Significant Determinants of Student Retention and Efficient Engagement Strategies in Online Second Language Learning Courses

Kolbrún Friðriksdóttir

Dissertation towards the degree of Doctor of Philosophy

2021

UNIVERSITY OF ICELAND
SCHOOL OF HUMANITIES

FACULTY OF LANGUAGES
AND CULTURES

Significant Determinants of Student Retention and Efficient Engagement Strategies in Online Second Language Learning Courses

Kolbrún Friðriksdóttir

Dissertation towards the degree of Doctor of Philosophy

University of Iceland
School of Humanities
Faculty of Languages and Cultures
December 2021

Mála- og menningardeild Háskóla Íslands
hefur metið ritgerð þessa hæfa til varnar
við doktorspróf í annarsmálsfræðum

Reykjavík, 6. september 2021

Geir Sigurðsson
deildarforseti

Faculty of Languages and Cultures
at the University of Iceland
has declared this dissertation eligible for a defence
leading to a Ph.D. degree in Second Language Studies

Faculty of Languages and Cultures

Doctoral Committee:
Birna Arnbjörnsdóttir, supervisor
Jozef Colpaert
Kristín Ingólfsdóttir

Significant Determinants of Student Retention and Efficient Engagement Strategies in Online Second Language Learning Courses

ISBN 978-9935-9640-0-7
ORCID: https://orcid.org/0000-0002-1032-4397

Ágrip

Doktorsverkefnið er á sviði annarsmálsfræða og hagnýtra málvísinda og beinist að áhrifaþáttum í námi í opnum netnámskeiðum, eða svokölluðum *massive open online courses* (MOOCs), en með þeim er veittur opinn aðgangur að menntun á ýmsum fræðasviðum í háskólum víða um heim. Nánar tiltekið beinist rannsóknin að námi og kennslu annars máls í netnámskeiðum, svonefndum *language MOOCs* (LMOOCs). Hún fellur jafnframt undir svið tölvustudds tungumálanáms og -kennslu, eða *computer assisted language learning* (CALL). Alþjóðlegar rannsóknir á notkun MOOC-námskeiða hafa sýnt fram á að lítill hluti nemenda lýkur jafnan námskeiðum að fullu. Þetta hefur vakið upp áleitnar spurningar um gæði slíkra námskeiða, kennslufræðina og námsumgjörðina sem nemendum er sköpuð, ekki síst með tilliti til málakennslu og málanáms. Á það hefur einnig verið bent í þessu sambandi að hópur nemenda í netnámskeiðum af þessum toga sé margbrotinn, með ólík markmið, bakgrunn og námsþarfir sem taka verði mið af í umræðum um námsframvindu og virkni nemenda.

Meginmarkmið rannsóknarinnar var að greina þætti sem gætu haft áhrif á framvindu (e. *retention*) í opnum netnámskeiðum og varpa ljósi á kennsluaðferðir og námsumgjörð sem gætu verið mikilvægur þáttur í að auka virkni og þátttöku nemenda í námskeiðum. Rannsóknin byggðist á gögnum frá nemendum á sjö netnámskeiðum Icelandic Online (IOL) í íslensku sem öðru eða erlendu máli en námskeiðin voru þróuð á vegum Háskóla Íslands. Þau eru sjálfstýrð og gagnvirk, ætluð fullorðnum og miðast þau við mismunandi færnistig í íslensku (A1–C1). Öll námskeiðin standa til boða í opinni námsumgjörð án endurgjalds og án stuðnings kennara en tvö þeirra eru jafnframt tiltæk gegn gjaldi og undir umsjón kennara, það er í blandaðri námsumgjörð og fjarnámsumgjörð.

Rannsóknin er þríþætt og grundvallaðist á blandaðri rannsóknaraðferð. Stuðst var við a) megindleg gögn úr gagnagrunni IOL frá rúmlega 43.000 nemendum á öllum námskeiðum IOL, b) megindleg gögn úr spurningakönnun með lokuðum spurningum sem lögð var fyrir 400 nemendur á einu námskeiði auk c) eigindlegra gagna frá 174 nemendum á einu námskeiði sem aflað var með opnum spurningum í spurningakönnun.

Í fyrsta hluta rannsóknarinnar, þar sem greind voru gögn úr gagnagrunni, var sjónum beint að því að kanna virkni, framvindu og námshegðun nemenda á námskeiðunum sjö og í mismunandi námsumgjörðum. Jafnframt var námsgreiningu (e. *learning analytics*) beitt til að varpa ítarlegu ljósi á brotthvarfsmynstur meðal þeirra sem luku ekki námskeiðum auk þess að veita innsýn í það hversu langt þeir fóru í námsefninu áður en þeir hættu.

Í öðrum hluta rannsóknarinnar, þar sem stuðst var við spurningakönnun, var annars vegar leitað eftir viðhorfum nemenda sjálfra til sex þátta sem lúta að efnisinnihaldi og kennslufræði sem beitt er á námskeiðinu og hins vegar til fjögurra þátta sem varða stuðning kennara í tveimur námsumgjörðum. Kannað var hvort nemendur teldu þessa þætti mikilvæga til að hvetja þá áfram eða ekki og einnig hvort þeir hefðu áhrif á námsframvindu samkvæmt mælingum vöktunarkerfis (e. *tracking system*) IOL. Þannig

voru mæld áhrif alls tíu námskeiðsþátta auk þáttar sem varðar upphaflegt markmið nemenda um þátttöku í námskeiðinu. Einnig voru áhrif lýðbreytna, aldurs og kyns, á framvindu skoðuð.

Í þriðja hluta rannsóknarinnar var kallað eftir eigindlegum textagögnum frá nemendum sem annars vegar höfðu lokið námskeiði og hins vegar frá þeim sem höfðu haft það að upphaflegu markmiði að ljúka námskeiði en gerðu það ekki þegar upp var staðið. Þannig voru þeir sem höfðu lokið námskeiði beðnir um að lýsa því hvað hefði umfram allt orðið til þess að hvetja þá áfram allt til enda námskeiðs. Að sama skapi voru þeir sem luku ekki námskeiði beðnir um að tilgreina ástæður þess að upphaflegu markmiði um að klára það hefði ekki verið náð.

Niðurstöður fyrsta hluta rannsóknarinnar, þar sem byggt var á gögnum úr gagnagrunni (n = 43.468), sýndu að hlutfall þeirra sem luku IOL-námskeiðunum alveg var lágt, það er 2,4 til 18,2% eftir námskeiðum og námsumgjörðum, sem samræmist niðurstöðum sambærilegra rannsókna á þátttöku í MOOC-námskeiðum. Mælikvarðinn sem var beitt í þessum hluta rannsóknarinnar miðaðist við það að nemendur hefðu farið yfir námsefnið allt til síðustu námsefnissíðu, en hvert IOL-námskeið felur í sér tugi efnissíðna og jafnvel mörg hundruð verkefni. Samkvæmt niðurstöðunum reyndust nemendur í blandaðri námsumgjörð líklegri til að ljúka námskeiði en nemendur í annars konar námsumgjörð. Greining á brotthvarfsmynstri í námskeiðunum, þar sem miðað var við hlutfall brotthvarfs á hverri efnissíðu í námskeiði, leiddi jafnframt í ljós að nemendur væru í sérstakri brotthvarfshættu á allra fyrstu stigum námskeiðs, sérstaklega í opinni sjálfstýrðri námsumgjörð. Einnig kom í ljós að margir nemendur höfðu klárað meginþorra námefnis, jafnvel 80 til 99% þess, þegar þeir hættu. Greining á heildarþátttöku nemenda í öllum námskeiðunum leiddi síðan á sambærilegan hátt í ljós að þeir sem höfðu greinst sem brotthvarfsnemendur í gagnagrunni IOL, samkvæmt fyrri skilgreiningu, höfðu í mörgum tilvikum lokið meirihluta námsefnis þegar þeir hættu.

Þessar niðurstöður greiningar á gögnum úr gagnagrunninum gáfu tilefni til að endurskoða skilgreininguna á því hverja beri að telja brotthvarfsnemendur í opnum netnámskeiðum af þessum toga. Í tveimur síðari hlutum rannsóknarinnar var því litið svo á að nemendur sem hefðu lokið að lágmarki 80% námsefnis teldust hafa lokið netnámskeiðinu en hinir sem fóru skemur voru skilgreindir sem brotthvarfsnemendur. Niðurstöður í fyrsta hluta rannsóknarinnar gáfu einnig tilefni til að rýna nánar í námshegðun nemenda í beinum tengslum við námsefnið sjálft, þá hvort tilteknir námsefnisþættir eða námsumgjörð gætu skýrt hátt hlutfall brotthvarfs á tilteknum efnissíðum námskeiðs og lítið eða ekkert brotthvarf á öðrum.

Niðurstöður úr öðrum hluta rannsóknarinnar, sem snúa að gögnum úr spurningakönnun (n = 400), sýndu í fyrsta lagi að 55–85% þátttakenda töldu alla kennslufræðiþættina sex eiga mikilvægan þátt í því að hvetja þá til þátttöku í námskeiðinu. Þær aðferðir sem beitt er í IOL að kynna ílag í smáum skrefum og bjóða upp á fjölþætt viðfangsefni og æfingar voru meðal þeirra þátta sem flestir töldu mikilvæga hvata. Þegar skoðað var á hinn bóginn hvort þættirnir sex hefðu áhrif á framvindu nemanna sýndi tenging gagna úr spurningakönnun við gögn úr gagnagrunni fram á jákvæða fylgni þriggja

þáttanna og námsframvindu en ekki hvað varðar hina þrjá þættina. Þátturinn sem snertir kynningu ílags í smáum skrefum reyndist hafa tölfræðilega marktæk áhrif á námsframvindu í rannsókninni.

Í öðru lagi kom í ljós að 50–90% þátttakenda í spurningakönnuninni (n = 64), sem höfðu verið í blönduðu námskeiði eða fjarnámskeiði, töldu alla fjóra þættina sem snerta aðstoð kennara eiga þátt í því að hvetja þá áfram. Þættir sem varða tímasetta áætlun kennara um yfirferð námsefnis í hverri viku og einstaklingsbundna aðstoð við nemendur voru þannig meðal þátta sem flestir töldu mikilvæga til að halda þeim við efnið. Við nánari skoðun á því hvort þessir fjórir þættir hefðu áhrif á framvindu nemanna í námskeiðinu sýndi tenging við gögn úr gagnagrunni fram á jákvætt samband allra þessara þátta og námsframvindu meðal nemenda í blandaðri námsumgjörð en ekki meðal þeirra sem voru í fjarnámsumgjörð. Í þriðja lagi sýndu niðurstöður fram á að íslenskunemar sækja námskeið í IOL með ólík markmið í huga. Um 57% þátttakenda (n = 226) höfðu haft í hyggju að taka fullt námskeið þegar þeir hófu nám en hinn hlutinn reyndist hafa áform um að fara eingöngu yfir hluta námsefnisins eða hafði óljós markmið. Þeir sem hófu nám með það í huga að ljúka námskeiðinu voru líklegri til að klára en hinir og reyndist þessi þáttur hafa tölfræðilega marktæk áhrif á námsframvindu í rannsókninni. Í fjórða lagi sýndu niðurstöður fram á áhrif aldurs á námsframvindu. Yngri hópar í rannsókninni voru líklegri til að ljúka námskeiði en þeir sem eldri voru. Þannig var sýnt fram á með línulegri aðhvarfsgreiningu að aldur hefði neikvætt forspárgildi með tilliti til framvindu. Kyn þátttakenda hafði hins vegar ekki áhrif á framvindu þátttakenda í rannsókninni.

Í síðasta hluta rannsóknarinnar, sem grundvallaðist á eigindlegum textagögnum (174 nemar), leiddi þemagreining í ljós margvíslegar ástæður þess að nemendur luku námskeiði eða hættu áður en því marki var náð. Meginþemun sem greind voru með tilliti til þeirra sem luku námskeiðinu benda til þess að innihaldsríkt efni og kennslufræði í IOL, viljinn til að ná góðum tökum á markmálinu og einlægur áhugi á landi og tungu eigi mikinn þátt í því að hvetja nemendur áfram allt til enda námskeiðs. Í ljós kom einnig að nemendur í blandaðri námsumgjörð, sem er eini hópurinn í rannsókninni sem var í einingabæru námi í IOL, töldu sókn eftir einingum eiga stærstan þátt í því að þeir luku námskeiði. Hvað varðar hinn hópinn sem hafði ætlað sér að ljúka námskeiði en hætti benda meginþemun til þess að skortur á tíma til að helga sig náminu hafi fyrst og fremst komið í veg fyrir að þeir luku námskeiði. Einnig komu fram vísbendingar um að sumir brotthvarfsnemenda gætu hafa verið á röngu stigi með tilliti til færni í málinu sem hafi orsakað brotthvarf. Að auki sýndu gögnin að margir þeirra sem talist höfðu til brotthvarfsnemenda í rannsókninni reyndust enn vera virkir nemendur í IOL en höfðu valið að fara í gegnum efnið eftir eigin hentugleika.

Þegar á heildina er litið sýna meginniðurstöðurnar fram á margvíslega þætti sem geta haft áhrif á virkni og framvindu nemenda í opnum netnámskeiðum og jafnframt kosti þess að beita blandaðri rannsóknaraðferð í rannsókn af þessu tagi. Með greiningu á gögnum úr gagnagrunni hefur í fyrsta lagi verið sýnt fram á gildi þess að rannsaka ítarlega námsferli nemenda í tengslum við námsefnið sem þeir nota, jafnt þeirra sem ljúka námskeiði og hinna sem gera það ekki. Auk þess hefur verið sýnt fram á kosti þess að

netnámskeið búi yfir innbyggðu vöktunarkerfi sem veiti möguleika á að afla og safna rannsóknargögnum í gagnagrunn.

Niðurstöðurnar benda í öðru lagi á mikilvægi þess að við þróun slíks tungumálanámsefnis á neti sé tekið tillit til kennslufræði annars máls, að það feli í sér innihaldsríkt og fjölbreytt námsefni sem miðist við þarfir fullorðinna og ólíkan námsstíl notenda, og að nýttir séu kostir margmiðlunar til að bjóða upp á gagnvirkt tölvustutt nám og kennslu. Í þriðja lagi benda niðurstöður til gagnsemi þess að nemendur fái notið stuðnings kennara og einstaklingsbundinna leiðbeininga í opnum netnámskeiðum. Með því að nýta spurningalista í rannsókninni til að kanna hug og reynslu notenda af kennsluefni og stuðningi kennara er jafnframt lögð áhersla á gildi þess að kalla eftir viðhorfum notenda sjálfra til námsefnis og námsumhverfis í heild. Ekki síst geta slík gögn ýtt undir endurbætur á námsefni og aukinn stuðning við nemendur.

Að síðustu er með greiningu eigindlegra gagna sýnt fram á með rannsókninni að ólíkur hvati geti legið að baki málanámi og virkni þátttakenda og einnig að einstaklingsbundnir og ytri þættir geti skýrt brotthvarf nemenda úr námskeiðum. Með því að laða fram einstaklingsbundna sýn nemenda í rannsókninni hefur þannig verið varpað víðara ljósi á þá fjölbreyttu þætti sem geta haft úrslitaáhrif á virkni og þátttöku nemenda. Niðurstöðurnar í heild geta einnig verið leiðbeinandi fyrir fræðimenn um þróun opinna tungumálanámskeiða á netinu og gefið hugmyndir að frekari rannsóknum á þessu sviði.

Lykilorð: annarsmálsnám; opin tungumálanámskeið á neti; tölvustudd málakennsla og málanám; kennslufræðilegir þættir námsefnis; stuðningur kennara; blönduð rannsóknaraðferð; Icelandic Online

Abstract

This doctoral thesis addresses the issue of commonly low retention rates in massive open online courses (MOOCs), and attempts to identify crucial factors that affect engagement and retention in language massive open online courses (LMOOCs). The study relied on data from learners in the open and free online program Icelandic Online (IOL), which provides self-guided online courses for second language learners, of which some are presented in different modes of delivery.

This is a three-tiered, mixed-method study that relied on tracked retention data, survey data in correlation with tracking data, and qualitative data elicited through a survey. First, the tracked retention data on learners' online behavior and progression throughout the courses came from approximately 43,000 learners in all IOL courses and delivery modes. Second, the survey study was based on data from 400 learners on their experiences of the course content as well as tutor support in one course, and the influence of these factors on student engagement and retention. Other motivational factors were also addressed in the survey, namely learners' initial intent of course engagement and the influence of this factor on retention. The impact of participant demographics on student retention was also investigated. Finally, the qualitative data were elicited from 174 informants in one course through open questions included on the survey, to reveal learners' own views regarding the factors that either drove them to complete the course or prevented them from completing it.

Firstly, the findings of the tracking data analysis showed relatively low completion rates across all courses and modes of delivery, ranging from 2.4% to 18.2%, and that the blended learning modes were more effective in retaining learners as compared to other delivery modes studied. Furthermore, through the mining of these data and the use of learning analytics, the study identified a pattern of attrition among learners who did not complete courses to the very-end, as well as a pattern of user engagement across all courses and modes. The analysis therefore provides detailed information on the timing of student attrition as well as the extent to which non-completers engaged with the course material. While the findings showed that students commonly drop out early on in these courses, they also revealed that learners may disengage towards the end of a course. These findings called for reevaluation of the previous frameworks that measure students' attendance in MOOCs: Instead of defining course completion as 100% coverage of a course's content in the follow-up survey studies, it was redefined as completion of 80% to 100% of a course's content. The findings from the tracking data highlight the value of exploring learners' tracked progress and behavior in detail within the context of their learning materials in order to gain further understanding of student retention in MOOCs, with consideration to those who covered course content to the end, as well as those who did not.

Secondly, the survey study identified six content-related factors that most participants considered important for their engagement with the course, as well as four tutor-specific factors that apply to the blended and distance modes. Among these factors are gradual and scaffolded presentation of input and private interaction with a tutor. When

the survey data were measured against the tracking data, three of the content-specific factors were found to have a positive impact on student retention while the other three did not. All of the tutor-specific factors seemed to have a positive impact on retention in the blended mode, but none of them did in the distance mode. The study thus found that the instructional methodology and engagement strategies applied in IOL may benefit the language learner in terms of engaging him or her with the learning material, and therefore underlines the value of using CALL design within the LMOOC learning environment. The results also underscore the potential benefit of the presence and guidance of a tutor in the LMOOC learning environment on learner engagement with the material. The findings from the survey data that were also connected with the tracking data underline the value of such self-reported data in informing research about users' opinions and experiences of a learning material and its impact on student engagement and retention.

Thirdly, the results showed that learners join MOOCs with various goals in mind in terms of participation. While over half of the learners in the study entered the program with the intent to complete the full course, many did not. The study also found that the goal of completing a course had a significant impact on course completion. These results stress the importance of considering learners' initial intended participation as relates to MOOC retention. Fourthly, the factor of age was found to have a negative predictive value in the study, while gender was not found to impact retention.

Finally, the analysis of the qualitative data from learners who had completed a course revealed various motivators for continuing with the course, such as interesting course material or an interest in the language or culture. Furthermore, statements from learners who had had the initial goal of taking a full course but disengaged before completing show that factors unrelated to the course, such as time constraints, affected retention. By capturing learners' own thoughts on the reasons why they completed or did not complete a course, the study therefore provides broad individual perspectives on critical factors of LMOOC retention.

Overall, the study has identified multi-ranged determinants of student retention. The thesis provides a new framework for promoting student retention in LMOOCs, including engaging instructional strategies and the supervision of a tutor, which may provide a useful guide for educators and developers of LMOOC courses, and suggests avenues for future research.

Keywords: L2 online learning; LMOOC retention; CALL; content factors; tutored factors; mixed-methods research; Icelandic Online

Acknowledgements

I would like to thank my supervisor, Birna Arnbjörnsdóttir, who not only guided me through this process academically but also provided me with great personal support and encouragement. I also want to express special thanks to other members of my doctoral committee, Jozef Colpaert and Kristín Ingólfsdóttir, for their enthusiasm and constructive criticism. Furthermore, I wish to acknowledge the students in the course Icelandic Online 2 who participated in the survey study and my MA students in the course Computer Assisted Language Learning and Teaching at the University of Iceland for their input. Finally, I thank the following assistants: Daníel Páll Jóhannsson for mining tracking data from Icelandic Online's database, the Social Science Research Institute at the University of Iceland for their collaboration in carrying out the questionnaire for the study and data analysis, as well as Alexander Elliott, Katrín Axelsdóttir, Kelsey Paige Hopkins, Marcia Allison, and Randi Whitney Stebbins for proofreading the articles and the dissertation.

Furthermore, I am very thankful to my family, friends, and colleagues who have encouraged me in a number of ways. Most of all I would like to thank my late mother, Bergljót Ingólfsdóttir, and late father, Friðrik Kristjánsson, for their constant encouragement and support throughout the years.

In particular I would like to thank my children to whom I dedicate this work for their sincere enthusiasm and for supporting me in such a loving way through this process.

This work was supported by Nordkurs.

To Bergljót Vala and Kolbeinn

Table of Contents

List of Abbreviations and Use of Concepts

CALL:	Computer-assisted language learning
CEFR:	Common European Framework of Reference[i]
IOL; IOL 1–5:	Icelandic Online; Icelandic Online 1; Icelandic Online 2; Icelandic Online 3; Icelandic Online 4; Icelandic Online 5
LA:	Learning analytics
LMOOCs:	Language massive open online courses
L2:	Second language
MOOCs:	Massive open online courses
SLA:	Second language acquisition
UI:	University of Iceland

The discussion is based on the following key terms:

Attrition patterns refer to the analysis of the tracking data on the overall online behaviors of students in IOL 1 and 2 who disengaged before completing the course, displaying the rate of attrition on every content page in the courses. *Attrition rate* is calculated as the number of students who disengage on each content page divided by total number of students who began the course.

Content coverage refers to learners' progress as they access content pages or cover the content of a course. This study measures student retention and engagement in IOL, and not actual performance.

Content page(s) refers to the organization of content in the IOL courses and not to an actual page online. The courses are varied in length and contain 29 to 139 content pages, each with 3 to 10 learning objects. The first content page in each course has the number 111.

Content-specific factors represent course elements that concern structure and organization of IOL 2, as well as design and pedagogical principles, that is:

a) *Clear and salient learning objectives* regarding grammar, vocabulary and language usage, which are presented initially in each section.

b) *Continuing storylines* that serve as an engaging context for language input and practice across a section.

c) *Curated and sequenced course structure* where the material is segmented into five sections each with three lessons (Part I, Part II, Part III), and guides the learners to follow a sequenced learning path.

d) *Form-focused and scaffolded presentation of grammar* that draws learners' attention to the target grammar in a meaningful context with different colors, depending on the input, including scaffolded grammar introduction.

e) *Gradual and scaffolded presentation of input*, whether it is grammar or vocabulary, is put forward gradually in three steps in each section (Part I, Part II, Part III), and involves the use of scaffolding of information throughout a section.

f) *Variety in types of learning objects* that present broad range of activities and interchanging characteristics of learning objects throughout the course.

Course completion is defined differently depending on the article. In Article I it refers to students who cover 100% of a course's content, that is, up to and including the last content page. In Articles II and III the term refers to students who cover 80% to 100% of the course's content in IOL 2 (complete content pages # 51 to 63).

Curated course refers to structured and organized course content in IOL 2 with theme-based and scaffolded learning input adapted to the learners' levels and the learning objectives in focus within and across sections.

Engagement patterns refer to the analysis of tracking data on the overall engagement of the students in IOL who did not complete the courses, based on different parameters for coverage of course content (less than 50%; 50% to 74%; 75% to 89%; and 90% to 99%), displaying the extent to which they engaged with the course material.

Focus on form applies to instructional strategies that IOL employs in order to focus learners' attention on targeted grammatical forms in context.

Initial goal concerns students' stated intentions of course engagement when they enrolled, whether they had the goal of completing the course, to work on part of it, or had no clear goal.

Mode(s) of delivery refers to the different modes through which IOL 1 and 2 are delivered:

a) *open self-directed mode*, which offers a free non-credential course without tutor intervention, b) *blended learning mode*, which offers a credential course located on campus under a supervision of a tutor, or c) *distance learning mode*, where a non-credential diploma course is delivered at a distance and includes support of a tutor.

Reason(s) for completing refers to students who completed IOL 2 and their reported views on why they completed the course.

Reason(s) for non-completing refers to students who had the stated intention of completing IOL 2 and their reported views on why they left the course without completing.

Scaffolding applies to instructional scaffolding of content and involves carefully organized, step-by-step presentation of new materials and practice in order to provide learners with gradual support in their learning process.

Student retention refers to the persistence of students enrolled in IOL, of whom some complete the course while others disengage before reaching the end.

Tutor-specific factors concern features that are provided by the tutor in the blended and distance modes of IOL 2, namely:

a) *A detailed introduction* to the organization of the course, its objectives, and technical functions.

b) *A set syllabus* where the tutor has pre-organized the course content into timetabled, manageable sections involving a recommended amount of material to cover each week in order to complete the course.

c) *Private interaction with the tutor* that may occur face-to-face or via email in the blended course and via email or through the course editor in the distance course.

d) *Tutor support overall* during the course that includes any kind of assistance with the learner concerning course content or technical issues.

List of Figures

List of Tables

List of Original Papers

This thesis is based on the following original publications, which are referred to in the text by their Roman numerals (I–III):

I. Friðriksdóttir, K. (2018). The impact of different modalities on student retention and overall engagement patterns in open online courses. *Computer Assisted Language Learning*, *31*(1–2), 53–71. https://doi.org/10.1080/09588221.2017.1381129

II. Friðriksdóttir, K. (2019). The effect of tutor-specific and other motivational factors on student retention on Icelandic Online. *Computer Assisted Language Learning*, 1–22. Advance online publication. https://doi.org/10.1080/09588221.2019.1633357

III. Friðriksdóttir, K. (2021). The effect of content-related and external factors on student retention in LMOOCs. *ReCALL*, *33*(2), 128–142. https://doi.org/10.1017/S0958344021000069

1. INTRODUCTION

The rise of Massive Open Online Courses (MOOCs), which are among the most recent e-learning and distance education initiatives to attain popularity among universities across the world, has generated the interest of students, educators, and researchers internationally (Aldowah, Al-Samarraie, Alzahrani, & Alalwan, 2020; Wiebe, Thompson, & Behrend, 2015). MOOCs are aimed at large-scale participation, offering free and open courses on topics from a wide range of academic disciplines to a diverse audience (Bárcena & Martín-Monje, 2014; Ingolfsdottir, 2014). Only a few MOOCs are specifically dedicated to languages, known as Language MOOCs (LMOOCs), many of which are still in the early stages of their development as platforms tailored to the task of language learning and teaching (Colpaert, 2014; Sokolik, 2014). Due to the rapid expansion and popularity of MOOCs, the issue of student retention is of special interest in the field (Chen et al., 2020; Jordan, 2015).

