools, districts and some teachers, have been held accountable for student performance on standardized assessments that are externally developed and delivered. While these tests can boast gold-star validity and reliability, many have begun to question their relevance - and with it, their value.

To improve the meaning and usefulness of assessments, some are beginning to call for greater emphasis on locally-developed assessments that are embedded in the curriculum and require students to demonstrate their learning through performance tasks. In such systems, students' performance tasks, projects, or exhibitions are rated by teachers (and sometimes by students themselves or other community experts) on a 4-point scale. More sophisticated than Uber's five stars, performance assessment scales leverage **rubrics** that clearly articulate expectations at each level of performance. Proponents reason that these assessments are more authentic because they better match student learning processes and real-world situations when compared to multiple-choice tests. Further, they provide real-time, contextual feedback that can be used to adjust instruction and supports to each student - something that the end-of-course externally-administered assessments are ill equipped to do.

Still others remain wary of placing too much trust in any one judgment by any one teacher. Focused on issues such as reliability and comparability - for example, what if one teacher plays favorites, or generally tends to "grade easier" than another? - critics of teacher judgment have good reason to desire greater standardization. Especially when teacher judgments are tied to high-stakes rewards or consequences, **some research** has shown that those judgments can be easily swayed.

Services that rely on consumer reviews have also had their share of corruption. Both **Amazon** and **Yelp** have had to mitigate against faux reviewers who are paid by companies or restaurants to give five-star reviews. Yet we haven't seen these companies retreat from placing user ratings at their core. Average user ratings still determine which properties float to the top of Airbnb lists; which hotels and experiences are ranked highest on TripAdvisor; and what feedback and professional support Uber and Lyft drivers will receive from corporate headquarters. In fact, Uber places its own "high stakes" on the average passenger rating for its drivers, **booting drivers** from the service if their scores dip too low.

Instead, we've seen these services add guardrails, such as averaging across a large number of ratings; coupling user reviews with expert-granted badges or credentials such as Airbnb's "Super Host" designation; and by showing direct evidence of quality through photo and video.

granted badges or credentials such as Airbnb's "Super Host" designation, and by showing direct evidence of quality through photo and video.

These guardrails are not unlike what some schools, districts, and states have established to ensure reliability and comparability in locally-developed or locally-scored assessments. In addition to taking measures to ensure the quality of the assessments themselves, states like **New Hampshire** and others are finding ways to ensure that samples of student work are scored by multiple teachers and experts; are coupled with externally-delivered standardized tests in benchmark years and subjects; and rely on evaluating real student work products against clearly-articulated criteria.

You may view this "uberizing" of student assessment as a welcome advancement or a mistake we will all come to regret. The difference may hinge on getting the guardrails right. But there's one fact that I can't ignore: Uber and Airbnb are working.

What's your take? Tweet to @JDPoon or #CCSSOILN to keep the conversation going.

Categories: Accountability   Assessment   School reform   State policy

Tags: airbnb   assessment   trust economy   uber