I a our

Toward Effective Courseware at Scale: Investig ating Automatically Gener ated Questions as Formative Practice

9MNX NX FS T[JW[NJ\ TK YMJ WJXJFWHHWNFQ[JXYNJQFYNLSJSHYMJYTZX
1JFWS G^)TNSL RTWJ FKKTWIFGQ^ NS HTZWXJ\FWJ G^ YMJ ;NYF
YJFR FX UWJXJSYJI FY YMJ 1%8HFQJ c ;NWYZF Q *[JSY KWTR /Z
JSYNWJ UFUJW MJWJ

If you follow trends in online education, you already know that online courseware can, when used well, provide great learning gains for students. Full courseware provides students with content to learn frBB 7 Cp 7We the q WJFXTSX <J \FSYJI YT PST\ NK FZYTRFYNH VZJXYNTS LJSJWFYNTS & practice opportunities as well as those written by content authors. If they do, effective courseware suddenly becomes much more broadly accessible to students who otherwise might have only a text-based resource.

&X &. IWN[JS VZJXYNTS LJSJWFYNTS NX F WJQFYN[JQ^ WJHJSY KNJQI lab setting with analys Ms done on a small number of questions. What we wanted to know was how well YMJ FZYTRFYNHFQQ^ LJSJWFYJI &, VZJXYNTSX UJWKTWR \NYM XYZIJ HTRUFWJI YT MZRFS FZYMTWJI -& VZJXYNTSX 4ZW WJXJFWHM VZJXY VZJXYNTSX JVZN[FQJSY YT -& VZJXYNTSX \NYM WJXUJHY YT JSLFLJRJ to answer that, we looked at data from 109 students who worked through courseware which contained both &, VZJXYNTSX FSI -& VZJXYNTSX 8YZIJSYX \JWJ STY F\FWJ FS^ TK YM simply learning material as they needed for their regular class.

KEY TERMINOLOGY

Automatic question generation (AQG): a method of creating formative practice questions at scale within courseware using artificial intelligence to minimize investments in human time and cost

Automatically generated (AG) questions: formative practice questions created by natural language processing and artificial intelligence using the course's textbook as source material

Human-authored (HA) questions: formative practice questions created manually by an individual and taken from the textbook's ancillary materials or written by subject matter experts

Recall question types: questions that require students to fill in a missing word rather than select from a fixed group of choices (in this study: AG or HA fill-in-the-blank questions)

Recognition question types: questions that require students to evaluate provided terms or concepts and select an answer (in this study: AG or HA matching; HA multiple choice, multiple choice multi-select, multiple choice grid, drag and drop, and pulldown)

DATA

WHAT DID WE FIND?

The data we have shows that

-JWJ NX FS J]HJWUY KWTR TZWJWJXMFWHJMKYMZSY XZRRFWN

1J[JQX TK JSLFLJRJSY INK*HZQY^ FSI UJWXNXYJSHJ \NYM &, VZJXY
courseware used by the same students were found to be largely equivalent.
While there were differences among results for individual question types, there was no evidence that
XYZIJSYX UWJKJWWJI -& T[JW &, VZJXYNTSX
9MJ KTWRFY TK F VZJXYNTS WJHTLSNYNTS [X WJHFQQ MFI YMJ LV
FSI YMFY INK*HZQY^ MFI FS NRUFHY TS UJWXNXYJSHJ F KYWNH \M
impact on learning.

Q: Will students engage with generated questions in the same way they do with authored questions?

A: Yes, they will! In no way did students seem to engage in the questions differently based on their source.

Q: Are generated questions too easy or too hard as compared to authored questions?

A: No! Within our ability to measure difficulty based on student data, you could not make any generali ations about either source of questions being too easy or hard for students.