This cumulative thesis seeks to advance knowledge and provide new evidence of critical determinants of student retention in LMOOCs, with the main aim of identifying efficient engagement strategies in such learning environments. The research is rooted in the field of Second Language Acquisition (SLA), in particular the fields of Computer-Assisted Language Learning (CALL) and Applied Linguistics, drawing on sociocultural theories of language development. In the study, critical factors of retention in LMOOCs were considered in the context of an existing CALL program, Icelandic Online (IOL), developed at the University of Iceland (UI). The program comprises comprehensive and organized language learning materials, and was specifically designed for online teaching and learning of Icelandic as a second or foreign language. IOL provides seven open and free consecutive courses, all of which are self-guided and interactive. Two of these courses are also offered in tutorial modes as blended and distance courses, where the learners receive guidance from a tutor. The IOL courseware includes a built-in tracking system that collects and stores user data in a database. Among the main theoretical criteria that guided the development of IOL's pedagogy are curated and structured course content, form-focused grammar instruction, and scaffolding of information. This research study employs mixed methods, and aims to explore the value of these specific courseware components, as well as other influencing factors, in engaging language learners with the course; the purpose is not to evaluate the design itself of the IOL materials through theory-based perspectives. The study is based on tracking data gathered through IOL's tracking system as well as survey data, and focuses on learner engagement and critical determinants of retention in the context of learning material from a wide array of perspectives. This

includes factors related to course content and instructional methodology, mode-specific features pertaining to tutor involvement, and other motivational and non-course related aspects.

As the candidate considers language learners' engagement with learning materials to be an important prerequisite for learning, this study was predominantly motivated by a desire to uncover efficient engagement strategies and significant determinants of student retention in LMOOCs and online second language courses in general. The study was, moreover, interested in providing data on users and usage of the IOL program, while simultaneously shedding light on student online behavior and progress across open online language cources. IOL has attracted thousands of learners around the world, and was regarded as an ideal source of evidence for addressing the research questions in the study. At the time of this research, IOL's tracking system had recorded user data over eight years. This offered the candidate a unique opportunity to analyze substantial information about the learners and their usage of the curriculum. This thesis describes the first study ever conducted on IOL's database. The candidate has been involved in IOL from an early stage of the development and throughout its design process, and is currently a project manager on IOL. She is a linguist in Icelandic and an experienced on-campus teacher in Icelandic as a Second Language and in Second Language Studies at UI, as well as tutor in IOL's blended and distance courses, which may entail individual contact between the tutor and the learners.

The introduction of the dissertation is divided into four main chapters and is structured as follows: This chapter presents the research topic and Chapters 1.1 and 1.2 provide the theoretical background of the project as well as the issues addressed in the study. Chapter 1.3 describes the IOL program, its background, and the courses under investigation. Furthermore, that chapter outlines the courses' content and instructional methodology as well as the different modes of delivery, including the content-specific and tutor-specific factors addressed in the study. IOL's tracking system is also described in Chapter 1.3. Chapter 1.4 introduces the study's main objectives and the research questions that are addressed. Chapter 2 then outlines the research design and the research methods used in each part of the study. Chapter 3 presents and discusses the results of the study. Finally, Chapter 4 includes a summary of the key findings and discusses the main implications of the study, and considers the limitations as well as potential directions for future research. In the following section the study is placed within the context of relevant literature.

1.1 Context of the Study

While MOOCs provide new paths in teaching and learning in higher education by making courses available to anyone, anywhere, with enrollment by the thousands, this mode of delivering education also poses new challenges. Issues center around low completion rates in MOOCs, which typically range from two to eleven percent (Chen et al., 2020; Gaebel, 2014; Ingolfsdottir, 2014; Jordan, 2014, 2015; Koller, Ng, & Chen, 2013; Reich, 2014). These issues have called into question the quality of learning materials, instruction, and the methodological strategies for transmission (Castrillo, 2014; Colpaert, 2014; de Freitas, Morgan, & Gibson, 2015; El Said, 2017; Sokolik, 2014). More evidence is needed to further the research on the complex interrelationships between learners and the learning materials they interact with, and the various kinds of determinants that influence student retention in such learning environments. The literature on learner behaviors in MOOCs provides some insight into patterns of retention which have led to a questioning of the most appropriate measures of retention (Hone & El Said, 2016), where it has been argued that a broad range of learners supposedly come to these courses with various motives, and may not necessarily aim for completing a program (Bárcena & Martín-Monje, 2014; de Barba, Kennedy, & Ainley, 2016; Henderikx, Kreijns, & Kalz, 2017; Hew, 2016; Ingolfsdottir, 2014; Koller et al., 2013; Reich, 2014). Non-completers in MOOCs have in that regard drawn attention in the literature (Frydenberg, 2007; Koller et al., 2013; Reich, 2014), where researchers argue that the timing of attrition (Chen et al., 2020; Greene, Oswald, & Pomerantz, 2015; Jordan, 2015; Reich, 2014) needs to be addressed in conjunction with the course content and its users, as well as students' initial intent in terms of course engagement when they sign up (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014). Definitions of course completion vary in the MOOC literature (Henderikx et al., 2017; Jordan, 2015; Perna et al., 2014), and more evidence has been called for on user progression beyond the first and last milestones in such a course (Greene et al., 2015; Perna et al., 2014). There is, furthermore, a lack of studies based on tracking data (Chun, 2013; Fischer, 2007; Greene et al., 2015) that aim to reveal actual retention and engagement with course content. Similarly, there is a lack of large-scale, long-term empirical studies in CALL that also involve follow-up studies. Most of the existing studies are limited to a few subjects and a short time period (Gillespie, 2020). Researchers have underlined the value of using learning analytics to investigate learner-produced data as MOOC learners engage with a course as a source of insight into the effectiveness of course

materials (Godwin-Jones, 2017; Long & Siemens, 2011; Martín-Monje, Castrillo, & Mañana-Rodríguez, 2018; Thomas & Gelan, 2018). It has thus been highlighted that MOOC retention must be evaluated in light of such internal factors as learning material, instructional design and pedagogical practices, and student support, as well as in view of individual factors that may impact student engagement (Colpaert, 2014; Henderikx, Kreijns, & Kalz, 2018; Hew, 2016; Hone & El Said, 2016; Kim et al., 2017; Koller et al., 2013; Shapiro et al., 2017; Sokolik, 2014). Different modes of delivery for learning and teaching have also been a central point in the discussion of what kind of delivery mode may be most advantageous to engaging and retaining learners (Bettinger, Fox, Loeb, & Taylor, 2017; Garrison & Kanuka, 2004; Harker & Koutsantoni, 2005). However, some of these studies raise concerns since they commonly compare student experiences and retention in courses that are delivered in non-equivalent modes (Patterson & McFadden, 2009), or even in the same delivery platform but containing diverse online programs (Levy, 2007). The impact of modes of delivery on student retention has thus been insufficiently researched to date. The body of knowledge about the factors specifically assigned to the mode of delivery that may influence student retention also requires more data.

MOOC course design and content have been identified as critical factors for retention (El Said, 2017; Hone & El Said, 2016), and diverse course elements and strategies have been proposed as significant engagement features. In the field of LMOOCs, concerns have arisen about whether the language learners are provided with engaging forms of design strategies and pedagogy in order to facilitate the process of language learning, and to improve engagement in such learning environments (Castrillo, 2014; Colpaert, 2014; Sokolik, 2014; Teixeira & Mota, 2014). The benefit of using computer technology for language learning, or CALL, has been highlighted in this context, which has evolved from being "affordance-driven toward more pedagogy-based approaches" (Colpaert, 2010:259) to focus more on the learner and student-centered pedagogy. Greater knowledge is needed on the effect and value of the instructional resources in CALL and SLA pedagogical approaches for the language learner (Chun, 2012, 2016; Colpaert, 2010, 2018; Godwin-Jones, 2017). The challenge of engaging autonomous MOOC learners has, furthermore, called for research on the value and impact of instructor presence on student engagement and retention. Existing literature suggests that tutor presence and tutor support may be crucial elements that facilitate the engagement of MOOC learners and encourage

the learning process (Hew, 2016; Hew & Cheung, 2014; Hone & El Said, 2016; Ross, Sinclair, Knox, Bayne, & Macleod, 2014). There is, however, a lack of research on specific course design properties in MOOCs and how they affect learner engagement and outcome. In particular, more knowledge is needed on LMOOC learners' experiences in such environments, and how content-specific factors and instructional pedagogy may affect the language learner's engagement and retention (Castrillo, 2014; Colpaert, 2014; Martín-Monje et al., 2018). Similarly, more empirical evidence is needed on the effect of tutor support and guidance on learner engagement and retention in LMOOCs (Bárcena & Martín-Monje, 2014; El Said, 2017; Hew, 2016).

While motivating content or tutor support may be important for retention, individual factors also play a role. Students attend MOOCs with multiple motives and goals that also need to be considered in the discussion on retention (de Barba et al., 2016; Kizilcec & Schneider, 2015; Littlejohn, Hood, Milligan, & Mustain, 2016; Sokolik, 2014). Furthermore, greater understanding is needed as to why MOOC learners who intend to complete a course disengage before completion (El Said, 2017; Reich, 2014). This is the scope of the study.

Perna and colleagues (2014) have called for more evidence on the pattern of user progression through MOOCs, from the time they register or begin a course until they leave, in order to shed light on predictors of student persistence and completion. To do so, they explored particular milestones in MOOC courses, for example whether participants had accessed any, some, or all of a course's lectures or attempted any, some or all of a course's quizzes, and how these milestones predicted student participation and course completion. While their study provides insight into how learners move from a single milestone to another, more detailed information is needed about how learners engage with learning material. Research on student retention and low completion rates in MOOCs has brought attention to attrition patterns in such courses and to the timing of student drop-outs (Chen et al., 2020; Greene et al., 2015; Jordan, 2014, 2015; Koller et al., 2013; Reich, 2014). Previous studies on attrition behaviors in MOOCs have identified considerable variation among drop-outs across a courses early, middle, and late periods, while the highest rates of attrition have been found at the beginning of such courses (Frydenberg, 2007; Ihantola, Fronza, Mikkonen, Noponen, & Hellas, 2020; Jordan, 2014, 2015; Perna et al., 2014; Reich, 2014). Others (de Freitas et al., 2015; Greene et al., 2015) have reported in a similar way relatively rapid attrition at the start of MOOCs, while drop-outs became

increasingly unlikely as participants neared the end of the course. Scholars have questioned whether completion rate alone is a valid metric of student engagement and retention in MOOCs, arguing that while completion rates offer a convenient metric for comparison across wide array of MOOCs, they fail to capture the diversity of the goals and engagement patterns that students may have in these courses (Hew, 2016; Koller et al., 2013; Reich, 2014). Greene and colleagues (2015) highlight that research on retention in MOOCs should consider precisely when over the course of a program learners drop out, in order to predict retention and achievement. They point out that the common way of classifying participants groups those who drop out in the first week of a course together with those who drop out in the last week, that is, both groups are considered non-completers. This classification leads to a loss of information about drop-out timing in MOOCs. Sokolik (2014) furthermore underlines that applying the traditional metrics of higher education to MOOCs is entirely misleading. Considering that participation in a MOOC is voluntary and completion of and engagement in a course are not enforced, she points out that there is reason to question the concept of 'dropping out' in these sorts of courses. Other issues may also affect student retention, such as certain life events, shifting personal priorities, or trying other courses, and many enrollees may enrol without ever intending to complete the course (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014; Shapiro et al., 2017). Along these lines, researchers argue that retention in MOOCs is commonly evaluated without accounting for student goals, emphasizing that retention in MOOCs must be considered with respect to learners' intended engagement in the course (Koller et al., 2013; Reich, 2014). Evidence shows that many learners join a MOOC with the initial intention to complete it, and that those who set this goal are more likely to complete the course than those who have no such goal. On the other hand, findings also show that many learners who approach a course with the intention to complete it fail to do so (Reich, 2014). Previous research has identified various factors that possibly explain why these learners disengage from MOOCs. These factors relate to issues regarding course design, low course interactivity, lack of support, or to outside variables such as time constraints or lack of motivation (de Freitas et al., 2015; El Said, 2017; Gimeno, 2020; Henderikx et al., 2018; Shapiro et al., 2017). However, greater understanding is needed to explain what factors hinder such learners from completing a MOOC (El Said, 2017; Reich, 2014). Non-completers in MOOCs may thus provide valuable information regarding the factors that prevent them from completing a course (El Said, 2017; Henderikx et al., 2018; Reich, 2014).

With the recent interest in 'big data', defined as large sets of structured data intended for data analysis (Godwin-Jones, 2017), and analytics generated by MOOCs, new possibilities have emerged in the field of learning analytics (LA), which is generally understood as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (Long & Siemens, 2011:34). Long and Siemens (2011) and Godwin-Jones (2017) thus point out the importance of big data in providing valuable teaching and learning insights, including the potential of learning histories and personal profiles for tailoring the delivery of learning materials. Furthermore, while analysis of such data potentially provides researchers with a rich source of data that can better explain learner behavior, it also affords the opportunity to discover valuable patterns in the data and to visualize and analyze learners' online interactions in the learning context (Ebben & Murphy, 2014; Gelan et al., 2018; Godwin-Jones, 2017; Hone & El Said, 2016; Martín-Monje et al., 2018; Thomas & Gelan, 2018). Fischer (2007) has called for more objective data on the instructional value of CALL programs and how learners use specific components and features of a courseware. He emphasizes the need for computer-based tracking data in order to shed light on the actual usage and the value of such programs. As Fischer has argued, the comparison of students' actual versus self-reported use of a software leads to an overreliance on self-reported data. While tracking technologies can provide useful information about both second language acquisition and pedagogical design, as well as about the usage of a program, a remarkably large proportion of CALL studies do not report on tracking data (Chun, 2013). In the same vein, Gillespie (2020) points out that there is a lack of evidence based on large-scale, long-term empirical CALL studies involving a large group of students. In his review of recent empirical research in CALL, he found that most studies involve a small number of subjects, cover a short period of time, and are rarely followed up (Gillespie, 2020). The automatic collection of tracking data and analysis of entire student cohorts, which involves data produced by learners themselves in their learning context, has the potential to afford valuable insight into the users and usage of a learning environment, as well as the impact of the learning material itself.

The emerging field of LMOOCs, which have been described as "dedicated Web-based online courses for second languages with unrestricted access and potentially unlimited participation" (Bárcena & Martín-Monje, 2014:1), has raised concerns about

whether such learning environments provide the autonomous language learner with engaging forms of design strategies and pedagogy in order to encourage the language learning process (Castrillo, 2014; Colpaert, 2014; Sokolik, 2014). Colpaert (2014) argues that the design of LMOOCs and MOOCs in general mainly reflects the design of the tool applied, "assuming some globally applicable learning and teaching model" (p. 167), and criticizes the lack of adaptation to the relevant subject matter and a targeted user group or subgroups with different needs. These issues might explain the high attrition rates in (L)MOOCs. Along the same lines, Castrillo (2014) and Sokolik (2014) emphasize that LMOOCs require a platform that is particularly aimed at the complex undertaking of teaching and learning a language. Castrillo (2014) thus highlights the importance of creating a structure that intertwines technology and pedagogy, takes different learning styles into consideration, and is based on the Common European Framework of Reference (CEFR) (Council of Europe, 2018). The value of CALL for the language learner has been pointed out in that regard, as has the potential of effectively interweaving technology and pedagogy (Chun, 2012, 2016; Colpaert, 2014). Colpaert (2014) emphasizes that LMOOC courses should represent good design in a methodological and justifiable way that involves features that depend on local and specified contexts, and suggests that such courses take into consideration relevant findings in CALL. CALL has gradually embraced full integration of technology into second language teaching, learning, and research (Chun, 2016; Garrett, 2009). As Garrett (2009) highlights, ""CALL" is not shorthand for "the use of technology" but designates a dynamic complex in which technology, theory, and pedagogy are inseparably interwoven" (p. 719–720). Schulze and Sholz (2016) underline that behind learner-computer interactions (LCI) in CALL is a complex, adaptive system that must be considered when making design decisions and conducting research in CALL. Recently, increased focus has been put on *affordances* in CALL, which describes the potential opportunities that the learning environment offers the user. Blin (2016:57) has defined affordances in CALL as "a unique combination of technological, social, educational, and linguistic affordances". The main challenge for CALL design may, however, be to ensure that the affordances embedded in a system promote the emergence, perception, and realization of linguistic affordances. CALL researchers have highlighted the need for more evidence on how technologically enhanced devices and activities may support the language learning process, and call for more knowledge based on research on CALL's theories, methods, and models (Chun, 2012, 2016; Colpaert, 2018).

Factors associated with MOOC course design have been identified as crucial factors in engagement and retention (Adomopolous, 2013; El Said, 2017; Hone & El Said, 2016; Salmon, Pechenkina, Chase, & Ross, 2017), and diverse course elements and strategies have been proposed as significant engagement features. Researchers (Garrett, 1991; Hubbard, 2013; Ross et al., 2014) have underscored the need for content curation for autonomous learners in open online environments where course materials are selected and organized according to topic and the appropriate language level. Other factors that have been highlighted as important to the promotion of learner engagement are well-ordered and structured course content with organized objectives, and explicit explanations of what is expected of the learners (Dörnyei, Muir, & Ibrahim, 2014; El Said, 2017; Kizilcec & Schneider, 2015). Others have pointed out the need for a more scaffolded approach and more structured support to increase retention in MOOCs (Gimeno, 2020; Rosenshine & Meister, 1992; Salmon et al., 2017; Teixeira & Mota, 2014). Furthermore, evidence suggests that appealing learning materials serve as important stimuli for learners (El Said, 2017; Gimeno, 2020; Hew, 2016). Situational interest in a topic has thus been suggested as a strong motivational factor in MOOCs, and course activities and content may be key elements in activating and maintaining students' attention (de Barba et al., 2016; Hone & El Said, 2016; Höfler, Zimmermann, & Ebner, 2017). Previous studies, moreover, suggest that technology quality and multimedia presentation of materials are important determinants of student rentention in MOOCs (de Freitas et al., 2015; El Said, 2017; Henderikx et al., 2018). However, more evidence is needed on the possible effect of content factors in MOOCs (El Said, 2017; Hew, 2016; Hone & El Said, 2016; Ross et al., 2014), particularly on learners' perspectives in LMOOCs and whether course design factors and pedagogy impact engagement and retention (Bárcena & Martín-Monje, 2014; El Said, 2017; Hew, 2016; Hone & El Said, 2016).

MOOC courses usually involve minor or no direct interaction between tutors and students (Castrillo, 2014; Hew & Cheung, 2014; Littlejohn et al., 2016; Rodriguez, 2012; Ross et al., 2014). While it is generally acknowledged that strong teaching presence is a critical component in improving the learning experience and student participation in online learning in higher education (Hew & Cheung, 2014; Rubio, Thomas, & Li, 2018; Sokolik, 2014), the potential impact of factors such as tutor support and tutor-learner interaction has received attention in the MOOC literature (El Said, 2017; Hew, 2016; Hew & Cheung, 2014; Ross et al., 2014). Ross and colleagues (2014) argue that the discussion on MOOCs

commonly focuses on the students or the technology, but is silent on the matter of the teachers and their role. They maintain that the teacher in MOOCs is typically conceived as either "a distant celebrity figure or is automated or facilitated out of existence" (p. 67), and is not intended to be available to the students in any relational or communicative way. Hew (2016) stresses the potential for instructor accessibility to increase the likelihood of student engagement in fully online courses. In his study, he identified effective strategies used in several high-ranking MOOCs that relate to instructor presence. Based on participants' views, the findings revealed that they perceived design factors such as a clear course description including a course syllabus, which shows the specific topic in focus for each week and expected hours of workload, as engaging elements in a course. That study also emphasized that it may be crucial for learners in fully online courses to be able to seek help and clarification on a topic when needed. Other studies have shown in a similar way that most learners who drop out of a MOOC do so because they have no one to turn to for assistance, and that they feel isolated (Henderikx et al., 2018; Hew & Cheung, 2014; Hone & El Said, 2016). Furthermore, studies have found that interaction with an instructor may be a significant predictor of student retention in MOOCs (Hone & El Said, 2016). Researchers, however, call for more research on the effect of tutored factors on student retention (Hew, 2016; Hone & El Said, 2016; Ross et al., 2014). Empirical evidence on learners' perspectives towards tutor support in LMOOCs and the potential importance of such factors for their engagement and persistence is needed in particular (Bárcena & Martín-Monje, 2014; Sokolik, 2014).

The open nature of MOOCs, which attract wide variety of participants, leads to diversity in motivations and expectations among learners, who have diverse motives and personal goals, backgrounds, and prior experience, as well as diverse skills, learning strategies, and abilities that the discussion on engagement and retention in MOOCs must also take into consideration (Beaven, Codreanu, & Creuzé, 2014; de Barba et al., 2016; Kizilcec & Schneider, 2015; Littlejohn et al., 2016; Salmon et al., 2017; Sokolik, 2014, 2014). Self-determination (Ryan & Deci, 2000) has been considered a crucial factor in explaining learners' motivation and engagement in MOOCs (Chen et al., 2020; Durksen, Chu, Ahmad, Radil, & Daniels, 2016; Joo, So, & Kim, 2018). In Self-Determination Theory (SDT), Ryan and Deci (2000) distinguish between different types of motivation; namely, intrinsic and extrinsic motivation. While intrinsic motivation refers to doing something because it is "inherently interesting or enjoyable", extrinsic motivation refers to

doing something because it "leads to a separable outcome" (p. 55). Based on this, learners' motivation may "vary not only in *level* of motivation (i.e., how much motivation), but also in the *orientation* of that motivation (i.e., what type of motivation)" (p. 54, emphasis in original), where orientation of motivation means that learners' attitudes and goals serve as a call to action. In a learning context, some MOOC learners may therefore be intrinsically motivated because of individual or situational interest in a topic for the sake of understanding or to develop competence, while others may be extrinsically motivated for the sake of chosen career or in the interest of achieving high grades or academic credentials (de Barba et al., 2016; Kizilcec & Schneider, 2015; Salmon et al., 2017; Wang & Baker, 2015). Learners usually possess a combination of motivations (Salmon et al., 2017). As already noted, MOOC learners usually receive minimal personal support from an instructor (Castrillo, 2014; Hew & Cheung, 2014; Littlejohn et al., 2016; Rodriguez, 2012; Ross et al., 2014), which places the responsibility on individual learners to determine their own learning path in this self-paced online learning environment. In order to do so, learners must self-regulate their learning, which requires them to determine when and how they engage with their learning material (Littlejohn et al., 2016). Previous research (Beaven et al., 2014) indicates that many participants in MOOCs, who attend of their own free will, are well-motivated prior to starting course and demonstrate a degree of self-determination and intrinsic motivation. However, considering the diversity of MOOC learners and their reasons for attending such courses, further studies are needed on the complex relationship between student motivation and online retention (Chen et al., 2020; de Barba et al., 2016; Doiz, Lasagabaster, & Sierra, 2014; Dörnyei et al., 2014; Wang & Baker, 2015). More nuanced understanding is needed on how diverse engagement patterns in MOOCs may reflect different factors that help to spark motivation and drive learners towards the end. Similarly, it is essential to explore a little-known issue which concerns the views of learners who had the initial goal of completing but stepped out of the course earlier. Such users may provide valuable information as to what factors prevented them from completing a course (El Said, 2017; Reich, 2014). For this, students themselves are the most significant informants (Colpaert, 2010; Doiz et al., 2014; Salmon et al., 2017).

These are among the specific research problems addressed in the study. The main objective of this research project is to identify critical determinants of student retention in LMOOCs and to discover efficient engagement strategies in the environment of autonomous online second language learning. The study addresses several research

questions using data from the population of students who were enrolled in one or more of the seven courses or different modes of delivery in the IOL program from 2006 to 2018. The primary focus is on developing an understanding of what factors have a substantial impact on retention in such learning contexts, whether they are attributed to content-specific or mode-specific factors, or to other outside factors ascribed to learners' individual motives or circumstances. To address these issues, a mixed-methods study was carried out using multiple data sources, including tracking data, survey data, and qualitative data elicited through open question forms in a survey.

First, to address the issue of student retention in open online language courses, cumulative tracking data were examined based on questions on a) the level of engagement and the overall retention in all the IOL courses, b) whether the mode of delivery affects student retention, and c) what the overall engagement pattern in IOL suggests about retention. These questions were addressed and findings presented in Article I, entitled "The impact of different modalities on student retention and overall engagement patterns in open online courses", which was published in the journal *Computer Assisted Language Learning* in 2018. The study thus explored a large set of tracking data that included approximately 43,000 learners in IOL which had been collected over eight years through IOL's tracking system on the overall progress and engagement patterns in all seven courses in the program. Since two of these courses are delivered in three different modes that are all online and based on identical learning materials, they provided the unique opportunity to compare tracking data between different modes on the possible impact of the mode of delivery on student retention. Furthermore, with the use of LA, retention data were further examined in order to obtain a more nuanced picture of the engagement behavior of those who did not remain to the end of the courses or modes. The evidence gathered in this first part of the study was thus intended to provide baseline information for the two subsequent studies (Articles II and III), which focused on learners and their engagement with the learning material in one course, IOL 2.

Second, to address the question about the possible influence of mode-specific factors on student retention, survey data (n = 64) were examined on learners' experiences in two tutored modes of delivery in IOL 2, that is, the blended and distance modes, in relation to the tracking data (n = 64) on the same learners. This examination was based on the questions of a) whether learners in these two modes considered certain tutor-specific elements that are provided in the modes important for their motivation to carry on in the

course, and b) whether those who considered these particular factors important were more likely to complete the course than those who considered them unimportant, as measured by the tracking data. These issues were dealt with in the second article entitled "The effect of tutor-specific and other motivational factors on student retention on Icelandic Online", published in the journal *Computer Assisted Language Learning* in 2019. Two additional questions were also addressed in Article II. First, the question of students' initial intent in terms of course engagement and the potential impact on retention was addressed, which drew on survey data (n = 400) from all learners in IOL 2 as well as tracking data (n = 400) from the same learners. The other additional question in Article II asked those who had completed IOL 2 to the end to explain the key factors that drove them towards completion. For this the study explored text data from 112 informants in IOL 2, which were elicited through an open question section in a survey.

Third, to address the question of the value of the instructional design and pedagogical methodology for engaging the IOL 2 learner, and the effect that this has on student retention, survey data (n = 400) in relation to tracking data (n = 400) on the same learners were investigated in regard to a) whether or not learners considered certain content-specific factors in the course important for their motivation to carry on in the course and b) whether those who considered these factors important were more likely to complete than those who considered them unimportant, as measured against the tracking data. This matter was addressed in the third article, entitled "The effect of content-related and external factors on student retention in LMOOCs", which was published in the journal *ReCALL* in 2021. In addition, Article III addressed the issue of what non-completers who had the stated intention of completing IOL 2 considered to be their main reason for failing to do so. In that regard, the study explored text data from 62 informants in IOL 2 elicited through an open question section in a survey. In the following chapter the key issues in this area of the study are addressed.

1.2 Second Language Learning Online: Main Challenges

LMOOCs have placed the spotlight on the autonomous language learner and the challenges he or she faces in engaging with a course in such self-directed learning environments (Godwin-Jones, 2011). LMOOC developers are able to promote the retention of second language learners through engaging instructional pedagogy and tutor support, though the influence of individual motives and non-course related factors on

learner progress and engagement must also be acknowledged. This study addresses several gaps, namely the lack of a more nuanced understanding of student retention in LMOOCs. For this reason the study examined users' engagement patterns across several courses and modes, including the effect of mode of delivery on retention. Also addressed are the effects of diverse course-related factors on retention, such as those associated with instructional methodology and tutor support, as well as other motivational or external factors.

The low retention rates typical in MOOCs (Chen et al., 2020; Ingolfsdottir, 2014; Jordan, 2014, 2015) have raised concerns about whether such courses are suitable learning environments for most learners (Adamopoulos, 2013). Others, however, have pointed out that traditional approaches to measuring retention in higher education may not apply to MOOCs (Sokolik, 2014), given the fact that the courses are aimed at highly diverse target groups (Bárcena & Martín-Monje, 2014; Colpaert, 2014; Ingolfsdottir, 2014) and participants are not obliged to complete or engage in a course. Similarly, the evaluation of student retention without taking into account their intentions in terms of course engagement has been questioned, as well as the tendency to overlook participants who have not completed courses to the end (Henderikx et al., 2017; Koller et al., 2013; Perna et al., 2014; Reich, 2014) but may even have completed a great part of a course's content (Greene et al., 2015). Different means of delivering such courses online have also been a focal point in the literature where the central debate concerns the discussion on what kind of delivery mode benefits learners the most and is most successful in retaining them (Bettinger et al., 2017). In studies where student retention in different modes of delivery has been compared, such as in blended and distance courses, a blended mode of delivery has been found to be most effective in retaining students (Garrison & Kanuka, 2004; Harker & Koutsantoni, 2005). The literature on this subject, however, is confusing and suffers from inconsistent use of terminology; for example, the meaning of the term 'blended learning' (Colpaert, 2014; Vorobel & Kim, 2012). In addition, the referencing and comparison of results from different studies is complicated because some studies compare student experiences and retention in courses with inequivalent modes, such as an online course with a face-to-face learning course (Patterson & McFadden, 2009), or in courses with the same delivery modes but diverse online programs (Levy, 2007). These are among the specific issues addressed in the present study (Article I) which used a tracking device built into the CALL software (Fischer, 2007; Garrett, 1991) to collect large amounts of tracking data and applied the mining of these data and LA in order to reveal how

learners in IOL progress overall and by mode of delivery, as well as to identify engagement patterns that emerge from the data. While retention in other studies may be measured by course event logs (Castrillo, 2014; Koller et al., 2013; Perna et al., 2014; Reich, 2014), such as viewing a lecture, completing an assignment or posting to a discussion forum, this study is grounded on tracking a student's actual involvement throughout a course where the instructional system monitors activity on every content page in a course. While there is a limited understanding of user progress through a MOOC beyond the first and last course events (Perna et al., 2014), this study explores user progress from the first milestone to the last (starting point and completion) in several courses, based on tracking data, and therefore addresses those who may not have completed a course to the very end. Furthermore, while most empirical CALL research consists of small-scale preliminary studies performed on a small number of students, focusing on a short period of time, and which are not usually followed up (Gillespie, 2020), this study measures student retention in a CALL program founded on a large set of tracking data collected over eight years, whereby retention is monitored in seven equivalent courses, all of which employ the same instructional methodology. Furthermore, this study follows up the analysis of these tracking data with a survey of the learners in one course about the learning environment. While retention in other studies may be measured without considering students' initial intentions in terms of engagement (Belanger & Thornton, 2013; Jordan, 2014, 2015), this study investigates student intent and the impact of this factor on actual retention as measured by the tracking data. While other studies may also compare student experiences and retention in non-equivalent modes of delivery (Patterson & McFadden, 2009), or in courses with different online materials (Levy, 2007), this study not only compares the same course delivered in different modes, all online, but also breaks down retention data with a view to reveal a more nuanced picture of the online behavior of the students in different modes of delivery who do not stay to the end of the course (Greene et al., 2015; Martín-Monje et al., 2018). Furthermore, this study tracks retention in view of different coverage of course content where parameters used to measure retention are not only adjusted to a 100% completion of course content, but also, for example, to 90% to 99%, or 75% to 89% coverage of course content. This was done in order to recognize the overall engagement pattern in different courses and modes of delivery, and to reveal the extent to which non-completers engaged with the course material. Accordingly, this study considers at what point during the course learners drop

out (Greene et al., 2015), and also addresses the issue what 'course completion' actually means (Hew, 2016; Koller et al., 2013; Perna et al., 2014; Reich, 2014). Instead of defining course completion as 100% coverage of a course's content, it is defined as completion of 80% to 100% of a course's content in the follow-up studies. In a similar vein, this study also considers the definition of 'experienced learners' (Perna et al., 2014) in research of a learning environment by distinguishing data from those who may only be registrants or leave in the very beginning of a course, and those who may provide evidence based on actual experience with the learning material in focus; the non-experienced group was excluded from the study.

The need for new and engaging forms of pedagogy, and design strategies in order to improve retention in MOOCs has, furthermore, been argued in the context of low retention rates (Castrillo, 2014; Colpaert, 2014; de Freitas et al., 2015; El Said, 2017; Hew, 2016; Hone & El Said, 2016; Kim et al., 2017). More evidence has also been called for to better understand how course design properties and pedagogical practices affect student engagement and outcome (El Said, 2017; Hew, 2016; Hone & El Said, 2016; Ross et al., 2014). For the purpose of promoting engagement and to facilitate learning, researchers (Dörnyei et al., 2014; Garrett, 1991; Hubbard, 2012, 2013; Kizilcec & Schneider, 2015; Ross et al., 2014) have thus stressed the importance of curated content for the autonomous online language learner where the course content is organized and structured, the learning objectives and expectations are put forward to the learner and the learning material is carefully scaffolded and adapted to the learner's language level (Arnbjörnsdóttir, 2004; Ross et al., 2014; Teixeira & Mota, 2014). Moreover, it has been highlighted that the complex undertaking of teaching and learning a language calls for a particular LMOOC platform (Castrillo, 2014; Colpaert, 2014; Sokolik, 2014) where technology and pedagogy are interwoven. Computer-assisted language programs have, in this respect, the potential of using various means to provide student-centered pedagogy and aid the language learner in developing the skills that can lead to a successful self-guided language study (Chun, 2012; Colpaert, 2010, 2014; Godwin-Jones, 2017). Similarly, the necessity of tutor support and guidance for the autonomous language learner has been underlined (El Said, 2017; Hew, 2016; Hew & Cheung, 2014; Höfler et al., 2017; Ross et al., 2014), and thus suggested that properly introducing learners to course prerequisites and syllabuses, as well as providing them with access to a tutor assistance and support, may encourage course engagement in MOOCs (Henderikx et al., 2018; Hew, 2016; Hew & Cheung, 2014; Hone & El Said,

2016). These are among the specific issues addressed in the study which on the one hand applied a survey questionnaire to elicit LMOOC learners' views towards the value of specific content-related and tutor-related factors in keeping them in the program, and on the other hand utilized data from a tracking system to evaluate the effect of these factors on retention. While more information is needed on learners' experiences with content factors in LMOOCs and whether they act as engaging elements for the learners (El Said, 2017; Hew, 2016; Hone & El Said, 2016; Ross et al., 2014), this study explores the learners' views towards specific content-related factors in IOL 2 and their influence on actual retention. While personal contact between learners and teachers in MOOCs is highly unusual and very little is known about learners' experiences with tutor presence and interference in LMOOCs, and the potential impact on student engagement and retention (Bárcena & Martín-Monje, 2014; Hew, 2016; Hone & El Said, 2016; Ross et al., 2014; Sokolik, 2014), this study investigates learners' experiences of tutor-related factors in IOL 2, including the option of personal contact between tutor and learner and individualized help for the learners, and the effect on retention, as measured by the tracking data.

The fact that diverse MOOC learners attend a course with various motives should be considered in the discussion of student retention, and calls for more studies on how motivation promotes student engagement and drives them towards course completion (Beaven et al., 2014; de Barba et al., 2016; Ingolfsdottir, 2014; Kizilcec & Schneider, 2015; Littlejohn et al., 2016; Sokolik, 2014). Considering that students attend MOOCs with variable intentions in terms of course engagement, the question has also been raised as to why learners who had the initial goal of completing such a course disengage before completing (El Said, 2017; Reich, 2014). The learners themselves are the most significant informants needed to shed light on these issues (Colpaert, 2014; Doiz et. al., 2014). These questions are dealt with in the present study (Articles II and III), which elicited learners' written descriptions on the matter through open question forms in a survey. While there is a lack of research considering learners' own views on why they completed an online course (de Barba et al., 2016; Kizilcec & Schneider, 2015), this study attempts to elicit learners' thoughts on the reason why they completed IOL 2 (Article III). It also addresses another gap in the literature: namely, the lack of research examining committed learners' own views as to why they do not complete such courses (El Said, 2017; Reich, 2014) (Article III). This study hence makes an effort to reveal learners' self-reports on the reason why they did not complete IOL 2 as they intended to do.

As will be presented in Chapters 2.2.1 and 3.2.4, the analysis of registration data from learners in IOL 2 revealed that learners' mean age is relatively high, and when analyzed by mode of delivery, the learners in the distance mode were found to have the highest mean age in comparison to the learners in the other two groups. In light of considerations in the literature on whether age and gender differences exist in the acceptance and use of new technologies in learning (Khechine, Lakhal, Pascot, & Bytha, 2014; Wang, Wu, & Wang, 2009), and in view of the fact that learners' mean age is relatively high in IOL, a decision was made to consider two demographic factors in the process of interpreting results in the study, that is, age and gender, in order to explore the possible impact of these factors on student retention. Previous findings indicate (Khechine et al., 2014; Wang et al., 2009) that older students are more concerned about the use of new technologies in an academic setting, and may also be in more need of facilitating conditions and technological support than the younger students. Findings also show (Khechine et al., 2014) that gender does not have a significant effect in terms of technology acceptance in this context. From that perspective, this study makes an effort to shed light on whether learners' age and gender in IOL 2 may have a predictive value in relation to student retention.

In summary, the existing literature on the area of this research highlights the challenges in higher education for the autonomous second language learner in open online courses and the concerns about the low student retention typical in such learning environments. More research is needed on the online behavior of these students and their progress through a course. Greater understanding is also needed on what factors may be crucial to promote course engagement, whether they are assigned to content-specific or mode-specific factors or even individual factors outside the learning material itself. This study addresses several gaps in the literature in the attempt to provide new data on influencing factors of student retention. First, the study presents data on student retention in seven equivalent courses where tracked retention is measured from the time they enter a course until they leave. The study also measures tracking data on the effect of different modes of delivery on retention and thus compares student retention in the same course delivered in three different modes, all online. Furthermore, the study breaks down retention data in order to reveal a more nuanced picture of the online behavior of the students who do not complete a course, and thus provides evidence on user progress throughout a course. Consequently, the study reconsiders the parameters commonly used to evaluate

course completion in MOOCS in subsequent survey studies. Moreover, the study provides survey data on learners' experiences towards the impact of various course-related factors on learners' engagement and the impact of these factors on tracked retention, such as content and tutor interference, and motivational factors such as the initial intent to complete a course. Finally, this study addresses the issue of student retention by exploring qualitative data on learners' thoughts on the reason why they completed a course, or left earlier. The following section presents the IOL program, which provides the source on which the research is based.

1.3 The Icelandic Online Project

The IOL program offers seven consecutive, free, asynchronous, self-guided online courses for adult learners, all of which are delivered in open non-tutorial modes. In addition, two IOL courses are also offered in tutorial modes of delivery, in which learners receive the support and guidance of a tutor. The courses are fully online and interactive, and are curated and sequential. They are skill-based and linked to the levels of the Common European Framework of Reference (CEFR) (Council of Europe, 2018). All the IOL courses are under investigation in this study. This large web-based program (pre-web 2.0), created for desktop and laptop computers, is provided in an independent design module, specifically developed for the IOL project (Arnbjörnsdóttir, 2008; Arnbjörnsdóttir, Friðriksdóttir, & Bédi, 2020), and contains features specifically designed for language learning and teaching (Colpaert, 2014). The initial decision to develop and exploit an individual web-based learning module for IOL, and not to use an available language learning app adapted to Icelandic, relied on the IOL's project policy to provide a holistic program founded on accepted and relevant theories of second language learning and pedagogy, and would give an opportunity to provide a rich variety of activities and learning objects in the learning context. Additionally, the intention of collecting and storing cumulative tracking data from the end user called for the use of the Internet. The IOL courses are defined here as language MOOCs, based on the following characteristics: First, the IOL program was specifically designed for online teaching and learning of a second or foreign language, in which the curriculum is entirely contained in an independent design platform. Second, IOL comprises a SLA theory-based CALL program, fully online and interactive, as well as a curated and sequential curriculum. Finally, the seven IOL courses are skill-based and are linked to the levels of CEFR (A1-C1) with

unrestricted[ii], free access and potentially unlimited participation. As noted in the literature, there are examples of universities or individual higher education institutions that offer their own MOOCs, without any connection to any of the large MOOC platforms, including some that already existed before the MOOC movement started, and so are not referred to by that name (Gaebel, 2014).

In the following chapters, IOL is presented from three perspectives: First, Chapter 1.3.1 discusses the background of the project and reception. Then, Chapter 1.3.2 focuses on the development and the courses of the IOL project. Finally, Chapter 1.3.3 is divided into three sub-chapters which shed light on IOL with regards to the components under investigation and their relevance to the goals of this study. Thus: Chapter 1.3.3.1 reviews IOL's instructional methodology and course content; Chapter 1.3.3.2 describes the different modes of delivery offered in IOL; and Chapter 1.3.3.3 presents IOL's tracking system.

1.3.1 Background and reception

The development of the IOL program, which began in 2001, was initiated with the main aim of providing scholars and Icelandic lecturers abroad with new and accessible online materials in Icelandic, as well as offering anyone with Internet access the possibility of participating for free in a global community of learners of Icelandic (Arnbjörnsdóttir, 2008). In this respect, the development of the program is rooted in a long history of the teaching of old and modern Icelandic at universities abroad, supported by the Icelandic government. Modern Icelandic is now taught at 40 universities in Europe, America, and Asia, where Icelandic authorities currently support the teaching of modern Icelandic in Austria (Vienna), Canada (Winnipeg), China (Beijing), Denmark (Copenhagen), Finland (Helsinki), France (Caen, Paris), Germany (Berlin, Kiel, München), Norway (Bergen), Sweden (Gothenburg, Lund, Uppsala), and the UK (Cambridge, Edinburgh, London).

IOL is a collaborative project developed at UI with the participation of the Centre for Research in the Humanities, the University's Faculty of Icelandic and Comparative Cultural Studies, the Árni Magnússon Institute for Icelandic Studies, and the Vigdís Finnbogadóttir Institute of Foreign Languages, in collaboration with several international universities that offer courses in Icelandic. Funding for the project has been provided by EU's Lingua Project, Nord Plus (Nordic Languages and Adult), the Rector of the University of Iceland, the Icelandic Research Fund, the Icelandic Ministry of Education,

the Árni Magnússon Institute for Icelandic Studies, the University of Wisconsin-Madison, the Directorate of Labour, and the Vigdís Finnbogadóttir Institute, which has housed the project from the outset.

The implementation of IOL marks an important milestone in the context of Icelandic higher education since IOL is the first, and still the only, program that offers consecutive and continuous online learning courses in Icelandic as a second language. The initial stance taken to offer free and open access to the IOL courses, in times long before the MOOC movement, is now seen to be advantageous and may partially explain the constantly growing student participation in the courses. The courses have benefitted thousands of learners around the world; currently, more than 260,000 have enrolled in one or more IOL courses worldwide since the launch of the first course in 2004. This growth, and the concurrent increase in enrollment in the on-site program in Icelandic as a Second Language at UI, led to further expansion of the program than was intended initially (Chapter 1.3.2). The IOL program received special recognition from the Icelandic government in 2014 for its contribution and support in maintaining the Icelandic language. Single courses have, furthermore, been honoured; Icelandic Online 2 received the University of Iceland's annual Applied Science Prize in 2005 (second place), and the Survival IOL course received recognition from the Icelandic Language Council in 2019 for its support to newcomers in Iceland. The following section discusses the development and use of the IOL program.

1.3.2 Development of IOL and the courses in focus

The initial idea behind the development of the IOL program was to provide only one course for beginners and to offer it through the means of two different themes – a Culture theme and a Nature theme – with the intention to appeal to the varied interests of students of Icelandic. However, a positive reception of this course both in Iceland and abroad called for further development, which led to the development of all seven IOL courses that were launched between 2004 and 2013. The first course that was developed, Icelandic Online 1 (IOL 1), was thus launched as two identical beginner courses in 2004 (IOL 1 Nature theme and IOL 1 Culture theme). Icelandic Online 2 (IOL 2) was launched in 2005, and IOL's integral tracking system was implemented in 2006. Three more courses were then launched in 2010; Icelandic Online 3 (IOL 3) and 4 (IOL 4), and the Survival IOL course. Finally, Icelandic Online 5 (IOL 5) debuted in 2013. As previously mentioned, all the courses, IOL

1–5, are provided in non-tutorial, self-directed modes of delivery with an open and free access. In addition to that, two of these courses, IOL 1 (Culture) and IOL 2, were specifically chosen for further development in order to offer learners the option of tutor support. These two courses were extended to provide a) credential tutorial courses (10 ECTS each) to be used in a new Practical Diploma Program at UI, referred to as 'blended courses' in the study, and b) tutored non-credential courses delivered at a distance, referred to as 'distance courses' in the study. These courses were released in 2008. All seven courses are included in the first part of the research (Article I), including different modes of delivery of IOL 1 and 2, while the two other parts of the study focus on one course, IOL 2 in different modes of delivery (Articles II and III).

The seven IOL courses are offered at five proficiency levels (A1–C1) where the learner has the option to continue from one level to the next: a) Survival IOL is aimed at newcomers in Iceland who are absolute beginners, and focuses on practical communicative needs, b) IOL 1 – Nature theme and c) IOL 1 – Culture theme are both beginner courses, CEFR level A1, d) IOL 2 is a lower intermediate course, CEFR level A1–A2, e) IOL 3 is an intermediate course, CEFR level A2–B1, f) IOL 4 is a higher intermediate course, CEFR level B1–B2, and g) IOL 5 is an advanced course with a primary focus on reading Icelandic literary texts at CEFR level C1. The courses vary in length and contain 29 to 139 content pages depending on the course, where each content page includes three to ten learning objects. (Please see Appendices A and B for further description of the courses and Chapter 1.3.3.2 for details on different modes of delivery.)

It must be noted that after the present research was undertaken, the IOL system was upgraded to a multi-platform system compatible with computers, tablets, and mobile phones, with a view to improve access to learning and meet the changing technical needs of its target users. For this, the prior web-based desktop/laptop program of IOL was adapted for use on mobile devices in order to be used specifically as a web-based app. Consequently, six out of the seven previously discussed IOL courses, including three different modes of delivery for two of the courses, are now available in a new multi-system for mobile devices. These new IOL courses, which were released from 2016 to 2018 and are accessible via the link http://www.icelandiconline.com, have now replaced the previous version of IOL (http://www.icelandiconline.is).[iii] At the same time, the new IOL project was reproduced for other languages – Faroese and Finland Swedish as second languages – in collaboration with the University of the Faroe Islands and the University of Helsinki.

Faroese Online (https://faroeseonline.com) was released in 2016 with the Survival IOL course used as a prototype. For Finland Swedish Online (https://finlandswedishonline.fi), two courses were launched in 2017 and 2019, using Survival IOL and IOL 1 as prototypes. The above-mentioned collaborators made use of the technical part of IOL as well as the pedagogy and design, and either translated the course content or adapted it to the appropriate target group(s). It is worth noting that no major amendments were made to the content or organization of the course content in the new version of IOL compared to the former one. The three following subsections focus on the instructional framework of the IOL program, the three different modes of delivery, and IOL's tracking system, all of which are the core of this research on influencing factors in regard to student retention and online behavior.

1.3.3 IOL's instructional framework, delivery modes, and tracking system

The initial stage of the research project (Article I) addressed tracking data on user progression and overall engagement patterns in the seven IOL courses, including the three different modes of delivery in IOL 1 and 2. In accordance with the main objective of the research, the two follow-up studies that were then carried out attempted to investigate student retention in the context of the course itself and the users; namely the possible impact of tutor-related factors (Article II) and content-related factors (Article III) on retention. It should be noted that the organization of the following presentation of content and modes is slightly modified from the order of the published articles. In this overall introduction of the study and its results, it was considered to be more comprehensible for the reader to get insight first into IOL's course content, the design and instructional methodology (Chapter 1.3.3.1), and then to examine the different modes that the course content is mediated through (Chapter 1.3.3.2). The presentation of the results will thus follow the same order (Chapters 3.2.1 and 3.2.2). The discussion of IOL's tracking system in Chapter 1.3.3.3 then concerns the study as a whole, as discussed in all three articles.

1.3.3.1 Theoretical foundations and content-specific factors in focus

This section focuses first on the design behind the IOL program and summarizes the theoretical background and methodological decisions made in the design process. Next, the use of technology as a medium for teaching and learning is outlined in the context of IOL. Finally, the section presents the associated factors related to the course content that were under investigation in the study.

IOL was developed primarily with two groups of learners in mind: Scholars and other adult beginner learners of Icelandic who are interested in the language and culture, and exchange students at UI who benefit greatly from easy access to language education and support in their studies. Among the main challenges encountered in the development of IOL's instructional design and pedagogy were presenting complex grammar in a meaningful context online, motivating and engaging self-directed adult learners with the course material, and ensuring ease of interaction with the CALL tool for learners of all levels of computer literacy (Arnbjörnsdóttir, 2004, 2007). From a constructivist and sociocultural view of language learning and instruction, which assumes networked language learning and a minimal capacity for active language use (Arnbjörnsdóttir, 2004; Kern, Ware, & Warschauer, 2004), the teaching of morphologically complex languages online poses specific challenges for current approaches in CALL. Considering the morphological structure of Icelandic, this issue especially pertains to the teaching and learning of the language at the beginner level, where the recognition of a number of forms of the same nominal is a prerequisite for a basic level of comprehension (Arnbjörnsdóttir, 2004).[iv] This means that a beginner learner in Icelandic can not easily continue in a communicative language course without an explanation of all of the different forms. Efforts to implement solutions to the presentation of inflectional morphology in a meaningful context in IOL are inspired by Chapelle's model of 'relevant SLA hypothesis' in reconciling CALL and the multitude of different approaches to SLA theory of acquisition and pedagogy (Arnbjörnsdóttir, 2004; Chapelle, 1998, 2003).

The theoretical criteria adopted by the developers of IOL is based on Schmidt's (1993) ideas of attention and noticing, which emphasizes that the learner must be aware of the feature of the target language in the input in order to learn it, and on Chapelle's (1998) suggestions of using online devices to help the learner notice specific elements of the input (Arnbjörnsdóttir, 2004). In that regard, the approach of focus on form (Doughty & Williams, 1998) guided the development of IOL (Arnbjörnsdóttir, 2004), meaning that learners' attention is drawn to grammatical forms in the input in a meaningful context (Ellis, Basturkmen, & Loewen, 2001). Another theoretical guideline behind the development of IOL is based on the notion of scaffolding (Bruner, 1974; Lantolf, 2000); namely, an instructor supports a learner in completing a new task and gradually places the responsibility for learning into the learner's hands as he or she gains mastery (Arnbjörnsdóttir, 2004). Scaffolding as applied to CALL is commonly understood as the

instructional support provided by a CALL program during learner-computer interactions (Heift, 2016). The instruction in IOL thus involves scaffolding of information under carefully organized step-by-step guidance intended to support the language learning process (Rosenshine & Meister, 1992). The philosophy of scaffolded instruction is evident in the structure and organization of the content in IOL (Aebersold & Field, 1997; Arnbjörnsdóttir, 2004), the gradual presentation of input, and the scaffolded presentation of grammar.[v]

The technical implementation was chosen to best serve the a priori pedagogical principles that provide the foundation of the IOL courseware. In light of the complex morphological structure of Icelandic, an effort was thus made at lower levels in IOL to use effective affordances of technology to facilitate learners' comprehension of the variety of ways in which words can change forms, and consequently, to support learners in constructing and negotiating meaning. These include 'input enhancement' with focus on form (Chapelle, 1998; Doughty & Williams, 1998) where selected forms in texts are made salient to the IOL learner. This was implemented by highlighting the focused grammar in a meaningful context where each of the different foci is assigned a different color (Arnbjörnsdóttir, 2004). Directly linked to this form-focused presentation of grammar, the system was used to provide learners with a scaffolded, bi-leveled presentation of the grammar in focus (please see below). By doing so, the design takes into account that not all learners are necessarily grammatically inclined and aims to avoid overwhelming the lower-level learner with information that may not be useful for all of them (Arnbjörnsdóttir, 2004).

The overall instructional approach in IOL reflects the intent to facilitate a gradual promotion of language learning through appealing course content, which makes an effort to engage adult learners and keep them engaged. In this vein, one of the decisions made in the design process was to pedagogically organize the material and guide the learners through structured and curated course content (Garrett, 1991) where they are introduced to explicit learning objectives. From the perspective of the architecture, the whole course was thus thoroughly planned and segmented into five theme-based sections where each section is considered as a specific learning module, or unit, with its own internal organization, particular topic and activities, and specific learning objectives to pursue. This internal organization comprises gradual and scaffolded presentation of input in three steps, or parts (Part I, II and III), in a section (please see below).

In the attempt to motivate and engage adult learners with the content in IOL, they are confronted with a range of appealing and varied learning material which reflects the authentic use of the target language, and is adapted to learners' language level and the learning objectives in focus (Chapelle, 2009; Garrett, 1991). This effort to engage the learners with the course is evident in both the choice of the subject matter in IOL and the variety of the learning objects included in the course. Storylines, or plots, as a context for language input and practice are used throughout the course in order to keep the learners engaged as they interact with a variety of oral and written texts accompanied by assorted comprehension and accuracy exercises. In the attempt to provide the broadest possible variety of activities and tasks (Colpaert, 2006), diverse characteristics of learning objects were designed and implemented, focusing on different aspects of the language and activities. Learning objects were thus specifically pre-programmed to serve the pedagogy of IOL, resulting in a wide array of exercises across the course, derived from approximately 40 different basic software design patterns or templates. Additionally, certain types of learning objects were developed to approach different inputs as well as to serve the target language from a linguistic point of view. For example, animation was considered well-suited to demonstrate how and why noun declensions result in change in word endings in Icelandic. Overall, the course was carefully planned beforehand to present different and interchanging characteristics of learning objects which account for different learning styles. The aforementioned elements of IOL are outlined further below in the overview of the factors explored in the study.

Furthermore, the IOL courseware includes multimedia resources as embedded links to online dictionaries and a database of Icelandic morphology. In addition, a glossary adapted to the course content for the lower levels was developed and included in the program (Chun, 2001).

While considering IOL's target group of voluntary learners interested in learning a lesser-taught language like Icelandic in an online environment, it was anticipated that many prospective enrollees would be highly motivated and possess a degree of self-determination and intrinsic motivation (Beaven et al., 2014; Ryan & Deci, 2000). Also taking into account that not all potential users will be used to digital learning environments (Arnbjörnsdóttir, 2004, 2007), the designers of the IOL software aimed to create a tool that would be easy to use and interact with.

Finally, an effort was made to develop and implement a tracking device, which was

built into the IOL software (Fischer, 2007; Garrett, 1991) to serve the need of collecting tracking data for future research on the usage of the program and its users. IOL's computer-based tracker is outlined in Chapter 1.3.3.3. The IOL courseware was developed by an interdisciplinary team at UI, including second language learning experts, Icelandic linguists, and experienced teachers of Icelandic as a second language at UI. Also included in the development team were content providers, artistic designers, and software engineers who are also experts in language and technology. Additionally, students of computer science at UI were a part of the development team. The web-based platform for IOL was written in PHP 5.6.40 and uses the webserver Apache 2.2.15. The built-in computer-based tracking system uses the database management system MySQL 15.1 for writing and querying data. The software design patterns and templates for diverse learning objects are developed in HTML, Java, and Adobe Flash.

Based on the particular design elements of the IOL program presented here, the study (Article III) investigated learners' views about the effect of six content-specific factors in IOL 2 in engaging them with the course and the potential impact of these factors on student retention. The factors relate to course structure and organization of the course content on the one hand, and instructional design and pedagogical principles on the other. The two factors concerning structure and organization of the course content are curated and sequenced course structure, and clear and salient learning objectives. The four factors representing the instructional design and pedagogical principles in IOL are gradual and scaffolded presentation of input, continuing storylines, form-focused and scaffolded presentation of grammar, and variety in types of learning objects. These features are summarized below.

Curated and sequenced course structure and *clear and salient learning objectives*: In accordance with IOL's pedagogical approach of guiding learners through a sequenced and structured learning path, the course material is organized and segmented into five thematic sections each with three lessons, or parts (see below). Learners are introduced to explicit section-specific learning objectives in the areas of vocabulary, grammar or language usage at the beginning of each section.

Gradual and scaffolded presentation of input: Based on the sequenced course structure and organization of the course material, the topic is introduced in three steps, or parts (Part I, II and III), throughout each section in the attempt to present the input gradually and provide scaffolding of information.[vi] Apart from the learning objectives that are presented in Part I

of each section, this part introduces, or pre-teaches, what the learner is going to focus on in a section and includes visual input and activities that involve learning goals pertaining to vocabulary, grammar, and language usage. Part II then includes texts, written and oral, where the vocabulary and grammar introduced in Part I are presented in context, followed by comprehension exercises. Finally, Part III involves a summary of what has been learned and focuses on output and affords opportunities for learners to practice the items they have been exposed to in the two previous parts of the section.

Continuing storylines: Consistent with one of the key pedagogical principles intended to motivate learners and engage them with the content of IOL, the course contains continuing storylines as a context for language input and practice. Activities in each section thus centre around a theme, or a story throughout a section, based on "authentic" texts that are adapted to learners' levels and the learning objectives in focus. For example, the learners follow a protagonist throughout a section as he goes about his academic and social life in Iceland. The storylines are supported by illustrations that are expected to appeal to a wide range of interests in order to provide a stimulating, entertaining, and engaging learning environment for adult learners.

Form-focused and scaffolded presentation of grammar: In an effort to draw learners' attention to the grammar in focus in a meaningful context and to explain linguistic features, the forms are highlighted in the text and grammatical information is then provided in a scaffolded presentation. First, in order to point out the grammar in focus, learners are provided with the option of moving their mouse cursor over the text which then displays the grammar items highlighted in the text with different colors depending on the input. When doing so, a one-liner (referred to as 'grammar help' in the survey) appears underneath, which explains in a short paragraph the basics of the specific grammatical feature. Finally, if the learners want to know more about this grammar feature, they have the option of going deeper into the grammar by clicking on the one-liner, which directs them to a larger grammar resource (referred to as 'read more' in the survey).

Variety in types of learning objects: Learning objects were specifically pre-programmed to serve IOL's pedagogical approach of providing learners with varied and engaging learning material and learning opportunities, resulting in a wide range of learning objects throughout the course. This involves different types of media and includes short videos (with subtitles), as a source of authentic language material, visual and interactive learning objects that present grammar and vocabulary, and various tasks for language practice.

Thus, the course as a whole was carefully planned beforehand to account for different learning styles by presenting different and interchanging characteristics of learning objects. The program provides audio files and feedback on nearly every learning object in the course. The next section discusses the different modes of delivery in IOL which were under investigation in the study, including the different target groups they are aimed at, and the tutor-specific factors provided in the blended and distance modes that are the focus of this study (Article II).

1.3.3.2 Different modes of delivery and tutor-specific factors in focus

MOOCs come in various forms. While some courses are designed for self-paced learning with no deadlines or tutoring, others have a fixed schedule and deadlines for completion of course components (Greene et al., 2015; Ihantola et al., 2020; Watson, Yu, & Watson, 2018). The learning material in IOL is mediated through diverse learning modes, or pedagogical models, all of which are based on self-directed asynchronous learning: An open self-direced mode, blended learning mode, and distance learning mode. As previously noted, all seven IOL courses are available in open self-directed modes in which learners receive no tutoring, but have free access to the learning materials and can begin the course and complete course materials whenever they see fit. Two of the courses, IOL 1 and 2, are offered with tutoring for a fee: The blended and distance learning courses, which include syllabuses and begin and end on fixed dates. The choice of mode depends on the learner, his or her individual goals and interests, or other motivations that a learner may bring to the course.

The initial stage of the research (Article I) compared tracking data on the overall retention and engagement patterns among the three different modes of delivery in IOL 1 and 2. Given that the three modes of each course are all fully online and include primarily the same learning materials, the IOL 1 and 2 courses provided a unique opportunity to investigate the leverage of mode of delivery on retention in the study (Bettinger et al., 2017; Garrison & Kanuka, 2004; Harker & Koutsantoni, 2005). The other phases of this study (Articles II and III) then followed up on the findings of the analysis of the tracking data, focusing specifically on factors associated with tutor support and supervision provided in the blended and distance modes of IOL 2, as well as the effect of these factors on retention (El Said, 2017; Hew, 2016; Hew & Cheung, 2014; Ross et al., 2014). The characteristics of the different modes of delivery and the factors in question are outlined below. For the sake of space, the discussion here of the different delivery modes of IOL focuses on IOL 2 specifically.

Open self-directed mode: The open self-directed mode for IOL 2 supplies an open and free access to the course, which has no fixed beginning or ending and there are no prerequisites for the course. This non-credential course guides learners through a structured learning process where they have no access to a tutor. The open self-directed course, as the other two modes of IOL 2, contains 63 content pages where each content page includes three to ten learning objects, with the total of 134 tasks to work on (Appendix A). The structure of the open self-directed mode is the base on which the other two modes, blended and distance, were subsequently developed with the added feature of a tutor. These modules, called 'Plus courses', are run independently of the basic course, although built on the same material and sequence. The Plus courses are so called because they include an additional layer of content that only applies to learners in the blended and distance modes.

Blended learning mode: The blended mode for IOL 2 delivers a self-study program on campus under the supervision of a tutor, which is integrated into a credit-bearing (10 ECTS) on-campus course at UI, supplementing traditional materials. The blended learning mode was specifically developed in order to support regular course work at UI in the two-semester Practical Diploma Program in Icelandic as a second language.[vii] In that regard, it is anticipated that the majority of learners in the IOL 2 blended course takes two other traditional courses in the Practical Diploma Program during a semester (a vocabulary and grammar course as well as a pronunciation and speech course). Students in this mode have met the requirements stipulated by UI for admission to undergraduate programs and pay an annual registration fee (469 €). The blended course is hybrid in the sense that IOL's learning materials form half of the course material and half is traditional material (mainly provided on UI's intranet), such as books and PDF files, which relates to learners' choice of an extra topic, whether they choose to focus on listening, reading, or grammar. It must be noted that the focus in this study is only on the IOL portion of the blended course, its learning material and learners' interaction with the online program. The blended mode includes 63 content pages, including 134 tasks, which is the same as the two other modes but expanded with 78 extra learning objects: so called 'Plus exercises'. All in all, the blended mode of IOL 2 therefore contains a total of 212 tasks to complete (Appendix A). Based on student evaluation in the development phases of the program, the estimated workload of the online material in the blended course is approximately 3 to 4 hours of work per week. It is scheduled as a 13-week course where learners receive planning and guidance from a tutor throughout the semester. They meet the tutor on campus for an

introductory session about the content and the organization of the course, learning objectives, and technical functions as well as support resources such as online dictionaries. They receive a set syllabus where the tutor has pre-organized the course into manageable, timetabled sections. This means that the tutor recommends how much material should be covered per week in order to complete the course within a given time frame, with an equal workload for the learner. During the course, learners have the option of consulting a tutor at all times, individually or via email. Students in the blended course submit a graded midterm paper based on the IOL part of their course (as well as their chosen elective topic part) and take a final exam on campus. In order to pass the blended course, learners must pass the midterm assignment and the final exam; the decision to complete the online material is optional. The study (Article II) investigated learners' views towards four tutor-specific factors in the blended learning mode and the potential impact of these factors on student retention, that is, the following factors: *a detailed introduction of the program* on campus, *private interaction with the tutor* face-to face or via email, *a set syllabus*, and *overall tutor support* during the semester.

Distance learning mode: The distance mode for IOL 2 was developed to meet the needs of learners outside UI who do not wish to study in the open self-directed mode and may prefer support from a tutor in such an online learning environment. The distance course is offered through UI's Language Centre as a non-credential diploma program with no prerequisites for the course, and there is an optional final exam. The distance course is scheduled as an eight-week course with a fee (247 €). Via email, students are provided with a detailed introduction about the organization of the course, learning objectives, and technical functions, and a set syllabus like the blended learners. The distance course contains 63 content pages, with 134 tasks as the other modes, and includes an additional 78 learning objects ('Plus exercises'), as mentioned above regarding the blended mode. Besides that, the distance course includes 12 built-in written assignments submitted through an editor within the course where learners receive private feedback from the tutor. Therefore, this mode includes 224 tasks for the learner (Appendix A). The estimated workload in the distance course is approximately 5 to 6 hours of work per week. Completion of the online material is optional, but in order to receive a certificate for the participation, distance learners must submit at least half of the written assignments and complete the final exam. Interactions between the tutor and learners are via email any time and in relation to individual feedback on learners' written assignments through the course

editor. The study (Article II) explored the distance learners' views towards the same four tutor-specific factors that were investigated in the blended learning mode and presented above, but in the distance mode the support was provided via email or through the IOL system: *a detailed introduction of the program* delivered via email*; private interaction with the tutor* via email and the course editor system, *a set syllabus*, and *overall tutor support* at a distance during the course.

In the discussion of the two tutored modes of IOL, blended and distance, that now have been considered, two more elements should be mentioned. The blended and distance courses are specifically developed as independent modules in a so-called 'Plus system' within the IOL system itself. In order for learners to become 'blended learners' or 'distance learners' the tutor needs to register them in the relevant mode of the Plus system. Furthermore, the Plus system incorporates a specific integral tracking system, which has the potential of tracking learner progress in the Plus courses. The collection of these tracking data, however, relies on students' self-evaluations during the course, where they are offered to mark whether they have completed all, some, or none of the tasks on each content page as they move along the course. These data on learners' progress are apparently unreliable, since they depend on the students' initiative but nevertheless give the tutor the opportunity to contact individual learners early if there are signs that a student is struggling. The following section describes the operation of IOL's tracking system which monitors learners' interactions with the program and partly gathers data for the research project.

1.3.3.3 Tracking system

In accordance with the primary goal of IOL's tracking system to provide course developers and instructors with feedback on the users and their usage of the learning resources, the system tracks and automatically records data based on the one hand on user registration, and on students' actions, or progress, on the other. User registration data are collected on a registration form on the website itself. The system collects users' email, password (encrypted in the database), name, gender, birth year, education status, country raised in, and first language(s). Regarding learners' progress through the program, the tracking system collects data on each user based on ID number and records a) when a user first logs into the program (date registered), b) where he or she is situated in a course (on which content page he or she was last active), and c) when a learner logs out or leaves a course (date last active). If a user is enrolled in multiple courses, the system captures activities in

all the courses. The administration of the database management system is based on the structure of the course content and captures users' access logs when they click on each content page in a course. As discussed in Chapter 1.3.2, the IOL courses vary in length, and contain from 29 to 139 content pages, depending on the course (Appendix A). Each content page in a course has its reference number in the database management system (which is also visible to the learner). The first content page in each course has the number 111 (section 1, lesson 1, sub-lesson 1) and the last content page in IOL 2 (including 63 content pages), to name an example, has the reference number 534 (section 5, lesson 3, sub-lesson 4). Therefore, when student retention in IOL 2 is monitored, the position of a learner on content page 63, with the reference number 534 in the database, means that he or she has completed the very last content page of the course. The tracking system stores the registration data along with the retention data in the database accumulating massive amounts of data on users and their online behavior. The database management system gives a data specialist the opportunity to export reports from IOL's server to Excel for analysis.

1.3.4 Summary

To summarize, the IOL program, which is the primary source of evidence for this study, provides seven open self-instructed and online courses for second language learners. The program is specifically designed for language learning and teaching, founded on accepted theories of SLA and pedagogy, where the curriculum is all contained in an independent design platform which functions as a technologically-enhanced device to support the learning process. IOL's methodology and instructional design aims to facilitate a gradual promotion of language learning and to engage the learner with the course by various means. The key pedagogical principles used in IOL for this purpose can, first, be seen in the structure and organization of the course content where the curriculum is curated and sequenced based on salient learning objectives, and thus reflects IOL's ideology of guiding the learners to follow a sequenced and structured learning path. The two elements representing the structure and organization of the course that are under investigation in the study, with the focus on IOL 2, are the factors of curated and sequenced course structure and clear and salient learning objectives. Second, IOL's instructional design and pedagogical principles can also be viewed in the gradual presentation of input in three steps (Parts I, II, III) and the three-tiered, form-focused, and scaffolded presentation of grammar in context. Finally, IOL's principle of motivating and engaging learners with the

course content can also be seen in the topics chosen, which center around continuing storylines, as well as the rich variety of tasks throughout the course. The four elements representing the instructional strategies in IOL 2 that were explored in the study are the factors of gradual and scaffolded presentation of input, form-focused and scaffolded presentation of grammar, continuing storylines, and variety in types of learning objects.

Two of the IOL courses, IOL 1 and 2, are delivered in three different modes and consequently as three individual courses: as an open self-directed course (non-tutorial); a blended course (tutorial); and a distance course (tutorial). These courses are all fully online and mainly based on the same learning materials which made them ideal to investigate the leverage of three different modes of delivery on student retention in the study. Two of the modes, blended and distance, involve a tutor who provides learners with various means of maintaining focus during the learning process, assist them with the learning materials and technical issues and offer individual guidance throughout the course. Four tutored factors in IOL 2 and their potential impact on learner retention were investigated in the study, that is, a detailed introduction of the program, private interaction with the tutor, a set syllabus, and overall tutor support.

IOL includes a tracking system, which monitors each learner's progress throughout a course and stores all the data into a database. The study's focus is on overall student retention and engagement patterns in the IOL program, including the different modes of delivery, as well as on learners' views about content-specific and mode-specific factors in one of the courses, IOL 2, and the effect of these factors on retention as measured by the tracking system. In addition, the study explored other motivational and external factors outside the learning material itself and the possible impact on student retention. The following section presents the research objectives, research questions, and relevant background.

1.4 Aim of the Study and Research Questions

The main objective of this study is to illuminate crucial factors of student retention in online second language learning courses in the attempt to identify efficient engagement strategies in such learning environment. The research drew on data from the IOL program where the primary interest was in recognizing which course-related factors may have a significant impact on retention in this learning context, whether they are assigned to content-specific or mode-specific factors, or to other outside factors attributed to learners' individual motives or circumstances. To address these issues, the study used mixed

methods where the adopted methodology was inductive, and whereby several hypotheses were developed. The overall aim was met by achieving the six objectives introduced below.

As discussed in Chapters 1.1 and 1.2, the issue of overall low student retention in MOOCs has raised concerns in the literature (Chen et al., 2020; Ingolfsdottir, 2014; Jordan, 2014, 2015) and called for further studies on diverse influencing factors that may affect retention, such as what kind of delivery mode engages learners the most in such an environment (Bettinger et al., 2017). The existing literature mainly includes comparison of student experiences and retention in courses with inequivalent modes (Patterson & McFadden, 2009), or comparison of diverse online programs (Levy, 2007). This study therefore attempts to provide new knowledge of the potential impact of different modes of delivery on retention by comparing courses that are provided in three different delivery modes, all online, and also include the same learning material. More evidence is also needed about actual retention in MOOCs as measured by a computer-based tracker (Fischer, 2007; Chun, 2013), and patterns of user progression throughout a course (Ebben & Murphy, 2014; Godwin-Jones, 2017; Hone & El Said, 2016; Martín-Monje et al., 2018; Thomas & Gelan, 2018). Course engagement and completion in MOOCs is commonly measured by course event logs (Castrillo, 2014; Koller et al., 2013; Perna et al., 2014; Reich, 2014) and thus focuses less upon the proportion of learners who do not go to the very end of such courses (Chen et al., 2020; Greene et al., 2015; Koller et al., 2013; Perna et al., 2014; Reich, 2014). Based on this context, the first objective and relevant research questions addressed in the study were as follows:

Objective 1

The first objective of this study was to reveal actual student retention in the IOL program, both overall in the seven courses and with regard to three different modes of delivery in IOL 1 and IOL 2, in order to uncover the potential impact of the mode of delivery on retention, as well as to illuminate what the overall engagement patterns suggest about retention. This was performed by mining tracking data and applying learning analytics on approximately 43,000 users who were enrolled in one or more of the courses or modes from 2006 to 2014. The results were presented in Article I. Data analysis regarding the second research question in (A) below was grounded on the established theory that mode of delivery affects retention and that a blended mode is more effective in retaining students than a distance mode (Garrison & Kanuka, 2004; Harker & Koutsantoni, 2005). Based on

previous research, it is therefore expected that the present study will reveal overall low retention in the IOL program, and that the blended mode will be more effective in retaining learners than the other modes in the study. The research questions were as follows:

(A) 1. What is the overall student retention in the IOL courses?

2. Does the mode of delivery affect student retention. If so, how?

3. What does the overall engagement pattern in IOL suggest about retention?

As highlighted in Chapters 1.1 and 1.2, the literature has drawn attention to the value of the organization of course content (Dörnyei et al., 2014; Garrett, 1991; Hubbard, 2013; Kizilcec & Schneider, 2015; Ross et al., 2014) and instructional pedagogy involving scaffolding of the learning input (Arnbjörnsdóttir, 2004; Castrillo, 2014; Rosenshine & Meister, 1992; Sokolik, 2014; Teixeira & Mota, 2014) to the facilitation of learning in LMOOCs. The value of CALL for the LMOOC learner has been stressed in this context, especially the potential of effective interlacing of technology and pedagogy (Chun, 2012, 2016; Colpaert, 2010, 2014, 2018; Godwin-Jones, 2017). However, more empirical evidence is needed on how technologically enhanced devices and instructional methodology may assist the learner in his or her learning process (Chun, 2012, 2016; Colpaert, 2010, 2014, 2018), and whether course design factors and pedagogy influence student engagement and retention in LMOOCs (Bárcena & Martín-Monje, 2014; El Said, 2017; Hew, 2016; Hone & El Said, 2016). Founded on this context, the second objective and relevant research questions of this study were as follows:

Objective 2

The second objective of the study was to illuminate whether a learner's experience of specific factors related to the course content and instructional design in IOL 2 is important for their motivation to engage with the course, and to reveal a potential influence of these content-specific factors on student retention. This was performed by gathering and combining survey data and tracking data for analysis from 400 users who were registered in the course from 2010 to 2018. Data collection and analysis regarding the research questions in (B) below related to suggestions that MOOC learners benefit from curated structure and apparent organization of a course's content and instructional pedagogy in order to engage with the course (Castrillo, 2014; Colpaert, 2014; Garrett, 1991; Hew, 2016; Hubbard, 2013; Rosenshine & Meister, 1992; Sokolik, 2014), and that

technologically enhanced devices have the potential to support the language learner in LMOOCs (Arnbjörnsdóttir, 2004; Chun, 2012, 2016; Colpaert, 2010, 2014; Godwin-Jones, 2017). Based on the literature, the present study predicts that the specific elements of the IOL design concerning the course structure and instructional methodology will have a positive impact on student engagement and retention. The findings were introduced in Article III. The following research questions were thus formulated:

(B) Do learners in IOL 2 consider the following factors pertaining to the structure and organization of the course and the design and pedagogical principles important for their motivation to carry on in the course? If so, are they more likely to complete than those who consider them unimportant?

a. Curated and sequenced course structure,
b. Clear and salient learning objectives,
c. Gradual and scaffolded presentation of input,
d. Continuing storylines,
e. Form-focused and scaffolded presentation of grammar,
f. Variety in types of learning objects.

Chapters 1.1 and 1.2 underlined the need for new engagement strategies for learners in LMOOCs with the aim of enhancing retention (de Freitas et al., 2015; Hew, 2016; Hone & El Said, 2016; Kim et al., 2017). For that purpose, elements such as tutor support and supervision, including conditions for interactions between the tutor and learner, as well as proper introduction to course prerequisites and syllabuses, have been proposed as being beneficial factors for the autonomous online learner (El Said, 2017; Hew & Cheung, 2014; Höfler et al., 2017; Joo et al., 2018; Ross et al., 2014; Rubio et al., 2018; Sokolik, 2014). Empirical evidence, however, is needed on the potential impact of tutor-specific factors on learner engagement and retention in LMOOCs (Bárcena & Martín-Monje, 2014; Sokolik, 2014). Based on this background, the third objective and related research questions were as follows:

Objective 3

The third objective of the study was to discover whether learners believe that specific factors related to tutor support, provided in the blended and distance modes of IOL 2, are important for their motivation to engage with the course and whether the results reveal a possible impact of these tutor-specific factors on student retention. This was executed by

gathering and combining survey data and tracking data for analysis from 64 users who were in either of these modes from 2010 to 2018. Data collection and analysis that address the research questions in (C) below were performed with respect to the suggestions that strong teaching presence in MOOCs, which involves the direct interaction of tutors and students, and a focused introduction to course requirements and course syllabuses may be a critical component in the encouragement of student engagement with course content and retention (El Said, 2017; Hew, 2016; Hew & Cheung, 2014; Hone & El Said, 2016; Ross et al., 2014). Based on this background it is expected in the present study that the factors relating to tutor support and personal assistance for learners provided in the tutorial modes of IOL will have a positive impact on student engagement and retention. Related research questions were as follows:

(C) Do learners in IOL 2, in the blended and distance modes, consider the following tutor-specific factors important for their motivation to carry on in the course? If so, are they more likely to complete than those who consider them unimportant?

a. Private interaction with the tutor,

b. Overall tutor support during the course,

c. A detailed introduction of the program,

d. A set syllabus where course content is pre-organized into timetabled manageable sections.

Chapters 1.1 and 1.2 pointed out that student retention in MOOCs is commonly evaluated without considering what the learners' actual intentions were in terms of course engagement when they entered a course (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014). Existing research shows that learners join MOOCs with various goals in mind in terms of participation, and that those who have the goal of completing a course are more likely to complete than those who do not have such a goal (Henderikx et al., 2017; Reich, 2014). However, more knowledge is needed on student intent in LMOOCs, particularly in different modes of delivery, and the potential impact of the initial goal to complete the course on actual course completion. Based on this background, the fourth objective and relevant question were as follows:

Objective 4

The fourth objective of the study was to uncover whether learners in IOL 2 had the initial

goal of completing the course when they started, and to reveal the potential effect of this factor on student retention, with a special focus on learners' intent across different modes of delivery. This was carried out by gathering and combining survey data and tracking data for analysis from 400 users who were registered in the course from 2010 to 2018, including the three different modes of delivery. Data collection and analysis addressing the research questions in (D) below were in reference to indications that many MOOC learners sign up for such courses with the intention of completing a full course, while others have different intentions (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014). Based on the literature, it is expected that the diverse participants in the study came to the IOL course with different goals in mind, and that the factor of an initial goal to complete a course will have a positive impact on student retention in the present study. The research questions were:

(D) Do learners in IOL 2, both overall and within each mode, have the initial goal of completing the course? If so, are they more likely to complete than those who do not have such a goal?

As discussed in Chapters 1.1 and 1.2, MOOC learners represent diverse users with various motives which need to be taken into account in the discussion on student retention (Beaven et al., 2014; de Barba et al., 2016; Ingolfsdottir, 2014; Kizilcec & Schneider, 2015; Ryan & Deci, 2000; Salmon et al., 2017), calling for more studies on how different engagement patterns in online environments reflect learners' motivations (Chen et al., 2020; Doiz et al., 2014; Durksen et al., 2016; Sokolik et al., 2014; Wang & Baker, 2015). More evidence is thus needed on how various motives may promote student engagement and drive them towards course completion (de Barba et al., 2016; Ingolfsdottir, 2014; Kizilcec & Schneider, 2015). In this context, greater understanding is also needed as to why learners who had the initial goal of completing such a course, disengage before completing (El Said, 2017; Reich, 2014). Founded on this background, the fifth objective and related research questions were as follows:

Objective 5

The fifth objective of the study was to reveal learners' own views as to what they see as crucial factors that either contributed to their completion of IOL 2 or prevented them from completing the course. This was performed by gathering qualitative data through open question sections in a survey. A total of 112 informants who were enrolled in the course

from 2010 to 2018 provided data on the reason why they had completed the course, while 62 informants offered data on the reason why they left before completion. Data elicitation and analysis regarding the research questions in (E) below were in reference to findings that course completion may reflect various types of motivation (Beaven et al., 2014; Chen et al., 2020; de Barba et al., 2016; Doiz et al., 2014; Kizilcec & Schneider, 2015; Littlejohn et al., 2016; Ryan & Deci, 2000; Wang & Baker, 2015), and that attrition in such courses may be attributed to factors outside the learning context itself (Belanger & Thornton, 2013; de Freitas et al., 2015; Henderikx et al., 2017, 2018, 2019; Perna et al., 2014; Reich, 2014; Shapiro et al., 2017). Based on previous research and the highly diverse target group in MOOCs, it is expected that the present study will reveal various motives for course completion, as well as different reasons for student attrition. The guiding research questions were:

(E) 1. Why do learners in IOL 2 complete the course? What is the motive from their point-of-view?

2. Why do learners who intend to complete IOL 2 not complete it? What is their primary reason for leaving?

As mentioned in Chapter 1.2, the effect of demographic variables such as age and gender have been discussed in relation to user attitudes towards the acceptance and uptake of modern information technologies (Oshlyansky, Cairns, & Thimbleby, 2007), where findings show that older learners seem to be more anxious towards such learning settings in higher education (Khechine et al., 2014; Wang et al., 2009). Furthermore, previous findings indicate that gender does not have a significant effect in terms of technology acceptance in this context (Khechine et al., 2014). From this perspective and based on the analysis of registration data that revealed a relatively high mean age of learners in IOL 2, it is expected that age differences in terms of retention will be identified in the different groups of learners in the study; the older learners will be less likely to complete the course than the younger learners. Similarly, based on the existing literature (Khechine et al., 2014) it is also expected that no relation will be identified between age and retention in the study. From this perspective, the sixth objective and the related question were as follows:

Objective 6

The sixth objective of the study was to reveal whether the age and gender of the users in IOL 2 influence student retention. This was executed by collecting registration data and

retention data for analysis through IOL's tracking system on 400 users who were in IOL 2 from 2010 to 2018. The following question was addressed:

(F) Do demographic factors such as gender and age have a predictive value in relation to student retention?

1.5 Summary

The IOL program, which was the primary source of evidence for this study, includes seven open online and curated courses for second language learners, including three different modes of delivery in two of the courses. IOL has a built-in tracker which records data on learners' progression and thus provides information on the users and usage of the program. The study was guided by six fundamental objectives with the primary focus on understanding what factors have a substantial impact on retention in such learning contexts, whether they are attributed to IOL's instructional design and pedagogical methodology, specific mode-related factors including tutor support, or even other factors ascribed to learners' individual motives or circumstances. First, the study attempted to reveal overall student retention and user progress in the seven IOL courses, including the three different modes of delivery, and thus explore what the engagement patterns in IOL may suggest about student retention. Second, the study made an effort to investigate user experience concerning six content-specific factors regarding the structure and organization of the IOL 2 course, as well as the instructional design and pedagogical principles, and the potential impact on engagement and retention. Third, the study attempted to explore the same user experiences of four mode-specific factors related to tutor support in the blended and distance modes of the course, and the possible effect on engagement and retention. In addition, the study addressed the question whether learners in IOL 2 had the initial intent of completing the course once they started, and the possible effect of this factor on student retention. Furthermore, to gather additional understanding of the topic, the study asked learners who completed IOL 2 to reflect on the reason(s) why they completed the course on the one hand, and asked non-completers why they disengaged before completing on the other. Finally, the study addressed the issue whether demographics, that is, age and gender, influence student retention in IOL 2. The following section describes the research methods and methodology used in the study.

2. CONDUCTING A MIXED METHODS STUDY

This is a mixed methods study, which addressed the overall question of determinants of student retention in online language learning environments. The study used multiple sources of evidence; namely, tracking data, survey data, and qualitative data obtained through a survey. The methodology used has been described as "research that involves collecting, analyzing, and interpreting quantitative and qualitative data in a single study or in a series of studies that investigate the same underlying phenomenon" (Leech & Onwuegbuzie, 2009:265). The rationale for collecting both quantitative and qualitative data in the research project is that each dataset on its own is insufficient to explain the phenomenon, and multiple perspectives are required to acquire a more complete understanding (Creswell, 2006; Leech & Onwuegbuzie, 2009) of student retention in online language learning courses. It is the type of design where quantitative and qualitative data are gathered, analyzed, and interpreted either independently or in parallel and are partially mixed at certain points in the research process (Guest, 2012). The research is empirical with an inductive approach which is aimed at generating a new theory emerging from the data (Godwin-Jones, 2017; Long & Siemens, 2011).

First, using the computer-based tracking system in IOL, the study collected data on students' usage and behavior in the program (Fischer, 2007). Data mining and LA were applied in the initial part of the study to analyze the large set of tracking data in the attempt to reveal how learners progressed through the IOL courses and to identify any patterns emerging from the data. Massive data collection and the application of LA holds the potential not only to use these intelligent learner-produced data to visualize and analyze learners' online interactions in context, but also to improve instructional materials and approaches and to benefit research methods and outcomes that inform both second language acquisition theory and CALL practices (Godwin-Jones, 2017; Long and Siemens, 2011; Martín-Monje et al., 2018; Thomas & Gelan, 2018). Data mining is increasingly used in CALL research, in which the investigation of such datasets creates opportunities to track and measure the data with unique accuracy in order to address practical questions about the usage and impact of CALL programs (Fischer, 2007; Godwin-Jones, 2017).

Second, the survey study used two research methods, a survey instrument and a written questionnaire to collect and analyze self-reported data from individuals in order to inform the study about users' opinions and experiences of various course-specific elements

in focus (Dörnyei & Csizér, 2012). Survey methodology and questionnaire design have their origins in the social sciences and are commonly used in the context of SLA research (Dörnyei & Csizér, 2012). In addition, a tracking system was used in the study for the collection and analysis of retention data from the same individuals in the interest of measuring the potential impact of the specific elements on student retention (Chun, 2013; Fischer, 2007).

Furthermore, the study was developed within the qualitative tradition of inquiry where the philosophical and theoretical underpinnings of the study are rooted in ontology, with the aim of generating further insight into the issues being studied (Creswell, 2013; Friedman, 2012; Godwin-Jones, 2017; Hsieh & Shannon, 2005). In order to gather non-numerical data, the study used an open-ended questionnaire in the survey where narratives from participants were provided in writing in order to examine the learners' own reflections on the topic. This methodology has the advantage of allowing participants more time to formulate responses and has the potential of eliciting more detailed descriptive data from the informants than closed-ended questionnaires (Creswell, 2013; Friedman, 2012). In the present study, which is based mainly on methods involving large sets of tracking data and survey data, a qualitative method was thus considered necessary supplement to the quantitative data (Godwin-Jones, 2017) in order to obtain a more complete picture of student retention and engagement in LMOOCs.

The sources and the research procedure used to investigate the subject were as follows:

a) A large set of tracking data, from 43,468 learners, was collected through a computer-based tracking system and subsequently analyzed. The purpose of this set of data was to unveil overall student retention and engagement patterns in IOL and to use the results as a premise for subsequent data sampling, collection and analysis procedures.

b) Based on the findings in the first phase of the study (a) above), one set of survey data and one set of tracking data involving a focused identical sample of 400 learners were collected and analyzed independently. Then, these two datasets were brought together, merged, and analyzed. These datasets were intended to reveal in more detail the participants' experiences of certain factors pertaining to the structure and content of IOL, and tutorial features, along with other motivational factors, and their potential influence on retention.

c) Two sets of qualitative data were elicited simultaneously through open question forms in the survey, involving a total of 174 informants, and then analyzed separately in the attempt to further illuminate the survey and tracking data results, and thus to provide a broader perspective on critical factors in retention.

d) The results of each stage of the research project were then assembled for overall interpretation which are presented here. The following figure describes the research design.

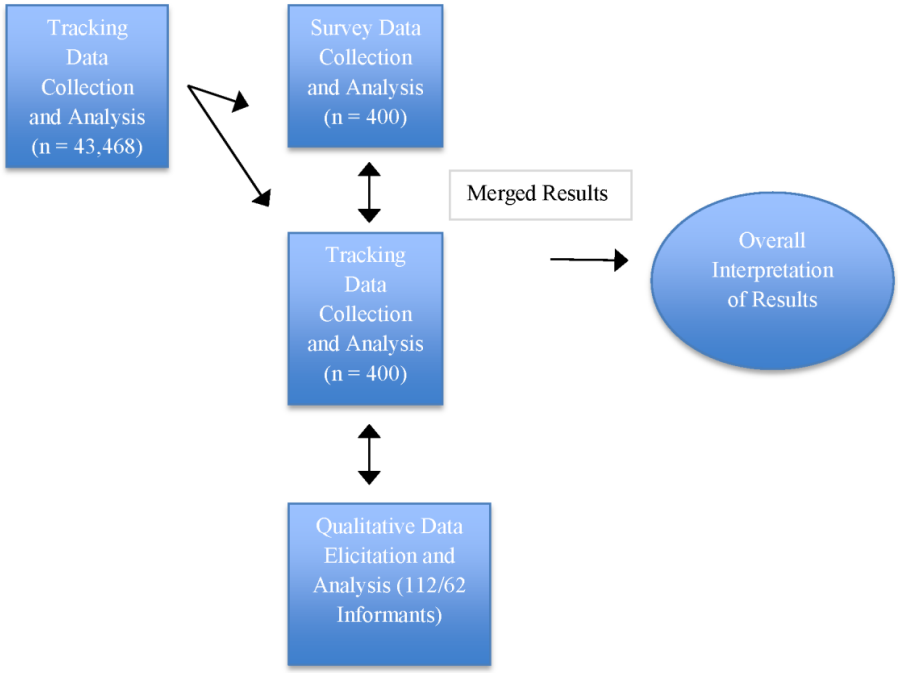


Figure 1. The study's mixed methods design based on points of interface between datasets and timing of mixing (Guest, 2012).

The subsequent chapters outline the methods used for data collection and analysis in each phase of the study. First, Chapter 2.1 describes the collection and analysis of the tracking data on learners in all the seven IOL courses and the three different modes of delivery of IOL 1 and 2, gathered through IOL's tracking system. Then, Chapter 2.2 attends to the survey data in relation to the tracking data on learners in IOL 2. Finally, Chapter 2.3 considers the qualitative data on two groups of learners in IOL 2.

2.1 Tracking Data

This section describes the methods used for data collection and analysis of the tracking data on learners' behavior in IOL overall and the three different delivery modes in the first part of the study (Article I). The research questions were as follows:

1. What is the overall student retention in the IOL courses?
2. Does the mode of delivery affect student retention. If so, how?
3. What does the overall engagement pattern in IOL suggest about retention?

2.1.1 Participants and data collection

The participant sample included 43,468 online learners out of the 139,941 who registered for one or more of the seven IOL courses from 2006 to 2014. Participants were distributed across the Survival course (n = 38,167), IOL 1 – Nature theme (n = 11,519), IOL 1 – Culture theme (n = 6,954), IOL 2 (n = 3,805), IOL 3 (n = 2,386), IOL 4 (n = 2,417), and IOL 5 (n = 1,185). Some of the participants in the study were working on one or more of the courses, or even all of them over the study period. Since IOL 1 (Culture theme) and IOL 2 are delivered in three different modes of delivery, participants in IOL 1 (n = 6,954) studied in either the open self-directed mode (n = 6,419), the blended mode (n = 434), or the distance mode (n = 101). Similarly, participants (n = 3,805) in IOL 2 overall were either in the open self-directed mode (n = 3,462), the blended mode (n = 281), or the distance mode (n = 62). The population in this part of the study was scattered around the world, with 63% coming from ten countries: the US, Germany, Poland, the UK, France, Canada, Sweden, Norway, Spain, and Italy. The division by gender was almost equal, with 51% of learners female and 49% male. In addition, 69% were under the age of 31 and 54% had a university degree.

Data were gathered through the tracking system, which monitors each participant's progress in all the courses. As discussed in Chapter 1.3.3.3, the tracking system captures each user's access logs to each content page of a course and stores the relevant retention data along with the registration data in a database that holds large amounts of data on the user's online behavior. The tracking system does not separate registration data and user data. When a user enrolls in IOL, he or she is automatically positioned in the database as a user on the first content page (reference number 111 in the database) on all the seven courses, regardless of which course the user may then be working on, and whether he or

she does anything more than register. Due to this element of uncertainty, content page 111 was excluded from the data analysis, involving 96,473 registrants, for all the courses and different modes of delivery in the study.

2.1.2 Tracking data analysis

For tabulating student retention, all data in the database were transferred to an SPSS statistics program (IBM Corp.) for analysis. As already noted, only data from users (n = 43,468) who went beyond content page 111 in the courses were included in the analysis. Tracking data for analysis were then guided by the three research questions.

1. To answer the first research question in Chapter 2.1, regarding overall student retention for the IOL courses, all the users who persisted to the end of the seven courses were recognized and the percentage of those who accessed the last content page in each course was calculated.

2. To answer the second question, whether the mode of delivery affects student retention, tracking data for analysis were selected by identifying users in IOL 1 (Culture theme) and IOL 2 who remained to the end of each course. The percentage of those who accessed the last content page was then calculated on a) the open self-directed courses, b) the blended courses, and c) the distance courses.

3. In order to answer the third research question of what overall engagement patterns in IOL suggest about retention, the data for analysis were selected by using two methods:

First, data were compiled on learners who started but did not complete the IOL 1 (Culture theme) and IOL 2 courses, for the three different modes of delivery. The overall attrition pattern was analyzed separately for IOL 1 and IOL 2 in the three modes of delivery, by calculating the rate of attrition on each content page in each course across all modes.

Then, data were compiled to reveal the overall engagement patterns based on diverse parameters for coverage of course content. Thus, parameters for course completion were adjusted to learners covering a) less than 50% of course content, b) from 50% to 74% of course content, c) from 75% to 89% of course content, and d) from 90% to 99% of course content.

The following section turns to the survey study and focuses on the methods used to gather and analyze survey data on the one hand, and retention data on the other.

2.2 Survey Data in Relation to Tracking Data

This section discusses the methods used for data collection and the analysis of the survey data in relation to the tracking data (Articles II and III). This part of the study addressed learners' experiences regarding specific course-related factors along with motivational factors attributed to learners themselves, and the potential impact of these factors on student retention in one course, IOL 2. In addition, this section outlines the methods used to explore the potential effects of two demographic factors, age and gender, on student retention. The demographic factors were explored in the process of overall interpretation of the research results.

The section starts with an introduction of the research questions that are addressed in the survey study. These concern a) the content-specific factors in the course (A below), b) the tutor-specific factors provided in the course (B below), and c) the factor that concerns learners' initial goals in terms of course engagement (C below). In addition, this section addresses the selected demographic factors and the relevant research question (D below). Chapter 2.2.1 presents the survey study participants. Then the survey instrument is outlined in Chapter 2.2.2. Finally, Chapter 2.2.3 involves the methods used for the survey and tracking data analysis. The following research questions were addressed:

A. Content-specific factors:

Do learners in IOL 2 consider the following factors pertaining to the structure and organization of the course and the design and pedagogical principles important for their motivation to carry on in the course? If so, are they more likely to complete than those who consider them unimportant?

a. Curated and sequenced course structure,

b. Clear and salient learning objectives,

c. Gradual and scaffolding presentation of input,

d. Continuing storylines,

e. Form-focused and scaffolded presentation of grammar,

f. Variety in types of learning objects.

B. Tutor-specific factors:

Do learners in IOL 2, in the blended and distance modes, consider the following tutor-specific factors important for their motivation to carry on in the course? If so, are they more likely to complete than those who consider them unimportant?

a. Private interaction with the tutor,

b. Overall tutor support during the course,

c. A detailed introduction of the program,

d. A set syllabus where course content is pre-organized into timetabled, manageable sections.

C. Initial-goal factor:

Do learners in IOL 2, both overall and within each mode, have the initial goal of completing the course? If so, are they more likely to complete than those who do not have such a goal?

D. Demographic factors:

Do the demographic factors of age and gender have a predictive value in relation to student retention?

Administration and sampling procedures for the study included a collection of the survey data and tracking data which were both aimed at the same specified target group, which is discussed below. The two datasets were analyzed separately and then combined for further analysis. The data for the analysis on participants' age and gender were collected through the course registration form and merged with the tracking data.

2.2.1 Participants

The specified target group for the survey study was derived from the findings revealed in the first part of the research project (Chapter 3.1.3) and included learners (n = 2,605) in IOL 2 from 2010 to 2018 and who had covered 15.9% to 100% of the course content. Of those, a total of 400 learners completed the survey, yielding a 15.4% response rate. The study population was diverse, with 63% originating from twelve countries: Germany, the US, the Philippines, the UK, France, Canada, Poland, Denmark, Sweden, Russia, Italy, and Switzerland. Of this population, 60% were female and 40% were male, and the age range varied from 16 to 87 years, with a mean age of 39. Moreover, 74% have a university degree.

2.2.2 Survey design

An anonymous post-course online questionnaire (Appendix F) was developed for the study, with nineteen items mainly using a five-point Likert scale (1 = 'strongly agree' to 5 = 'strongly disagree', 1 = 'definitely' to 5 = 'definitely not', or 1 = 'very easy' to 5 = 'very difficult'). Three items used a four-point Likert scale (1 = 'frequently' to 4 = 'never') and two items used a three-point Likert scale (1 = 'very important' to 3 = 'not important').[viii] A self-report measure of learner navigation, resource use, tutor contact, initial goal, and retention (the total of 5 items) was also created and included in the questionnare (Hone & El Said, 2016). The questionnaire was made up of closed-ended questions. In addition, the survey included open-ended sections involving two questions in order to elicit qualitative data (Chapter 3.3).

Based on the exploratory nature of the research project, the survey was extensive and designed to gather data for the two follow-up studies (Articles II and III) and for future investigations. Items in the survey were primarily constructed in view of IOL's specific design factors. It was also constructed in light of previous conclusions drawn on the engagement patterns identified in IOL 2 (Article I), and of a review of past literature. The construction of the questionnaire was additionally based on informal evidence (Colpaert, 2010; Doiz et al., 2014) extracted from learners' comments on specific elements of the IOL course, complaints and testimonies gathered via email or face-to-face, along with informal surveys and pilot interviews (Friðriksdóttir, 2015). In addition, alpha and beta tests during software and content development phases of the program were also informative to the survey. Furthermore, the content of the survey was driven by established findings and indications in the literature on potential determinants of retention in online language learning environments.

The construction of the survey represents three components based on research questions (A), (B) and (C) in Chapter 2.2 above: 1) The content-specific factors; 2) the tutor-specific factors; and 3) the factor assigned to learners' initial goals in terms of course engagement. Each component includes two to eleven relevant items under investigation in the study. The full scales can be found in Appendix C and are referred to in the following summary:

1. Content-specific factors refer to certain pedagogical methods applied in IOL 2 that were presented in Chapter 1.3.3.1, and that are intended to motivate and engage learners with the course content. First, these factors concern the structure and organization of the course

content (a) and b) below), where six relevant items were measured, SO1–6. Second, they concern the instructional strategies applied in the course (c) to f) below), where eleven relevant items were measured, IS1–11. First, the factors regarding the structure and organization of the course content are as follows:

a) Curated and sequenced course structure (labeled SO1–3 in Appendix C),

b) Clear and salient learning objectives (labeled SO4–6),

Then, the factors concerning the instructional strategies are as follows:

c) Gradual and scaffolding presentation of input (labeled IS1),

d) Continuing storylines (labeled IS2–3),

e) Form-focused and scaffolded presentation of grammar (labeled IS4–6),

f) Variety in types of learning objects (labeled IS7–11).

As argued in Chapter 1.4, the theory which was put forward implies that LMOOC learners consider the content factors studied important in allowing them to engage with the course content, and that these factors are likely to encourage retention (Arnbjörnsdóttir, 2004; Castrillo, 2014; Chun, 2012, 2016; Colpaert, 2014; Dörnyei et al., 2014; Garrett, 1991; Godwin-Jones, 2017; Hubbard, 2013; Rosenshine & Meister, 1992; Sokolik, 2014).

2. Tutor-specific factors encompass elements that relate to the delivery mode of the course presented in Chapter 1.3.3.2 that involve support and supervision from a tutor, which is provided in the distance and blended modes. Five relevant items were measured, TI1–5. The factors concerning tutor interventions are as follows:

a) Private interaction with the tutor (labeled TI1–2 in Appendix C),

b) Overall tutor support (labeled TI3),

c) A detailed introduction of the program (labeled TI4),

d) A set syllabus (labeled TI5).

As claimed in Chapter 1.4, the theory which is proposed indicates that LMOOC learners consider the tutored factors studied important in enabling them to engage with the course and, that these factors are likely to support retention (Bárcena & Martín-Monje, 2014; El Said, 2017; Hew & Cheung, 2014; Höfler et al., 2017; Ross et al., 2014; Sokolik, 2014).

3. The initial goal factor refers to the learners' intentions in terms of course engagement once they started the course. Two relevant items were measured, IG1–2.

The theory argued in Chapter 1.4 implies that the diverse LMOOC learners come to such courses with different goals in terms of engagement (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014), and that those who have the initial goal of completing a course are more likely to do so than those who do not have this goal (Reich, 2014).

4. The factors of age and gender of the learners in IOL 2 were added in the process of interpreting results in order to evaluate whether such demographic factors may interact with student retention. The theory argued in Chapter 1.4 implies that age differences exist in technical learning environments and that gender does not have a significant effect in terms of technology acceptance (Khechine et al., 2014). Hence, these factors are considered in relation to student retention in the study.

To measure the students' views on the importance of the content-specific factors in terms of retention, participants in IOL 2 were asked to rate how important each factor was to them. To measure the students' experiences of the tutor-specific factors in terms of retention, participants in the blended and distance modes of IOL 2 were asked to rate how important each factor was to them. To measure the students' initial goals in terms of course engagement, participants in IOL 2 were asked to state whether they initially had the goal of completing the course, of completing part of it, or had no clear goal. To measure plausible age and gender differences in IOL 2 with regards to student retention, participant demographics were explored based on registration data. After piloting the questionnaire, it was distributed to learners who had completed more than 15.9% of the course content in IOL 2 (Chapter 3.1.3) with the web survey tool Qualtrics (Qualtrics Labs, 2012). The survey was completed in English and displayed several images from the IOL 2 course to assure that learners related to the questions.

The analysis of the data included retention as the dependent variable. The content-specific and mode-specific factors, and the initial goal factor, were included as independent variables. Furthermore, the study also included demographic variables of age and gender as independent variables.

2.2.3 Survey and tracking data analysis

Based on research questions (B), (C) and (D) presented in Chapter 1.4, the data for analysis were based on a) whether or not students experienced the content-specific factors in IOL 2 as important for their motivation and whether these factors affected retention, b) whether or not students in the blended and the distance modes experienced the tutor-specific factors important for their motivation and whether these factors affected retention, and c) whether or not students had the goal of completing the course when they started and whether the goal to complete affected student retention.

1. The tracking data were first explored to measure the level at which each student in IOL 2 was last tracked in the course. From these empirical data, a sample group was created containing students who had completed 15.9% to 100% (Chapter 3.1.3) of the course content. The sample was then uploaded into the web survey tool Qualtrics.

2. Survey data for analysis were then found by identifying respondents with respect to a) the content-specific factors and related items measured (SO1, SO2, SO3, SO4, SO5, SO6, IS1, IS2, IS3, IS4, IS5, IS6, IS7, IS8, IS9, IS10, IS11), b) the tutor-specific factors and related items measured (TI1, TI2, TI3, TI4, TI5), and c) the initial goal factor (IG1, IG2) (Chapter 2.2.2 and Appendix C). The data collected from the survey were then exported into SPSS Statistics (IBM Corp.) and R software package (R Core Team, 2020) was used for the analysis.

3. Finally, the survey data were merged with the retention data in Excel.

Moreover, based on research question (F) presented in Chapter 1.4 regarding the factors of age and gender, the data for analysis were found by identifying respondents with respect to their registration data and the retention data, to which a linear regression was applied. The following section discusses the data elicitation and analysis for the qualitative part of the study.

2.3 Qualitative Data

This section reviews the method which was developed in the qualitative tradition of inquiry (Articles II and III), with the aim of expanding understanding and pursuing a broader view of other motivational or external factors that may drive learners to complete a course or hinder them from completing. Based on research question (E) presented in Chapter 1.4, data for analysis were elicited through open question sections in the survey.

The survey instrument was described in Chapter 2.2, including a description of the target group (Chapter 2.2.1), which is mentioned here since the data gathering and analysis in Chapters 2.3.1 and 2.3.2 below rely on the same sample. The data elicitation and analysis focused on the following research questions:

1. Why do learners in IOL 2 complete the course? What is the motive from their point-of-view?

2. Why do learners who intend to complete IOL 2 not complete it? What is their primary reason for leaving?

2.3.1 Data elicitation

The qualitative data were elicited through the survey's open-ended question forms where different target groups were invited to reflect on the reasons why they completed, or did not complete, IOL 2. These groups had reported in the survey whether they had completed the course or not[ix], and therefore received different questions. First, for the purpose of eliciting data about the reason(s) why students completed the course, participants who had self-reported course completion in the survey were invited to express their views in the survey's free writing form (Creswell, 2013; Doiz et al., 2014; Friedman, 2012) by using the fill-in sentence (Colpaert, 2010): "When I think about the reason why I completed the course to the end I think about the following ..." (Appendix F).

Similarly, in order to gather data about the reasons why committed learners left the course before completing, participants who had stated the initial goal of completing IOL 2 and who had self-reported non-completion of the course in the survey were offered to express their perspectives in the survey's open question form by using the fill-in sentence (Colpaert, 2010): "If it turned out that you did not complete the course after all, please write three keywords to describe why you think your initial goal changed." (Appendix F). Based on the study's research questions, the text data from these informants were then examined for common patterns, or themes (Hsieh & Shannon, 2005).

2.3.2 Data analysis

To address the first research question in Chapter 2.3 above on why learners in IOL 2 completed the course, the text data for analysis were based on users who had self-reported course completion in the survey (Appendix C) and were found by identifying responses from these users in a) the IOL 2 course overall, and b) the blended, distance and open self-directed modes. Sample groups were then created.

Then, to address the second research question in Chapter 2.3 about the reason why committed learners in IOL 2 did not complete the course as was their goal, the data for analysis were based on users who had reported that they had the initial intent of completing the course and had stated non-completion of the course in the survey (Appendix C), and were selected by recognizing responses from these users. A sample group was then created.

The study was grounded on analytic induction (Creswell, 2013; Taylor & Bogdan, 1998) in order to discover meaningful patterns or themes across the data samples under investigation (El Said, 2017; Hsieh & Shannon, 2005; Namey, Greg, Thairu, & Johnson, 2008). A summative approach was used to interpret meaning from the content of the text data with attention to expressions of engaging or disengaging elements or other affective viewpoints (Duff, 2012; Hsieh & Shannon, 2005) in terms of retention.[x] The survey instrument that was used to elicit these data gave the researcher the opportunity to export reports from the informants to Microsoft Word for analysis. The study employed initial coding (or open coding) (Creswell, 2013), whereby the data were coded manually based on labeling relevant words or repeating content that emerged from the data, followed by an interpretation of the contextual meaning of the content (Hsieh & Shannon, 2005). Finally, to extract content categories from the data, relevant codes were connected into overarching themes until the new data reached saturation (Creswell, 2013; Taylor & Bogdan, 1998). Themes that occurred most often across the datasets were in the focus of this study.

2.4 Summary

The study used mixed methods to address the six underlying objectives of the research project and comprised three different data sources: tracking data, survey data in relation to the tracking data, and qualitative data elicited through the survey. First, in order to answer the research questions in (A) in Chapter 1.4 on the overall retention and engagement patterns in IOL, tracking data from approximately 43,000 learners were collected through IOL's tracking system. The data were explored with the use of LA in order to uncover the overall retention in the seven IOL courses, and to identify engagement patterns in the courses, including the three different modes of delivery in IOL 1 and 2. The purpose of this set of data was also to provide the premise for sampling and data collection and analysis for the subsequent studies.

Second, to answer the research questions in (B) and (C) in Chapter 1.4 on the influence of the course-specific factors in IOL 2, survey data and tracking data were collected from 400 learners in order to reveal the participants' experiences of six content-specific and four mode-specific factors in the course, namely, whether or not they considered them to be important for their motivation to persist in the course. The tracking data from the same learners were then investigated and merged with the survey data in order to explore the potential impact of these factors on student retention. The same procedure was followed in order to answer research question (D) in Chapter 1.4 on whether or not learners in IOL 2 came to the course with the intentions to complete it and the potential impact of this factor on retention.

Third, to answer the research questions in (E) in Chapter 1.4 on the reasons why learners in IOL 2 completed the course or disengaged earlier, two sets of qualitative data were gathered through open question sections in the survey. Participants who had self-reported course completion of IOL 2 in the survey were asked to reflect on the reason why they completed the course. A total of 112 informants provided data on this question. On the other hand, participants who had stated non-completion of the course in the survey and who had also reported that they had had the initial goal of completing the course, were asked to consider why they ended up not completing the course. The study received 62 comments on this matter. These data were then analyzed for common themes.

Finally, to answer the research question in (F) in Chapter 1.4 on the potential impact of the demographic factors age and gender on student retention in IOL 2, registration data from 400 users were collected through IOL's tracker and measured against same learners' tracked retention data. The following section presents the results of the study.

3. RESULTS AND DISCUSSION

As outlined in the previous chapters, the research was based on mixed methods and comprised three different data sources: a) tracking data, b) survey data in relation to the tracking data, and c) qualitative data gathered through open questions in the survey. Therefore, this chapter is organized around the three datasets with reference to the six objectives that were introduced in Chapter 1.4, and includes three main chapters. The first main chapter, 3.1, presents the results of the tracking data on the overall retention and engagement patterns in IOL. The second main chapter, 3.2, introduces the results from the survey data on learners' experiences in IOL 2, as well as their goal in relation to the relevant tracking data. It also presents the results regarding the demographics in relation to the tracking data. The third main chapter, 3.3, concentrates on the findings from the qualitative data gathered through the survey. The organization is as follows.

First, Chapter 3.1, which focuses on the tracking data results (Article I), is divided into three sub-chapters: Chapter 3.1.1 highlights the findings on student retention across all seven IOL courses and the three different modes of delivery of IOL 1 and 2. Chapter 3.1.2 then discusses the overall engagement patterns in the IOL courses, including the attrition patterns revealed across the different delivery modes of IOL 1 and 2. Finally, Chapter 3.1.3 includes a summary and discusses the premises set for the subsequent survey studies covered in Chapters 3.2 and 3.3.

The second main chapter, Chapter 3.2, which focuses on the survey data results in relation to the retention data results, is split into five subchapters: Chapter 3.2.1 reveals learners' views about their experiences using content-specific factors in IOL 2 and their effects on retention (Article III), while Chapter 3.2.2 uncovers learners' views on tutor-specific factors in the course and their effects on retention (Article II). Chapter 3.2.3 centers around the learners' stated goal of course engagement when they started the IOL 2 course and the effect of the initial goal factor on retention (Article II). In Chapter 3.2.4, the demographic factors and their effect on retention are revealed. Chapter 3.2.5 involves a summary.

Finally, Chapter 3.3 concentrates on the results of the analysis of the qualitative data that were elicited through the survey's open questions, and reveals informants' views on what either drove them towards the end of the IOL 2 course or prevented them from completing it. That chapter is divided into three sub-chapters. Chapter 3.3.1 reveals

learners' reflections on the reasons why they completed the course (Article II), while Chapter 3.3.2 uncovers the learners' thoughts on why they did not finish the course as intended (Article III). Chapter 3.3.3 then concludes with a summary.

3.1 Tracking Data: Overall Retention and Engagement Patterns in IOL

The main objective of the first part of the study (Article I) was to uncover student retention and engagement patterns in the IOL program as measured by the integrated tracking system, both overall across the seven courses and within the three different modes of delivery in two of the courses. Based on objective 1 in Chapter 1.4, the data collection and analysis centered around three research questions:

1. What is the overall student retention in the IOL courses?
2. Does the mode of delivery affect student retention. If so, how?
3. What does the overall engagement pattern in IOL suggest about retention?

The data source comes from roughly 43,000 users who signed up for one or more of the courses from 2006 to 2014, of which approximately 11,000 were participants in the courses provided in the three different modes of delivery in IOL 1 and 2. The discussion in the following subchapters is organized around the three research questions. The results responding to research questions one and two above are discussed in Chapter 3.1.1 and the results connected with the third research question are presented in Chapter 3.1.2.

3.1.1 Retention across all courses and different modes of delivery

When addressing the first research question above on the overall retention for the IOL courses, the investigation of the tracking data on the learners (n = 43,468) in all seven IOL courses, revealed overall low completion rates across the seven courses, as shown in Table 1.

Table 1. Completion rates across all seven courses in IOL as measured by the tracking system.

	IOL Survival (n = 38,167)	IOL 1 Nature (n = 11,519)	IOL 1 Culture (n = 6,954)	IOL 2 (n = 3,805)	IOL 3 (n = 2,386)	IOL 4 (n = 2,417)	IOL 5 (n = 1,185)
Completion rates %	5.3% (n = 2,016)	2.4% (n = 281)	3.3% (n = 229)	5.1% (n = 195)	4.9% (n = 117)	18.2% (n = 440)	3.5% (n = 41)

Note that completion rates in IOL 1, Culture theme, and IOL 2 are based on an average of the three different modes of delivery in these courses. Note also that some of the users participated in more than one course during the period of the investigation, which explains why the total number of participants in the table is higher than 43,468.

Table 1 reveals that the completion rates are relatively low in all the IOL courses, ranging from 2.4% to 18.2%. The Survival IOL course exposes 5.3% completion rates (where 2,016 out of the 38,167 who started the course completed it); IOL 1, Nature theme, shows 2.4% completion rates (where 281 out of the 11,519 who started the course completed it) and the Culture theme 3.3% completion rates (where 229 out of the 6,954 who started the course completed it); IOL 2 shows 5.1% completion rates (where 195 out of the 3,805 who started the course completed it); IOL 3 shows 4.9% completion rates (where 117 out of the 2,386 who started the course completed it); IOL 4 exposes 18.2% completion rates (where 440 out of the 2,417 who started the course completed it), and IOL 5 shows 3.5% completion rates (where 41 out of the 1,185 who started the course completed it). Why retention rates were highest in the IOL 4 course is yet to be explored. However, it is possible that the learners in this high-intermediate level course have already put a lot of effort and time into learning the language by supposedly completing IOL 1, 2, and 3 before attempting IOL 4 and may, therefore, be more likely to proceed (Chen et al., 2020; Ihantola et al., 2020). It should also be mentioned in that regard that the IOL 4 course contains the lowest number of content pages compared to the other IOL courses (Appendix A). This, however, does not necessarily mean that the IOL 4 course is less demanding for learners than other courses with higher numbers of content pages. The low completion rates in IOL 5, on the other hand, might be because this course is designed for advanced learners at the end of their BA studies at UI, where learners are not obliged to complete the whole course but can instead choose to work on specific sections.

To summarize the answer to the first research question in Chapter 3.1 regarding the overall retention in the IOL courses, the data show, as expected, that student retention in the seven courses is relatively low, and therefore confirm established findings on low completion rates in MOOCs (Chen et al., 2020; Ingolfsdottir, 2014; Jordan, 2014, 2015; Koller et al., 2013; Reich, 2014).

In order to answer the second research question in Chapter 3.1 regarding the potential impact of the mode of delivery on student retention, the tracking data, involving the total of 10,759 learners, as shown in Table 2, were examined specifically for the efficacy of different modes of delivery of IOL 1 (Culture theme), including 6,954 learners, and IOL 2, including 3,805 learners.

Table 2. Completion rates across three different delivery modes of IOL 1 and 2 as measured by the tracking system.

	IOL 1 Open self-directed mode (n = 6,419)	IOL 1 Distance mode (n = 101)	IOL 1 Blended mode (n = 434)	IOL 2 Open self-directed mode (n = 3,462)	IOL 2 Distance mode (n = 62)	IOL 2 Blended mode (n = 281)
Completion rates %	2.9% (n = 189)	4.0% (n = 4)	8.3% (n = 36)	4.4% (n = 152)	4.8% (n = 3)	14.2% (n = 40)

As Table 2 shows, completion rates ranged from 2.9% to 8.3% across modes of the IOL 1 course, and from 4.4% to 14.2% in the IOL 2 course. Focusing first on IOL 1 on the left-hand side of the table, the data analysis showed that the completion rate was highest in the blended mode of IOL 1, or 8.3% (where 36 out of the 434 who began the course completed it), followed by the distance mode with a 4.0% completion rate (where 4 out of the 101 who started the course completed it) and the open self-directed mode with a 2.9% completion rate (where 189 out of the 6,419 who started the course completed it). Correspondingly, when observing the results for IOL 2, the highest completion rates were found in the blended mode, or 14.2% (where 40 out of the 281 who began the course completed it), as compared to a 4.8% completion rate in the distance mode (where 3 out of the 62 who started the course completed it) and a rate of 4.4% in the open self-directed mode (where 152 out of the 3,462 who began the course completed it).

Hence, the findings in Table 2 show that completion rates varied between the three different modes of delivery in both IOL 1 and 2, and that the blended learning mode of both courses was more effective in keeping learners engaged to the end in comparison to the other two modes of the same course. Learners in the blended learning mode were found to have significantly higher completion rates than learners in the open self-directed mode of both IOL 1 ($p = <0.001$) and IOL 2 ($p = <0.001$). Because of the limited sample size of the distance learners completing both courses, it was not possible to run those through an analysis of statistical significance. This result therefore supports the hypothesis presented

in Chapter 1.4 that the blended learning mode would be more effective in retaining learners than the other modes studied. Much of the literature echoes the findings that student retention varies according to mode of delivery, and that a blended learning mode is more efficient in retaining students than other delivery modes (Garrison & Kanuka, 2004; Harker & Koutsantoni, 2005). The findings revealed in this study, however, raise many questions as to why the blended mode is more effective than the other modes, and call for a follow-up study on the potential impact of tutor-related factors on retention in the distance and blended modes of IOL 2. We will come back to this issue in Chapter 3.2.2. When considering other factors that could explain the differences in retention rates between the three modes of delivery in IOL 1 and 2 shown in Table 2, it has been argued (Jordan, 2015) that completion rates vary significantly according to course length with longer courses (in weeks) having lower completion rates. In the case of the IOL 1 and 2 courses, which both have time limits, this does not, however, seem to be a plausible explanation of the differences in retention rates in this study between the blended and distance modes. As outlined in Chapter 1.3.3.2, the blended courses are longer (13 weeks) than the distance courses (8 weeks) (Appendix B). Furthermore, it is noteworthy in Table 2, when observing student retention in the open self-directed modes (which are non-tutorial) as opposed to the distance modes (which are tutorial), that there is little difference in the completion rates between these two modes. This issue will be brought up again in Chapter 3.2.2.

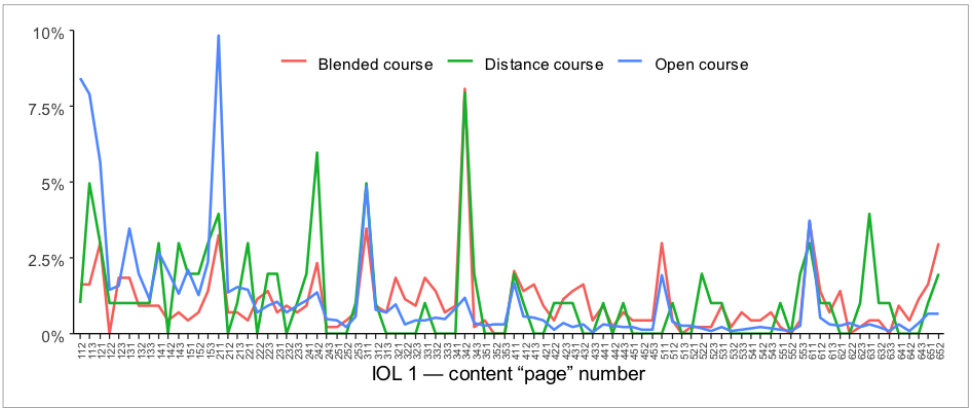
To summarize the answers to the first and second research questions in Chapter 3.1, the results showed that overall completion rates in the seven IOL courses were relatively low, spanning 2% to 18% depending on the course and mode. The results, furthermore, demonstrated that the blended mode of IOL 1 and 2 was more effective in retaining learners to the end than the other two modes. The differences in completion rates between the blended learners and the open self-directed learners in both IOL 1 and 2 were found to be statistically significant. The overall results, therefore, revealed that a large proportion of the participants in the study, or approximately 82% to 98% depending on courses or modes, did not go to the very last content page of a course. In order to develop a better understanding of user progress and retention, the population of non-completers was further investigated in the study, and the results are presented in the next section.

3.1.2 Attrition patterns in different modes and engagement patterns

The third research question addressed in Chapter 3.1 asked what the overall engagement pattern in IOL suggests about retention. In order to answer this question, first, the study investigated the attrition patterns among those who did not complete the course. The retention data were thus broken down in the attempt to reveal when during the program the learners disengaged. Two of the courses, IOL 1 and IOL 2 in the three different modes of delivery, were examined specifically for this purpose (Figure 2A–B). Then, the overall engagement behavior in all the seven courses was explored (Figure 3), including the IOL 2 course in the three different delivery modes (Figure 4).

Let us first turn to the results on the attrition patterns recognized in IOL 1 and 2 in terms of the three different modes of delivery. Using LA, the large set of tracking data from non-completers (n = 10,335) in the courses was explored in order to reveal their online behavior throughout the courses or modes. Figure 2A–B below reveals the patterns of attrition that were recognized in the data from these students. The red lines in the figure show the patterns for blended learners in IOL 1 and 2, the green lines reveal the patterns for distance learners in both courses, and the blue lines show the patterns for open self-directed learners in both courses. The x-axes in Figure 2A–B show the reference numbers of each content page in each course. To clarify, the last content page in the IOL 1 course has the reference number 653, and the last content page in IOL 2 has the number 534. The x-axes therefore show learners' behavior up to and including the second-to-last content page in each course, that is, to content pages number 652 (IOL 1) and 533 (IOL 2). The y-axes in the figure show the percentage of attrition on each content page in each course across all modes, namely, the proportion of each cohort disengaging from the course on each content page.

(A)



(B)

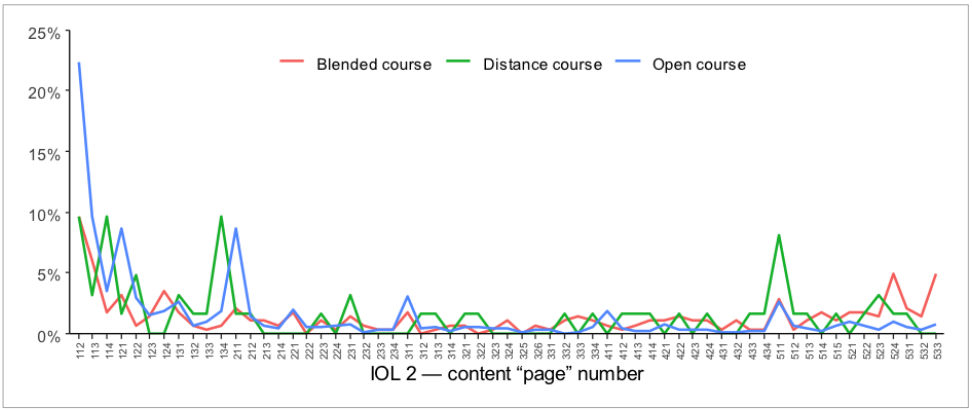


Figure 2 A–B. (A) Attrition patterns across the three different modes of delivery of IOL 1. (B) Attrition patterns across the three different modes of delivery of IOL 2.

The findings revealed in Figure 2A–B are first explained in terms of each mode of delivery of IOL 1, and then in terms of each mode of IOL 2. Turning to the findings in Figure 2A, the blue line shows the pattern for the open self-directed mode where 6,419 participants started the course. The results first revealed sharp drops on the first content page in the course, with an 8.4% (n = 539) attrition rate, and 22% (n = 1,408) cumulative attrition on the first three content pages (# 112–122), or when the learners had completed only 1% to 3% of the course's content. The next attrition peak in the open self-directed mode of IOL 1 appeared on content page # 211, when 17% of the course is covered, with a 9.8% attrition rate (n = 631). Finally, an attrition peak was identified on content page 311 with a 4.9% (n

= 312) attrition rate, when 34% of the course's content has been covered. Apart from that, there was more stability further on in the open self-directed course in IOL 1, with attrition rates ranging from 0.1% to 3.7% depending on the content page.

The red line in Figure 2A represents the blended mode of IOL 1, where 434 participants started the course. As illuminated in Figure 2A above, the results showed low attrition rates initially in the course, or cumulative attrition rate of 6.2% (n = 27) on the first three content pages. The overall attrition patterns further on in the blended course showed relatively low attrition peaks, with 0.2% to 3.7% attrition rates depending on content pages, except for content page # 342 with a 8.1% (n = 35) attrition rate when learners had covered 45% of the course's content. As is also shown in Figure 2A, some blended learners left the course on the second-to-last content page, # 652, when 99% of the course's content had been covered, with a 3% attrition rate (n = 13). Finally, the figure shows that there are some content pages in the blended mode where there was no attrition at all, even on two pages in a row. The data show seven such content pages in this mode.

The green line represents the distance mode of IOL 1 where 101 participants began the course. As Figure 2A shows, the results for this mode revealed relatively low attrition on the first three content pages, the same as in the blended mode, or as the cumulative attrition rate of 9% (n = 9). Three main drop-out points occurred first on content page # 242, when 28% of the course's content had been covered, with a 5.9% attrition rate (n = 6). Then, on content page # 311, when 34% of the content was covered, the attrition rate was 5% (n = 5). Finally, content page # 342, when 45% of the content had been covered, had an attrition rate of 7.9% (n = 8). Besides that, the attrition rates spanned 1% to 4% depending on the page. This includes the learners who left the course on the second-to-last content page when 99% of the course content was covered, showing a 2% attrition rate (n = 2). As Figure 2A above also illustrates, there are some content pages in the distance mode, even five pages in a row, where there is no attrition whatsoever. The data show 39 such content pages in this mode.

To sum up the findings on the attrition patterns revealed in IOL 1 (Figure 2A), the open self-directed mode stood out with respect to quite heavy attrition rates in the beginning of the course. An attrition peak was then identified in this mode when 34% of the course content had been covered. Apart from that, there was more stability further on in the course. On the other hand, the blended and distance modes of IOL 1 showed minor attrition early on in the courses. For the blended course, the main attrition peak was evident

on content page # 342 when 45% of the content had been covered. In the distance course, three attrition peaks were identified at 28% course completion, 34% course completion (on content page # 311), which involves the same content page and similar attrition rates mentioned for the open self-directed course, and at 45% course completion (on content page # 342), which relates to the same content page mentioned in the distance modes with similar attrition rates. Furthermore, the blended and distance modes of IOL 1 shared the fact that there was no attrition on many content pages, even many in a row, specifically in the distance mode. This result is interesting, and further consideration in the context of the learning material is needed in order to explain why some content pages seem to retain all learners. When the attrition data were further observed in that regard and compared between the blended and distance courses, it turned out that the two modes shared the same content pages of no-attrition in only a few cases.

Turning to the attrition patterns in IOL 2 that are revealed in Figure 2B, the blue line shows the pattern for the open self-directed mode of the course where 3,462 participants started. A huge drop-out, or 22.4% (n = 774) was revealed on the very first content page in the course (# 112), and a cumulative drop-out rate of nearly 36% (n = 1,229) on the three first pages (# 112–114) when 2% to 5% of the course content had been covered. Two other attrition peaks were identified early on in this mode, one on page # 121 when 6% of the content was covered with a rate of 8.7% (n = 301), and the other on page # 211 when 19% of the course's content has been completed with a rate of 8.6% (n = 298). There was more stability further on in the open self-directed mode, with attrition rates ranging from 0.1% to 3.1% depending on the content page.

The red line in Figure 2B concerns the pattern for the blended mode of IOL 2 where 281 participants started the course. The figure shows drop-out peaks early in this mode with a cumulative attrition rate of 17.4% (n = 49) on the first three content pages of the course. Two attrition peaks were then identified very late in the blended course, one on content page # 524 when 94% was covered with a rate of 5% (n = 14), and the other on page # 533 when 98% of the course content was covered. It also had an attrition rate of 5% (n = 14). There are four individual content pages in this mode without student attrition.

The green line in Figure 2B concerns the distance mode of IOL 2 where 62 participants began the course. The data for this mode revealed quite high drop-out peaks at the beginning as shown in Figure 2B, with a cumulative attrition rate of 22.6% (n = 14) on the first three content pages in the course. Then there are attrition peaks on content pages #

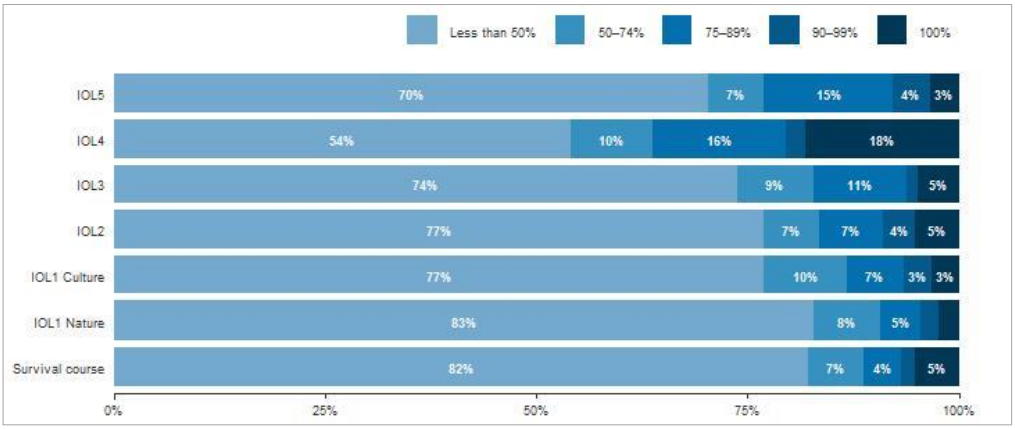
122 when 8% of the course content was covered with a rate of 4.8% (n = 3), and on page # 134 when 18% of the course content was completed with a rate of 9.7% (n = 6). A large attrition peak appeared late in the distance course, on content page # 511 when 81% of the course content is completed with a rate of 8.1% (n = 5). Apart from that, the attrition rates range from 1.6% to 3.2% across the course. However, there are 27 content pages in the distance mode with no attrition at all, with up to five content pages in a row.

To sum up the findings revealed in Figure 2B on the attrition patterns in IOL 2, quite sharp attrition was identified in the data for all the three modes of IOL 2 on the first three content pages when 2% to 5% of the course content had been covered, especially in the open self-directed mode. Other attrition peaks were then recognized early on in the open self-directed mode when 6% and 19% of the course content were covered; apart from that, there was more stability further on in the open self-directed course. As for the blended mode, the main two attrition peaks as identified in Figure 2B were found almost at the end of the course, when 94% to 98% of the course content had been covered. Finally, for the distance mode of IOL 2, two attrition peaks were found early on in the course, when 8% to 18% of the course had been completed, but a large attrition peak was also identified very late in the course, or when 81% of the course content had been covered. Equivalent to the blended and distances modes of IOL1, the blended and distance modes of IOL 2 had some common content pages with no attrition at all, particularly in the distance mode.

In conclusion, the analysis of the online behavior of non-completers in Figure 2A–B illuminated constant attrition throughout the courses and across delivery modes. Students from all groups were likely to disengage in the earlier stages of the courses, in particular in the open self-directed mode of IOL 2. The findings of this study are in accordance with previous research on the timing of drop-outs in MOOCs, which shows that attrition rates vary considerably across a course's early, middle, and late periods, while drop-out rates are highest at the beginning of the courses (Frydenberg, 2007; Greene et al., 2015; Ihantola et al., 2020; Jordan, 2014, 2015; Perna et al., 2014; Reich, 2014). However, the attrition patterns of learners in IOL 1 and 2 revealed in Figure 2A–B show that many learners disengaged when they are near the end of a course, particularly in the blended and distance modes; in this regard the results of the present study disagree with existing findings (de Freitas et al., 2015; Greene et al., 2015) that indicate that student attrition in MOOCs is highly unlikely when learners are near the end of a course. Based on the overall attrition patterns presented in Figure 2A–B, further questions arose about what is causing

large attrition on certain content pages, while there is no attrition on other pages. As shown in the study, attrition peaks were in some instances on the same content pages in two or all modes, while sometimes there was no consistency. The same is true as concerns the content pages of no-attrition in the blended and distances modes, in which only a few cases the two modes had in common the same content pages of no-attrition. This result raises the question of whether student attrition may not be attributable directly to content-related factors in the courses, but rather to the format of these courses. The findings presented in this chapter regarding the predominant point of initial drop-out in the open self-directed modes of IOL 1 and 2 can be considered in the context of previous research on retention in MOOC courses of different formats (Ihantola et al., 2020): a self-paced version with no deadlines and no acess to a tutor, and a fixed-schedule version where the course of study follows a predefined timeline. While that study observed that drop-out rates were greater at the beginning of the courses in both modes than towards the end, the initial attrition rates were higher in the self-paced MOOC. The results also showed that the fixed-schedule mode was much more likely to retain learners than the self-paced mode, especially early on in the course.

Let us now turn to the second part of the third research question in Chapter 3.1, which asked what the overall engagement patterns in IOL suggest about student retention. In the attempt to answer this question, the study investigated the pattern of user engagement across all the seven IOL courses in light of different parameters for coverage of the course's content, in order to illuminate the extent to which those considered non-course-completers engaged with the course material. Hence, as shown in Figure 3 below, instead of defining course completion as covering 100% of a course's content, parameters for course completion were adjusted to 90% to 99% coverage of a course's content, 75% to 89% coverage, 50% to 74% coverage, and finally to less than 50% coverage of a course's content.[xi] As the figure shows, the results uncovered various patterns of engagement across the courses when different parameters were used to examine retention. In addition to the large number that completed less than 50% of a course's content, it turns out that, apart from those who completed a course to the very last content page, many of the study's population completed for example 90% to 99%, 75% to 89%, or 50% to 74% of a course's content.



Note: coverage of course's content in IOL 1 – Culture theme and IOL 2 are based on an average of the three different modes of delivery in each course.

Figure 3. Overall engagement patterns in all seven IOL courses in terms of different parameters for coverage of course's content.

By using the same method to further investigate the engagement patterns in the three different modes of delivery of IOL 2, Figure 4 below shows identical patterns of engagement across the three modes when retention was examined in accordance with the above-mentioned parameters for coverage of the course's content. When focusing first on the blended learners, where 14% of those who started the course completed it to the very last content page, it turns out that 15% of them covered 90% to 99% of the course's content and 11% of them covered 75% to 89% of the content. Similarly, for the distance learners in Figure 4, where only 5% of those who started the course completed to the end, it was found that 19% of them covered 75% to 89% of the course's content. The engagement pattern for the open self-directed learners then showed that while 4% of those who began the course completed it, 7% of them covered 75% to 89% of the course's content.

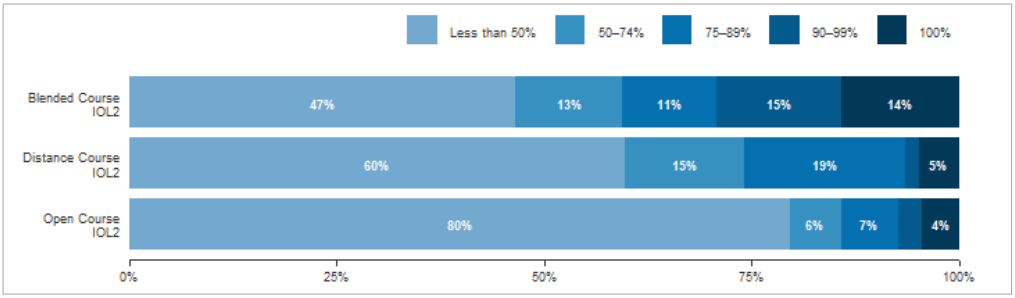


Figure 4. Overall engagement patterns in IOL 2 in the three different modes of delivery in terms of different parameters for coverage of course's content.

Overall, the findings shown in Figures 3 and 4 provide nuanced information on the engagement behavior of the learners in the IOL program. While many of those who did not complete the course were found to have covered only a small part of it, others had completed even the vast majority of the course content before they withdrew.

In summary, by mining tracked retention data on a large scale and using LA, the study presents a refined picture of users' online progress and the overall engagement patterns in the program, thereby providing valuable information about students' online behaviors and interactions in the context of their learning materials (Ebben & Murphy, 2014; Fischer, 2007; Gelan et al., 2018; Gillespie, 2020; Godwin-Jones, 2017; Hone & El Said, 2016; Martín-Monje et al., 2018; Thomas & Gelan, 2018). The findings in Figure 2 first revealed the patterns of attrition among the cohort of learners who did not fully cover the course content of IOL 1 and 2. These patterns illustrated constant high and low attrition rates at certain junctures across the three different modes of the courses, with high drop-out rates in common very early in the courses. This is in accordance with previous studies on student retention in MOOCs (Frydenberg, 2007; Greene et al., 2015; Ihantola et al., 2020; Jordan, 2014, 2015; Perna et al., 2014; Reich, 2014). In contrast with other studies (de Freitas et al., 2015; Greene et al., 2015), this study also found that many students who have covered the greater part of the course content may be likely to disengage before they have reached the final content page in the course. The patterns of student attrition that were revealed in the study raise further questions regarding factors that influence learner engagement and disengagement in such learning environments, and whether these factors relate to the course or to other individual factors (El Said, 2017; Hone & El Said, 2016; Kizilcec & Schneider, 2015; Littlejohn et al., 2016; Reich, 2014; Sokolik, 2014). Second, as shown in Figures 3 and 4, the study identified overall patterns of student engagement across all seven IOL courses, as well as the different modes of delivery in IOL 2. While the study revealed that most learners engaged with less than half of course materials, it also brought to light that learners may complete the majority of a course's content despite not seeing the course through to the very end (Greene et al., 2015). These findings relate to the discussion of low completion rates in MOOCs and the critical question of what it means to complete a course, and that mere completion rate may not be a relevant metric for student engagement and retention in MOOCs (Hew, 2016; Koller et al., 2013; Reich, 2014). The findings, therefore, called for a reevaluation of earlier parameters by which retention is measured and laid the foundation for the parameters to measure retention in the follow-up

studies. The next section sums up the main findings revealed in Chapter 3.1 and discusses the premises established for the two subsequent studies that are presented in Chapters 3.2 and 3.3.

3.1.3 Summary

To summarize the answers to the three research questions in Chapter 3.1 on the overall retention in the IOL program, the impact of the mode of delivery on retention and what the overall engagement pattern in IOL may suggest about retention, the tracking data first found that overall student retention was relatively low across all the seven courses and modes of IOL, ranging from 2.4% to 18.2% (Table 1). Second, the data revealed that the blended mode was most effective in retaining learners in both IOL 1 and 2, with an 8.3% completion rate in IOL 1 compared to 4% in the distance mode, and 2.9% in the open self-directed mode, and with a 14.2% completion rate in the blended mode of IOL 2 compared to 4.8% in the distance mode and 4.4% in the open self-directed mode (Table 2). Third, when mining the data on the online behavior of non-completers in the IOL 1 and 2 courses across different modes, the analysis of their progress throughout the course illuminated a regular pattern of attrition in all the different modes of delivery (Figure 2A–B). Accordingly, the study identified relatively high attrition rates at certain junctures in the courses, in particular at the beginning of the courses. In some cases, the attrition peaks were on the same content pages across modes, while in other instances there was no consistency. Interestingly, relatively high attrition peaks were also recognized very late in the courses, specifically in the blended and distances modes of IOL 2. Finally, the investigation on the overall engagement patterns on the IOL program illuminated that many of the learners in the program covered the majority of a course's content even though they did not complete the very last content page of a course (Figures 3 and 4). Thus, when student retention, both overall in the seven courses and across modes of IOL 2, was considered in the light of different parameters for course completion, the results showed various paradigms of engagement and revealed considerably higher retention rates when the parameters for course completion were modified to, for instance, 90% to 99% or 75% to 89% coverage of the course's content.

The results revealed in this chapter provided the premises for subsequent sampling, data collection, and analysis in the two follow-up survey studies discussed in the next two chapters. In that regard, a focused sample of learners in the IOL 2 course was specifically

chosen in order to investigate the potential effect of course-specific and other motivational factors on student retention. Thus, first, based on Figure 2B (Chapter 3.1.2), which revealed relatively sharp drop-outs initially in the course, a decision was made to narrow the target sample in the following studies to those who completed more than 15.9% of the course's content, therefore excluding data on those who covered only a few initial pages and had little experience using the learning material. Second, also based on Figure 2B and Figures 3 and 4 (Chapter 3.1.2), the relatively large attrition rates almost at the end of IOL 2 supported arguments for reevaluation of earlier parameters for the measurement of course completion in the follow-up studies. As the findings show, many of the learners had covered the course's content almost until the end when they disengaged. Thus, instead of considering course completion defined as 100% coverage of course's content, course completion was defined in the follow-up studies as learners who completed 80%[xii] to 100% of the course's content. For further clarification, the engagement pattern in IOL 2 is revised in Figure 5 below, which illustrates the measure set for the target sample as well as for the parameters for measuring course completion in the subsequent studies.

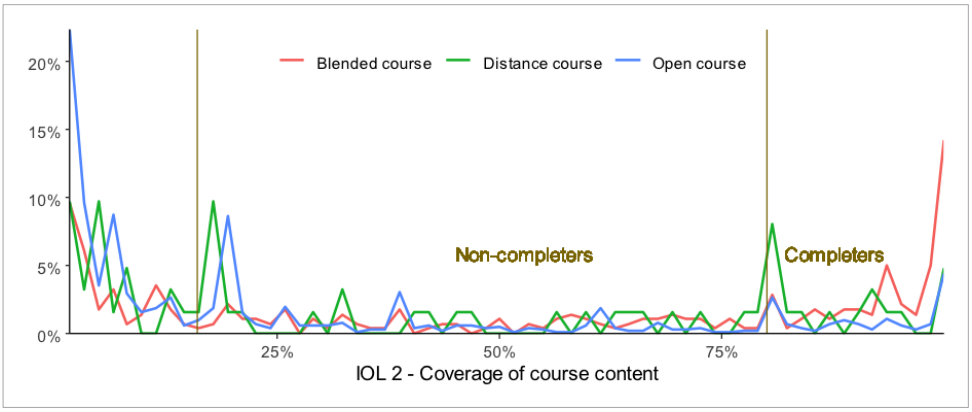


Figure 5. Illustration of the parameters set for the target sample and for those who are considered course completers in IOL 2 in the follow-up survey studies.

The following two main chapters highlight the results of the subsequent survey studies, which were geared towards learners in the IOL 2 course and addressed their views on particular components that related to the learning materials in focus as well as other motivational factors, and their potential impact on retention.

3.2 Survey Data: The Effect of Content and Tutor, and Other Motivational Factors on Retention

This chapter focuses primarily upon the survey data results which addressed the questions of whether factors related to the content in IOL 2 (Chapter 3.2.1) and tutor support in two of the course's modes (3.2.2) were important for learner engagement, and whether these factors affected retention. Furthermore, this chapter (Chapter 3.2.3) discusses the findings on the issue of whether learners in IOL 2 began the course with the aim of completing it, and the potential impact of the initial goal factor on retention. In addition, the latent issue of the effect of age and gender on retention is addressed briefly in this chapter (Chapter 3.2.4). The discussion in the following sections is structured around research questions B, C, D, and F presented in Chapter 1.4, which are reviewed within each sub-chapter. The findings are then summarized in Chapter 3.2.5.

The survey and retention data presented in the following chapters come from 400 learners who were enrolled in any one of the three modes of delivery of IOL 2 from 2010 to 2018.

3.2.1 Content-specific factors and retention

The study on the effect of content-specific factors in IOL 2 on student retention (Article III), was based on the following research questions.

> Do learners in IOL 2 consider the following factors pertaining to the structure and organization of the course and the design and pedagogical principles important for their motivation to carry on in the course? If so, are they more likely to complete than those who consider them unimportant? The factors are as follows:
>
> a. Curated and sequenced course structure,
> b. Clear and salient learning objectives,
> c. Gradual and scaffolded presentation of input,
> d. Continuing storylines,
> e. Form-focused and scaffolded presentation of grammar,
> f. Variety in types of learning objects.

To address the first part of this question of whether the learners considered the content factors important for their motivation to carry on in the course, the survey responses (n = 400) were tabulated in regard to the six content-specific factors. As Table 3 shows, the

sample contained respondents who thought of these factors as significant motivators (marked 'strongly agree'/'agree' or 'definitely'/'probably' in response to relevant statements), against respondents who did not consider them important for motivation (marked 'neither agree nor disagree'/'disagree'/'strongly disagree' or 'possibly'/'probably not'/'definitely not').[xiii] First, the two factors regarding the structure and organization of the course content are considered and then the factors concerning the instructional strategies.

Table 3. Students' reports on the importance of content-specific factors in IOL 2 for their motivation.

	Important for motivation	Not important for motivation
STRUCTURE AND ORGANIZATION:		
Curated and sequenced course structure	73.8% (n = 295)	26.2% (n = 105)
Clear and salient learning objectives	55% (n = 220)	45% (n = 180)
INSTRUCTIONAL STRATEGIES:		
Gradual and scaffolded presentation of input	84.8% (n = 339)	15.2% (n = 61)
Variety in types of learning objects	82.2% (n = 329)	17.8% (n = 71)
Continuing storylines	69.8% (n = 279)	30.2% (n = 121)
Form-focused and scaffolded presentation of grammar	66.2% (n = 265)	33.8% (n = 135)

When focusing first on the factor of curated and sequenced course structure, the results showed that 73.8% (n = 295) of the participants considered this factor to be important for their motivation. Connected to this factor, the survey asked two additional questions concerning participants' experiences of the ease of navigating the website, and whether they accessed the material in the suggested order. The findings revealed that 80% (n = 320) found the navigation easy and that 88% (n = 350) navigated through the course as suggested. For the second factor in Table 3, clear and salient learning objectives, the

results revealed that 55% (n = 220) of them considered this factor to be important for their motivation. In related sub-questions that asked whether the presentation of the learning objectives helped to keep them focused and whether they would have preferred to set their own learning goals, the results showed that 60% (n = 240) considered such presentation of objectives helpful and that only 22.8% (n = 91) agreed that they would have preferred to set their own learning goals.

In table 3, the third factor of gradual and scaffolded presentation of input seems to be highly appreciated by the learners, or by 84.8% (n = 339). Similarly, the fourth factor, variety in types of learning objects, turns out to be the second most important factor with a response rate of 82.2% (n = 329). Four sub-questions were addressed in the survey concerning the learning objects, namely a) whether participants thought that the diverse learning objects across the course made it fun to work with, b) if the learning objects were clearly exhibited such that the users understood what to do, c) whether they accessed the attached audio files, and d) whether they asked for feedback that was provided for most of the learning objects in the course. According to the data, almost 89% (n = 325) of participants stated that the variety of learning objects made the course fun and that around 81% (n = 325) thought that they were clearly presented, while almost 79% (n = 315) claimed that they utilized the audio files frequently and nearly 78% (n = 310) asked frequently for feedback.

As for the fourth factor of continuing storylines, Table 3 shows that 69.8% (n = 279) of the participants responded that this factor was important for their motivation. When asked in a sub-question whether using the storylines made the course fun, 72% (n = 288) agreed. Finally, concerning the factor of form-focused and scaffolded presentation of grammar, the data revealed that 66.2% (n = 265) of respondents felt that this kind of grammar presentation was supportive in keeping them motivated. Moreover, participants' answers to sub-questions in the survey regarding the grammar aid provided in the course indicate that almost 51% (n = 203) frequently continued to the first step in the grammar resource (referred to as 'grammar help' in the survey), and that around 63% (n = 253) sought more information in the grammar resource (referred to as 'read more' in the survey).

To sum up the answers to the first section of the research question above, the findings revealed that the majority of learners in the study considered all six content

factors in focus to be important motivators. The results indicate that the instructional methodology and engagement strategies employed in IOL 2 may be useful in encouraging the language learner. The findings relate to previous research underlining the value of MOOC structural design, and that a facilitative and salient structure of course content that includes clear learning objectives spelled out in advance may be of benefit in motivating learners and maintaining their focus (Castrillo, 2014; Dörnyei et al., 2014; El Said, 2017; Garrett, 1991; Hubbard, 2013; Kizilcec & Schneider, 2015; Ross et al., 2014). The data suggest, moreover, that language learners greatly appreciate a gradual and scaffolded presentation of input, which involves scaffolding of information under carefully organized step-by-step guidance. This supports previous evidence of the importance of providing learners with careful scaffolding of course content, information, and tasks in such online learning environments (Castrillo, 2014; Doiz et al., 2014; Hew, 2016; Rosenshine & Meister, 1992; Teixeira & Mota, 2014). Furthermore, the findings support earlier research that highlights the value of diverse and clearly presented learning tasks and appealing content in order to call forth and sustain learners' motivation (de Barba et al., 2016; Hew, 2016). In addition, the findings imply that technologically enhanced devices may encourage the LMOOC learner in the learning process (Arnbjörnsdóttir, 2004; Chun, 2012; Colpaert, 2010, 2014).

In addition to the six content factors that have been discussed so far and presented in Table 3 above, there were several sub-questions addressed in the survey that were intended to gather additional data on the learner's actions and usage of such a program. One question addressed whether learners navigate the course linearly. As the findings show, the vast majority of participants in the study accessed the course content in the suggested order. These findings are in harmony with previous findings (Perna et al., 2014) that indicate that most MOOC users progress through a course sequentially in the order identified by the course instructor, rather than determining their own approach to accessing content. Another sub-question asked whether learners utilize the supplementary material offered within the IOL program as extra support. Based on the findings, most learners in the study used the available resources attached to the learning objects in the course, such as the audio files and automated feedback on tasks, as well as the grammar help and additional information supplied by the grammar resource. According to previous research (Hew, 2016), MOOC learners appear to highly appreciate features that provide prompt feedback on their assignments and access to relevant resources (Chun, 2001). Hew's

(2016) study suggested that these may be crucial strategies for educators to promote student engagement in fully online courses. The additional data presented in the current study, particularly as regards participants' use of the option to receive feedback on learning objects and their use of supplementary resources, provide a basis for future studies on the impact of these features on student engagement and retention.

To address the second part of the research question, whether learners who valued these content-specific factors as important motivators were more likely to complete the course than those who disagreed, survey data (n = 400) were tabulated with regard to the factors and then measured against tracking data.

Table 4. Retention of learners who considered content-specific factors in IOL 2 important for their motivation as opposed to those who did not.

	Important for motivation		Not important for motivation	
STRUCTURE AND ORGANIZATION:	Completed	Did not complete	Completed	Did not complete
Curated and sequenced course structure	45.8% (n = 135)	54.2% (n = 160)	38.1% (n = 40)	61.9% (n = 65)
Clear and salient learning objectives	42.7% (n = 94)	57.3% (n = 126)	45% (n = 81)	55% (n = 99)
INSTRUCTIONAL STRATEGIES:				
Gradual and scaffolded presentation of input	46% (n = 156)	54% (n = 183)	31.1% (n = 19)	68.9% (n = 42)
Variety in types of learning objects	45% (n = 148)	55% (n = 181)	38% (n = 27)	62% (n = 44)
Form-focused and scaffolded presentation of grammar	43.4% (n = 115)	56.6% (n = 150)	44.4% (n = 60)	55.5% (n = 75)
Continuing storylines	43% (n = 120)	57% (n = 159)	45.5% (n = 55)	54.5% (n = 66)

When considering first the factor of curated and sequenced course structure, the results showed that 45.8% (n = 135) of those who found this factor to be important for their motivation completed the course while only 38.1% (n = 40) of those who did not agree

completed. For the second factor of clear and salient learning objectives, the findings showed that 42.7% (n = 94) of those who considered this factor to be important for their motivation completed the course, in comparison to 45% (n = 81) of those who disagreed.

Regarding the third factor, gradual and scaffolded presentation of input, 46% (n = 156) of those who identified this factor as important completed the course, as compared to only 31.1% (n = 19) of those who disagreed. For the fourth factor, variety in types of learning objects, the results show that 45% (n = 148) of those who reported this factor to be important for their motivation completed, while only 38% (n = 27) of the other group completed. For the fifth factor of form-focused and scaffolded presentation of grammar, the study found that 43.4% (n = 115) of those who considered it to be important for their motivation completed, while 44.4% (n = 60) of those who disagreed completed. Finally, when looking at the impact of the factor of continuing storylines, 43% (n = 120) of those who valued it important for their motivation completed, while 45.5% (n = 55) of the other group completed.

The findings in Table 4 indicate that there is a positive association between three of the factors and course completion, namely, the factors of curated and sequenced course structure, gradual and scaffolded presentation of input, and variety in types of learning objects. When measured against the tracking data, participants who considered these factors important for their motivation were more likely to complete the course than those who did not consider them important. A chi-square test of homogeneity indicates a statistically significant difference in completion rates ($p = 0.0439$) between those who experienced the factor of gradual and scaffolded presentation of input to be important for their motivation and those who did not. No statistically significant differences, however, were found in completion rates between the two comparison groups regarding the factor of curated and sequenced course structure ($p = 0.2129$) or the factor of variety in types of learning objects ($p = 0.3474$). The other features in Table 4, clear and salient learning objectives, form-focused and scaffolded presentation of grammar, and continuing storylines were not found to have a positive impact on course completion in the study. This result suggests that course developers should pay more attention to these factors. The findings of the study contribute to the arguments that course content must be curated for online language learners, and that careful scaffolding of instruction and wide-ranging course content may be crucial to engage and retain the LMOOC learner (Arnbjörnsdóttir, 2004; Castrillo, 2014; de Freitas et al., 2015; Doiz et al., 2014; Hubbard, 2014; Ross et al., 2014; Sokolik, 2014; Teixeira & Mota, 2014).

To summarize the answer to the question of the effect of IOL 2's course content on student retention, the majority of the participants in the study, or between 55% and 85%, found great value in all the content-specific factors and considered them important for their motivation to carry on with the course. This is of particular concern to factors related to gradual and scaffolded presentation of input and the variety in types of learning objects. Since the principal goal of the IOL course was to engage learners and support their learning process, the majority of survey respondents demonstrated that the course was successful. When the survey data were measured against the tracking data, the findings showed that learners who considered three out of the six factors to be engaging elements were more likely to complete than those who did not, where the factor of gradual and scaffolded presentation of input was found to have a statistically significant association with course completion. Overall, the results of the study partly support the hypothesis presented in Chapter 1.4, that is, that the specific elements of the IOL design concerning course structure and instructional methodology would have a positive impact on student engagement and retention in the present study. The study has identified multiple elements of LMOOC course design and CALL practices that appear to encourage learner engagement and, in some cases, retention. The next section presents the results on the learners' views about tutor-specific factors and their effect on retention.

3.2.2 Tutor-specific factors and retention

The study on the effect of tutor-specific factors in IOL 2 on student retention (Article II), was based on the following research questions.

> Do learners in IOL 2, in the blended and distance modes, consider the following tutor-specific factors important for their motivation to carry on in the course? If so, are they more likely to complete than those who consider them unimportant? The factors are as follows:
>
> a. Private interaction with the tutor,
>
> b. Overall tutor support during the course,
>
> c. A detailed introduction of the program,
>
> d. A set syllabus where course content is pre-organized into timetabled, manageable sections.

To address the first part of the question of whether or not the learners in the blended and distance modes of IOL 2 considered the tutored factors important for their motivation to carry on in the course, the survey responses (n = 64) were tabulated with regard to the four tutor-specific factors. As Table 5 shows, the sample contained respondents who thought of these factors as significant motivators (marked 'strongly agree'/'agree' or 'very important'/'moderately important' in response to relevant statements), against respondents who did not consider them important for motivation (marked 'neither agree nor disagree'/'disagree'/'strongly disagree' or 'not important').

Table 5. Students' reports on the importance of tutor-specific factors in the blended and distance modes of IOL 2 for their motivation.

	Blended mode (n = 41)		Distance mode (n = 23)	
	Important for motivation	Not important for motivation	Important for motivation	Not important for motivation
Set syllabus	90.2% (n = 37)	9.8% (n = 4)	91.3% (n = 21)	8.7% (n = 2)
Private interaction with the tutor[a]	87.5% (n = 7)	12.5% (n = 1)	86.7% (n = 13)	13.3% (n = 2)
Detailed introduction of the program	56.1% (n = 23)	43.9% (n = 18)	60.9% (n = 14)	39.1% (n = 9)
Overall tutor support	51.2% (n = 21)	48.8% (n = 20)	82.6% (n = 19)	17.4% (n = 4)

a. Only those who had stated in a previous question in the survey that they had asked for a private meeting with the tutor or interaction via email received this question, which explains the low response rate concerning this factor (Appendix E).

Regarding the first factor in Table 5 of set syllabus, the results show that 90.2% (n = 37) of participants in the blended mode and 91.3% (n = 21) in the distance mode considered this factor to be important for their motivation to carry on with the course. Similarly, with regard to the second factor in the table, the findings show that 87.5% (n = 7) of those who had received private instruction from a tutor in the blended mode and 86.7% (n = 13) in the distance mode believed that this factor was important for their motivation. The results on the third factor of detailed introduction of the program show that this factor was considered important by most of the participants in both modes, or by 56.1% (n = 23) in the blended mode and 60.9% (n = 14) in the distance mode. Finally, the factor of overall tutor support was believed to encourage course engagement by both groups, the distance mode in

particular. In the blended mode, 51.2% (n = 21) of the participants considered this factor important, while 82.6% (n = 19) in the distance mode found this factor valuable for their engagement.

To conclude the answer to the question of the importance of these tutor-specific factors for learner engagement, Table 5 reveals that the majority of learners in the blended and distance modes considered all factors to be important for their motivation. The results suggest that the tutor involvement and guidance provided in these two modes may benefit learner engagement with the course content. As the data show, learners in both modes specifically appreciated the factors of set syllabus, whereby predetermined learning material is assigned on a weekly basis, and of private interaction with the tutor, whereby the tutor is available for help at any time. These findings contribute to earlier arguments for the necessity of providing tutor support and supervision in MOOCs, and that syllabuses and proper introduction to a course's prerequisites may be helpful in such self-directed learning (El Said, 2017; Hew, 2016; Hew & Cheung, 2014; Ross et al., 2014; Rubio et al., 2018; Sokolik, 2014). The results of the study seem to support previous findings (Hew, 2016) that clear course information and a course syllabus which shows learners what they are supposed to do and how much effort they are expected to put into the course per week may be a critical factor in student engagement. However, the result on the factor of overall tutor support in Table 5 raised questions as to why the blended learners in the study appeared to value this factor less than the distance learners. As noted in Chapter 1.3.3.2, the blended course is delivered on campus where most of the learners in that mode supposedly take two other face-to-face courses in the Practical Diploma Program simultaneously, where they have direct contact with a teacher in a classroom. Based on this, one may wonder whether the blended learners' sense and expectations of 'tutor support' in this particular course, which is all online and self-guided, may in some ways be different from the expectations of the distance learners, who study on their own and whose only option for learning support is from an online tutor. Further research is needed on this matter (Henderikx et al., 2018).

In order to address the second part of the research question of whether learners who valued these tutor-specific factors as important motivators are more likely to complete than those who disagreed, survey data (n = 64) from the learners in the two tutored modes were tabulated with respect to the four factors and then measured against tracking data.

Table 6. Retention of learners in the blended and distance modes who considered the tutor-specific factors in IOL 2 important for their motivation as opposed to those who did not.

	Blended mode (n = 41)		Distance mode (n = 23)	
	Important for motivation	Not important for motivation	Important for motivation	Not important for motivation
	Completed	Completed	Completed	Completed
Overall tutor support	81% (n = 17)	55% (n = 11)	36.8% (n = 7)	50% (n = 2)
Detailed introduction of the program	78.3% (n = 18)	55.6% (n = 10)	28.6% (n = 4)	55.6% (n = 5)
Set syllabus	70.3% (n = 26)	50% (n = 2)	33.3% (n = 7)	100% (n =2)
Private interaction with the tutor	57.1% (n = 4)	0% (n = 0)	30.8% (n = 4)	50% (n = 1)

The findings on the potential impact of the tutor-specific factors on student retention presented in Table 6 will first be considered with regard to the data on the blended learners and then the distance learners. As shown in the table, the results revealed that 81% (n = 17) of the blended learners who viewed the factor of overall tutor support to be important for their motivation did, in fact, complete the course, while only 55% (n = 11) of those who considered this factor unimportant completed it. Concerning the second factor of detailed introduction of the program, the findings showed that 78.3% (n = 18) of the blended learners who considered this factor to be an engaging element in the course completed the course, in comparison to 55.6% (n = 10) of those who disagreed. Regarding the third factor, set syllabus, the results showed that 70.3% (n = 26) of those who reported this factor to be important for their motivation completed, in comparison to 50% (n = 2) of those who disagreed. Finally, the evaluation of the factor of private interaction with the tutor shows that 57.1% (n = 4) of the blended learners who found this factor to be engaging completed the course. No comparison group was available in relation to this factor.

The results regarding the learners in the blended mode indicate a positive association between all tutor-specific factors and course completion overall. When measured against the tracking data, the learners who considered these elements important for their motivation were more likely to complete the course than those who did not consider them important. Even though the sample size was limited in the blended mode,

part of these data was run through an analysis of statistical significance. Marginally nonsignificant differences in completion rates ($p = 0.05801$) were observed between those who considered the factor of overall tutor support important for their motivation and those who did not. Similarly, no statistically significant differences in completion rates ($p = 0.1127$) were found between those who considered the factor of detailed introduction of the program important for their motivation and those who did not.

When turning to the data on the distance learners in Table 6 and the potential impact of these same factors on their retention, the results revealed quite the opposite picture. It turned out that, even though the majority of the distance learners considered all four tutor-specific factors to be important for their motivation in the course, as shown in Table 5, these factors were not found to have a positive impact on course completion. This result suggests that course developers should give more attention to these factors in the distance learning mode. These interesting findings raise further questions. Since the learners in the distance and blended modes interact with the same learning material, both groups are provided with tutor support and the option of communicating with a tutor, and since both modes include paying groups (Koller et al., 2013), these findings may raise the question of whether there are other factors assigned to the mode of delivery or the users that may explain why the blended learning mode is more effective in retaining learners than the distance mode. Returning to the results in Chapter 3.1.1 (Table 2) that compared overall student retention in the open self-directed mode ($n = 3{,}462$), the distance mode ($n = 62$) and the blended mode ($n = 281$) in IOL 2, it turned out that there was little difference in the completion rates between the non-tutorial open self-directed mode and the tutorial distance mode of IOL 2. In relation to these earlier findings, and when considering the results illuminated in this section that show that tutor-specific factors do not seem to have a positive impact on retention in the distance mode, the results suggest that there are other elements than the tutor-related factors that may play a decisive role in student retention in the blended and distance modes. As will be discussed in the following chapter and Chapter 3.3, various motivational and outside factors were identified in the study that need to be considered in this context and in the discussion on student retention in such learning environments.

To conclude the findings regarding the research question in Chapter 3.2.2 on learners' views towards the tutoring factors in IOL 2 and their effect on retention, the

results showed that a large proportion of the participants in the blended and distance modes, or between 50% to 90%, depending on the factor and mode of delivery, found that all the four tutor-specific factors in question seemed to be essential for their motivation to continue the course. This particularly concerns the factors of set syllabus and private interaction with the tutor. Furthermore, when measured against the tracking data, the results showed that learners in the blended mode who considered these factors to be important were more likely to complete the course than those who did not consider them to be important for their motivation. The opposite was the case when the group of distance learners was explored, which suggests that other elements than the tutoring factors may be involved when trying to explain the difference in student retention between the blended and distance learners in the study. Overall, the findings of the study partly support the hypothesis presented in Chapter 1.4, that is, that factors relating to tutor support and personal assistance in the tutorial modes of IOL would have a positive impact on student engagement and retention. The study has identified various features of LMOOC learning environments that seem to be important for learner engagement with a course. The following section concentrates on student retention with regars to the question of whether learners in IOL 2 had the initial goal of completing the course once they started.

3.2.3 Learners' initial intent in terms of course engagement and retention

The study on learners' initial goals for course engagement and their effect on retention (Article II) was based on the following research questions.

> Do learners in IOL 2, both overall and within each mode, have the initial goal of completing the course? If so, are they more likely to complete than those who do not have such a goal?

The data were first analyzed in terms of learners overall in the IOL 2 course, based on a survey question that asked whether they came to the course with a) the goal of completing the course, b) the goal of working on part of the course or c) no clear goal. Second, the potential impact of the initial goal of course completion upon overall completion rates was explored. Third, the data were analyzed according to the three different modes in order to shed light on learners' intended course engagement in each mode. Finally, the potential influence of the goal of course completion upon actual course completion was investigated according to mode of delivery.

By first examining whether or not the learners in the survey (n = 400) had initially intended to complete the course once they started, the results showed that 56,5% (n = 226) of them had the stated initial intention of completing the course while 15% (n = 60) had the goal of working on part of the course. Almost 28,5% (n = 114) had no clear goal. It was thus noticeable that 43,5% (n = 174) of the participants in the study had no goal of completing the course when they attended.

When investigating the potential impact of the initial goal factor on overall student retention in IOL 2, the results presented in Table 7 below showed that this factor seemed to affect retention when measured against the tracking data. As evident from the table, the sample contained respondents who had the initial intention of completing the course (marked 'had the goal to complete' in response to relevant statement), against those who did not have such a goal (marked that they either had the goal to work on part of the course, or no clear goal).[xiv]

Table 7. Retention of learners overall in IOL 2 who had the initial goal of completing the course as opposed to those who had no such goal.

	Had the goal to complete (n = 226)	Had no goal to complete (n = 174)
Completed	52.7% (n = 119)	32.2% (n = 56)
Did not complete	47.3% (n = 107)	67.8% (n = 118)

The results revealed that learners who had a goal of completing were more likely to do so than those who had no such intentions. Almost 53% (n = 119) of those who intended to complete the course did so, as measured against the tracking data, while 32.2% (n = 56) of those who had no such goal completed the course. A chi-square test of homogeneity indicated a statistically significant difference in completion rates (p = 0.001) between those who had the initial goal of completing the course and those who did not have such goal. These results are in line with previous findings that students in MOOCs who have the initial aim of completing a course are more likely to do so than those who have not (Reich, 2014). Moreover, it is noticeable in Table 7 that many of the learners who had the goal of completing were not successful while some who did not intend to complete the course ultimately did so, which is also in accordance with Reich's findings (2014). These results therefore support the hypothesis presented in Chapter 1.4 that learners in IOL 2 would have different goals in terms of course engagement, and that the initial goal factor would

have a positive impact on student retention in the study. The findings furthermore contribute to the argument that the discussion of low completion rates in MOOCs must take into account; namely, that many students may never intend to complete a course in the first place. This argument underlines the significance of individual perspectives when evaluating progression and drop-out rates in MOOCs. Many attend these courses with an open mind, and report that they visit a program to find out what a course is about, to browse, or to complete only some course activities in order to develop specific skills and knowledge (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014; Wang & Baker, 2015).

On the other hand, when focusing on the learners in different delivery modes of IOL 2, shown in Table 8 below, and addressing the question on whether or not each group had the initial intent of completing the course and its impact on retention, the results showed that the majority of the blended learners, or 75,6% (n = 31), had a stated goal of completing the course when they enrolled. Similarly, the table shows that all of the distance learners (n = 23) had had this goal when attending the course. However, as the table also exhibits, less than half of the learners in the open self-directed mode, or 47,8% (n = 133), came into the course with the goal of completing it, which is in line with aforementioned considerations that some MOOC learners attend these courses having no goal of finishing (Henderikx et al., 2017; Koller et al., 2013; Reich, 2014; Wang & Baker, 2015).

When considering the potential impact of the goal to complete on course completion across modes, where the survey data on learners in the three modes were measured against their tracking data, it turned out, as shown in Table 8, that the blended learners and the open self-directed learners who had a stated goal of completing the course when they enrolled were more likely to do so than those who had no such goal in these modes. Regarding the blended learners, the results showed that 77.4% (n = 24) of those who had the goal of completing the course completed it in comparison to 40% (n = 4) of those who did not have this goal. Concerning the open self-directed mode, the findings showed that 48.9% (n = 65) of the learners who intended to complete the course completed, while only 32.4% (n = 47) of those who did not have this goal completed the course. These results support earlier findings that committed learners who have an objective of completing a course are more likely to succeed in doing so than those who have no such goal (Koller et al., 2013; Reich, 2014). A chi-square test of homogeneity showed a statistically significant difference in completion rates in the blended mode (p =

0.03435) and the open self-directed mode (p = 0.003765) between those who had an initial goal of completing and those who did not. All participating distance learners, however, stated an initial goal of completion when they started and did not allow for comparison.

Table 8. Retention of learners by mode of delivery in IOL 2 who had the initial goal of completing the course as opposed to those who had no such goal.

	Had the goal to complete		Had no goal to complete	
Blended mode (n = 41)	75.6% (n = 31)		24.4% (n = 10)	
	Completed	Did not complete	Completed	Did not complete
	77.4% (n = 24)	22.6% (n = 7)	40% (n = 4)	60% (n = 6)
Distance mode (n = 23)	100% (n = 23)			
	Completed	Did not complete		
	39.1% (n = 9)	60.9% (n = 14)		
Open self-directed mode (n = 278)	47.8% (n = 133)		52.2% (n = 145)	
	Completed	Did not complete	Completed	Did not complete
	48.9% (n = 65)	51.1% (n = 68)	32.4% (n = 47)	67.6% (n = 98)

To summarize the answers to the research questions of whether learners in IOL 2 had the goal to complete the course, and whether having such a goal had an impact on student retention, it turned out that most learners in the course overall, as well as in the blended and distance modes, had the stated intent to complete the course when they started. Accordingly, the results showed that many learners, such as in the open self-directed mode, had no initial goal of completing the course whatsoever, which is clearly an issue that the discussion of low retention rates in MOOCs must address. The results also showed a positive impact of the initial goal of completion both overall in the course and within the blended and open self-directed modes.

To further clarify and contextualize the investigation of the potential impact of the initial goal factor by mode of delivery in particular, the findings of the tracking data in the

first phase of the study (Article I), showing that the blended learning mode of both IOL 1 and 2 was more efficient in retaining learners than the other modes, called for further studies. An attempt was therefore made in the follow-up study (Article II) to explain why the blended learning mode had this unique position in the study.[xv] In order to do so, that study examined the potential effect of tutorial factors and other motivational factors, such as an initial goal to complete a course, on retention in the different modes of IOL 2. The findings revealed in the present chapter, that is, that most of the blended learners in the study had the initial intent of completing the course and succeeded in doing so, can therefore partially explain why the blended learners persisted in the course over learners in the other modes. In the following chapter, the influence of demographic variables on retention will be discussed.

3.2.4 The factors of age and gender and retention

The learners' mean age is relatively high in IOL 2, as mentioned in Chapter 1.4, which raised questions regarding the potential impact of demographics on student retention in the study. The research question was primarily rooted in the issue of whether age might explain some of the variation that was identified in the data regarding the different modes of delivery in the study. As was illuminated in Chapter 3.1.1, the blended learning mode was found to be in a unique position as the most effective in retaining learners in both IOL 1 and 2. Furthermore, as discussed in Chapter 3.2.2, the survey data analysis on the possible impact of the tutored factors on retention in the blended and distance modes of IOL 2 revealed that these factors seemed to have a positive impact on student retention in the blended mode, but not in the distance mode. For this reason, an analysis of retention data according to participants' demographic information for age and gender was included in the process of interpreting the overall results of the research based on the following research question.

> Do demographic factors such as age and gender have a predictive value in relation to student retention?

The survey data (n = 400) revealed that ages ranged from 16 to 87, with a mean age of 39, in the IOL 2 course, and that 60% of the population were female and 40% male (Chapter 2.2.1). When the participants' age was explored specifically by mode of delivery, the data showed that the lowest mean age was in the blended learning mode (33.3 years), the highest mean age was in the distance mode (43.7 years), and that the open self-directed

mode had the mean age of 40.7 years. The presentation of results will first focus on the factor of age. Grounded on the research question above and relevant discussion in Chapter 1.4, the hypothesis predicted to find age differences in terms of retention, and that the group of older learners is less likely to complete the course than the younger groups investigated in the study. To test this hypothesis, a linear regression was performed, including age as the independent variable and retention as the dependent variable.

When first considering the participants overall in IOL 2, regardless of mode, the findings showed a negative linear relationship between age and retention (slope -0.002589) and that the independent variable was statistically significant (p = 0,00524). Based on the downhill linear relationship revealed in Figure 6, the findings suggest that the higher the learners' age, the smaller the proportion of the course's content they cover on average. Even though the points in the figure are somewhat scattered in a wider band, showing a slight downhill (negative) relationship, they indicate that a linear relationship is present. As concerns the different modes of delivery, the study similarly found a negative linear relationship between these two variables in the open self-directed mode (n = 278) (slope -0.001912), and that the independent variable was marginally nonsignificant (p = 0.0714). Due to the limited sample size available for the blended and distance modes in the study, the findings on the regression analysis for these two modes were inconclusive, and further research involving a larger sample size is needed. These results therefore support the hypothesis presented in Chapter 1.4, that the group of older learners in the study would be less likely to complete the course than the group that included younger learners. Overall, based on the results presented here, the findings suggest that age might explain some of the variation in retention revealed between modes of delivery, and seem to be in line with earlier indications (Khechine et al., 2014) that older individuals in the higher education environment might have more issues with the acceptance of learning in online environments.

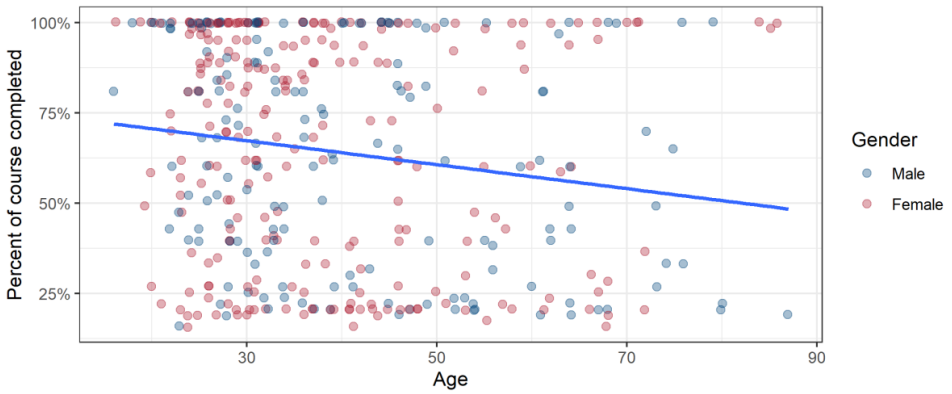


Figure 6. Scatterplot with a regression line representing the correlation of age and the proportion of the course's content covered in IOL 2. Each dot in the figure represents a student, colored according to gender.

With respect to the investigation on whether gender has a predicive value in relation to student retention, the results showed that no relationship existed between the variables gender and retention in the study, as was expected. These findings agree with earlier findings showing that the gender does not seem to have an impact on the attitude towards online learning in higher education (Khechine et al., 2014).

In sum, the main findings indicate that age has a negative predictive value on student retention overall in the survey study, and may therefore be one of the factors that could explain why the blended learning mode of IOL 2, which includes the youngest group of learners (mean age 33.3 years), is more effective in retaining students than the other modes in the study. Furthermore, considering earlier findings in Chapter 3.2.2 that showed that the tutor-specific factors in IOL 2 seemed to have a positive impact on retention in the blended mode of IOL 2 but not in the distance mode, it may be questioned whether the factor of age may come into play in that regard to some extent, since the mean age in the distance mode was highest (43.7 years). As expected, the findings indicate that gender is not a predictor of course completion in the study.

3.2.5 Summary

This section discussed the survey data on learners' views of content-specific, tutor-specific factors, and other motivational factors in IOL 2 and their effect on retention. The findings uncovered various factors affecting student retention in such learning environments. First, the data revealed that the majority of learners in the study considered all six content-

specific factors under investigation to be important elements to engage them with the learning material, in particular factors related to gradual and scaffolded presentation of input and the variety in types of learning objects. When measured against the tracking data, those who considered three out of the six factors to be important for them to engage with the content were more likely to complete the course than those who did not consider them important. The factor of gradual and scaffolded presentation of input was found to have a statistically significant association with course completion in the study. Overall, the results pertaining to content-specific factors in IOL 2 suggest that learners may benefit from engaging forms of pedagogy and design strategies in LMOOCs and that such factors may have a positive impact on their retention.

Second, the investigation showed that the majority of learners in the blended and the distance tutorial modes of IOL 2 believed that all four tutor-specific factors under investigation promoted their engagement with the course. This especially concerned the factors of set syllabus and private interaction with the tutor in both modes. This result suggests that it may be important to provide learners in LMOOCs with personal assistance from a tutor in order to increase the likelihood of continuation in the course. When measured against the tracking data, a positive association was found between all tutored factors and student retention in the blended mode. However, the opposite was the case in the distance mode when investigating the potential impact of these factors on retention. Even though most of the distance learners considered all tutor-specific factors to be important for their motivation in the course, these factors were not found to have a positive impact on course completion in that mode. These findings raise questions as to whether other elements assigned to the modes, or the users, should be taken into consideration in order to explain the variation in learner retention between these two delivery modes.

Third, concerning the factor of initial intent to complete a course and the potential impact of this factor on retention, the results revealed that learners had different goals in mind in terms of course engagement when they came to the IOL 2 course. While the majority was found to have the stated initial goal of completing the course, many either had the goal to cover only a part of it or had no clear goal at all. This result highlights the importance of considering learners' personal goals for engagement when discussing retention in MOOCs. Focusing on the three different modes of delivery, the results revealed that most of the blended and all of the distance learners had the initial intention of completing the course, while less than half of the participants in the open self-directed

mode had this goal. Those who had the goal of completing the course in the open self-directed and blended modes were more likely to do so than the comparison group, who had no such goal; no comparison group was available for the group of distance learners in that regard. The initial goal factor was found to have a statistically significant impact on course completion in the course overall and in the blended and open self-directed modes. Based on the findings that show that most of the blended learners in the study had the initial goal to complete the course and succeeded in doing so, the study addressed the question of whether the initial goal factor might partly explain why learners in the blended mode persisted in the course more so than learners in the other modes.

Finally, concerning the investigation of the influence of age and gender on student retention in IOL 2, the findings showed that age had a negative impact on retention overall in the study: The older a student is, the less likely he or she will be to complete the course. The lowest mean age was found in the blended learning mode, which may be a crucial factor for explaining why the blended mode of delivery had the unique position in the study in terms of retention. Gender, however, was not found to have an impact on retention in the study. The following section discusses the findings on learners' considerations as to what either drove them towards completion of IOL 2 or prevented them from finishing.

3.3 Learners' Reflections on Why They Completed or Withdrew Earlier

This section presents the results of the qualitative data from learners who were enrolled in IOL 2 from 2010 to 2018 and provided written comments on their course engagement. They were thus invited to put into their own words a response to an open-ended question in the survey on why they either completed IOL 2 or left the course without completing. The two groups had stated course completion or non-completion of the course in the survey, depending on sample groups. The former group that had stated course completion received the following fill-in sentence: "When I think about the reason why I completed the course to the end I think about the following ..." The other group that had reported non-completion in the survey, and had also stated the initial intent of completing the course once they started, received this fill-in sentence: "If it turned out that you did not complete the course after all, please write three keywords to describe why you think your initial goal changed."

The objective was therefore to reveal learners' own views in respect to what they see as critical factors that either contributed to or prevented their completion of IOL 2. The data elicitation and analysis were then based on the two following research questions:

1. Why do learners in IOL 2 complete the course? What is the motive from their point-of-view?

2. Why do learners who intend to complete IOL 2 not complete it? What is their primary reason for leaving?

The section is divided into two subsections: First, Chapter 3.3.1 focuses on the reflections of learners who successfully completed the course (Article II), while Chapter 3.3.2 considers perspectives provided by those who left without completing (Article III). It should be noted that the data elicitation and analysis regarding the first research question focused solely in Article II on learners' views with respect to different modes of delivery since that was the subject of the article. Here, however, in the overall presentation of the results of the thesis, the focus is primarily on the overall text data from the same informants, regardless of modes. We first turn to the first question on why learners completed the IOL 2 course and the results.

3.3.1 Reasons for course completion

In order to answer the first research question above, in which learners in IOL 2 were asked why they completed the course, the survey data on stated reasons for completing the course were tabulated depending on the descriptive data. Out of the 123 learners who received this question, 112 learners provided written comments. Patterns identified in the data from all informants in IOL 2, regardless of modes, who had stated course completion in the survey and reflected on the reason why they completed, revealed three leading themes which shed light on the main reason for course completion: a) interesting material; course satisfaction; b) want to learn the language; and c) motivated; interested in the language or culture.

First, the theme 'interesting material; course satisfaction' was identified as the most frequently mentioned reason given for completing the course, reported forty-one times. Second, the theme 'want to learn the language' was reported thirty-three times. Finally, the theme 'motivated; interested in the language or culture' was mentioned twenty-two times. In some instances, more than one theme was recognized in each participant's response. Furthermore, when the data were analyzed in terms of the three different modes in the

study (in Article II), one more theme was identified in the data from informants in the blended learning mode specifically: 'earn a degree', which only applies to this mode. This factor stood out as the most frequently mentioned reason for completion in the blended mode, where the total of 29 learners provided comments, referred to fifteen times. Other less common themes were identified in the data overall, such as 'useful' and 'ease of access or use'.

The following responses reveal various affecting factors identified by the learners in the study and are representative of the most frequently mentioned reasons for course completion for the learners in IOL 2, based on themes. (Abbreviations: Participant = P# (number in the database); female = F; male = M; age = A; country of residence = C; primary school = PS; secondary school = SS; undergraduate = UG; postgraduate = PG; blended learner = B; distance learner = D; open self-directed learner = O.)

Interesting material; course satisfaction

The following quotes from the data emphasize how satisfying course content engaged the learners with the course and supported their language learning, as well as the topics chosen and the instructional strategies used in IOL 2:

> "This course is THE best resource online to learn Icelandic. It has everything: texts with audio, exercises with answers, grammatical explanation, a nice design." (P#59, F, A: 31, C: FR, PG, O)
>
> "Excellent course with a wide variety of language teaching methods well implemented." (P#67, F, A: 30, C: US, UG, O)
>
> "I enjoyed the course tremendously. Having participated in several different types of language learning programs through the years, I found this one to be the most effective, particularly it its explanations of grammar (something which many more popular language-learning tools [...] tend to gloss over almost entirely)." (P#134, M, A: 26, C: US, UG, O)
>
> "I found the course's topics very interesting and was keen to learn more about them." (P#225, F, A: 48, C: DE, UG, O)
>
> "Icelandic is a challenging and beautiful language. The online material provided a fun way to develop an understanding of complex grammar and vocabulary in a fun and even paced manner." (P#351, M, A: 34, C: CA, UG, O)
>
> "I find the activities interesting, educational, and practical." (P#368, F, A: 27, C: PH, UG, B)

> "The course was lots of fun and I really enjoyed it. I also found myself advancing quite quickly thanks to the nature of the exercises, making it fun to learn [...]." (P#389, F, A: 37, C: DK, PG, D)

> "[...] the several ways the course proposes – such as reading texts, listening to dialogues, practicing by completing funny exercises and, last but not least, very clear explicative grammar helps – really represented to me a winning criteria take confidence with Icelandic and eventually learn it." (P#400, F, A: 30, C: IT, SS, B)

Want to learn the language

The responses below highlight how a yearning to master the target language drove learners to complete the course, even multiple times:

> "I want to learn the language and that is not done by an incomplete course." (P#11, F, A: 50, C: SE, UG, O)

> "Because I was going to study in Iceland and thought it would be helpful." (P#12, F, A: 31, C: ES, PG, O)

> "I did it several times actually because I wanted to learn Icelandic. I worked all way through it." (P#78, F, A: 28, C: DE, UG, O)

> "I am a serious learner of the Icelandic language and my goal is C2-level proficiency." (P#237, M, A: 49, C: US, PG, O)

> "I completed the course to the end because I know it would help me to learn if I do so." (P#368, F, A: 27, C: PH, UG, B)

Motivated; interested in the language or culture

The following reports from the learners reveal that some of them were mainly motivated by their sincere interest in the language and the culture:

> "I'm really motivated to learn Icelandic. My love for the language fuelled my desire to learn." (P#59, F, A: 31, C: FR, PG, O)

> "Interest in Icelandic culture and language." (P#96, M, A: 28, C: DE, PG, D)

> "My interest in Icelandic language and hope to follow a summer course at the university in Iceland." (P#158, F, A: 71, C: DK, PG, O)

> "As I do not have to learn Icelandic for some practical reason, it's just my fascination for the country and its language that drives me." (P#344, M, A: 20, C: BE, SS, O)

Finally, what concerns the last aforementioned theme, 'earn a degree', which was identified when the data were analyzed according to different modes of delivery (Article

II), the following quotes reflect the main reason mentioned for course completion by the blended learners in IOL 2.

Earn a degree

As the following quotes from the data indicate, the blended learners are mainly driven towards the end of the course by the awarding of credentials, but also because the course was useful for them:

> "I wanted to improve my Icelandic and get credits." (P#68, F, A: 24, C: CA, UG, B)
>
> "I had to do that in order to pass an exam." (P#157, F, A: 39, C: LT, UG, B)
>
> "I was enrolled at university. I always finish university courses." (P#283, F, A: 44, C: US, UG, B)
>
> "I have finished because it was mandatory to our class but it was also very helpful and intuitive." (P#324, F, A: 37, C: CZ, UG, B)
>
> "Because I had to do it to get the credits for the Practical Diploma in Icelandic." (P#397, F, A: 21, C: CH, UG, B)

To summarize the answer to the first research question in Chapter 3.3 so far in regard to what learners in IOL 2 considered to be the main motive for them to complete the course, the findings show that above all the learners were driven by multiple motivations. This supports the previous discussion of the need to consider student engagement in MOOCs in light of different types of motivation (Chen et al., 2020; Durksen et al., 2016; Ryan & Deci, 2000; Salmon et al., 2017). In relation to the factor of interesting material; course satisfaction, several responses from the informants in this study indicated that IOL's instructional design and methodology may include essential factors in helping language learners to engage with course content and thus aid the language learning process. These relate to the structure and organization of the course content which comprises gradual and scaffolded presentation of input (P#351 above), the strategies used to present and explain grammar in context (P#59; P#134; P#351; P#400), the variety in the tasks throughout the course, which also embodies diversity in teaching methods and practices (P#59; P#67; P#368; P#389; P#400), as well as the topics chosen for the adult language learners (P#225; P#368). Therefore, these findings are in many ways in harmony with the earlier findings presented in this study (Chapter 3.2.1), which showed that the instructional methodology and engagement strategies employed in IOL 2 may be helpful in encouraging language learners and driving them towards course completion. The results are in line with earlier findings that suggest that appealing and enjoyable course content, which affects course

satisfaction, may be one of the crucial factors in retaining students in such learning environments (de Barba et al., 2016; El Said, 2017; Gimeno, 2020; Hew, 2016). These data are considered to provide better understanding of the effectiveness of the course material in question, and underline that elements such as course content and course design may be crucial in engaging the MOOC learners (El Said, 2017; Gimeno, 2020; Hew, 2016; Hone & El Said, 2016; Teixeira & Mota, 2014).

Regarding the factor of wanting to learn the language, the results revealed that some learners are driven towards course completion by the strong desire to develop competence in the target language. As some of the informants' responses indicate, they saw themselves as 'serious language learners' (P#237) and were determined to become skilled users of the target language, and may even have completed the course again and again for that purpose (P#78).

Simlarly, with concern to the factor motivated; interested in the language or culture, the results revealed that some of the learners were fuelled by their attraction to the culture and its language, which kept them going towards the end of the course. This supports earlier findings in research on participants' motivations in MOOCs, which show that many MOOC learners are intrinsically motivated, and how these motivations may positively affect their course engagement and, ultimately, completion (Chen et al., 2020; de Barba et al., 2016; Durksen et al., 2016; Ryan & Deci, 2000; Joo et al., 2018; Salmon et al., 2017).

As the findings also illuminated, the majority of learners in the blended learning mode of IOL 2 who completed the course seemed to be extrinsically motivated and primarily driven towards course completion by the awarding of credentials. These findings support earlier findings that a credit-driven MOOC learning environment could increase the likelihood of completion in MOOCs (El Said, 2017; Joo et al., 2018). Based on previous research, the intention to obtain a certificate of completion of a MOOC is an important indicator of student retention (Greene et al., 2015; Kizilcec, Piech, & Schneider, 2013).

To conclude, the text data analyzed in the present study, provided by users of the program, offer in-depth insight into the experiences of learners who completed the program. The findings on the learners' perspectives towards what motivated them to complete the course first and foremost reflect intrinsic and extrinsic motivations. Therefore, the results overall support the hypothesis presented in Chapter 1.4 that the diverse group of participants in the study would express various motives for completing the course.

3.3.2 Reasons for non-completion of the course

In order to answer the second research question in Chapter 3.3, which asked learners in IOL 2 why they did not complete the course, the survey data on respondents' stated reasons for not finishing the course were tabulated depending on the descriptive data. Only those who had the stated initial goal of completing the course and who had stated non-completion of the course received this question. Of the 107 learners who received the question, 62 provided written comments and each response was explored for common themes. The thematic analysis revealed four leading themes: a) lack of time; b) inappropriate proficiency level; c) still working on the course; and d) lack of motivation.

First, the data analysis illuminated that the theme 'lack of time' was most frequently mentioned reason for non-completion, provided by thirty-eight respondents. Additional comments on this term from many of the respondents exhibited changed priorities in their personal or academic life, while others stated that they still intended to complete the course in due time. Second, the comments pertaining to the theme 'inappropriate proficiency level', reported ten times, illuminated that the course was mostly too difficult. Subsequently, two reasons for not completing, 'still working on the course' and 'lack of motivation', were both referred to nine times. In some instances, a single participant's response identified more than one theme. Other less common themes identified in the data were 'technical issues', 'course organization', 'lack of support or resources', 'uninteresting content', and 'needs fulfilled'.

The following quotes provided by participants illustrate the most prevalent reasons noted for not completing the IOL 2 course based on the above themes:

Lack of time

As the following quotes from the data imply, time constraints were a crucial reason for leaving the course:

> "Lack of time at the end; but planning on to complete it someday." (P#10, M, A: 20, C: DE, SS, O)
>
> "My professional duties prevented participation within the set deadlines." (P#35, M, A: 64, C: DE, PS, D)
>
> "I didn't have enough time due to my studying career, but I plan to continue the course this summer! I want to make all the courses." (P#172, M, A: 22, C: ES, SS, O)

> "Time limitations (I was preoccupied with my works and other tasks). The fact that I can always continue made me lazy and I postponed continuing." (P#311, F, A: 40, C: BG, PG, O)

> "Doing my PHD." (P#349, M, A: 32, C: DE, PG, O)

Inappropriate proficiency level

These quotes indicate that learners may leave the course simply because it is unsuitable for them:

> "[...] it got too hard. I found that to absorb the lesson meant quite a lot of other work – vocabulary, memorizing declensions, checking other resources, and that is way less fun." (P#6, F, A: 44, C: CA, SS, O)

> "It got too difficult with the grammar without it being better explained in smaller steps." (P#30, F, A: 38, C: CH, SS, O)

> "Because it's not really tailored to my needs as an Icelandic learner." (P#70, F, A: 28, C: CA, PG, O)

> "Grammar, vocabulary list confusing, course too long." (P#187, M, A: 23, C: AT, SS, O)

Still working on the course

The following comments reflect that some learners have, in fact, not left the course at all:

> "Still working on it!" (P#51, F, A: 84, C: CA, SS, O)

> "Not yet finished!" (P#118, M, A: 21, C: DE, UG, O)

> "I am still following the course and intend to complete it." (P#146, M, A: 68, C: GB, PG, O)

> "I will finish it, I'm just not done yet." (P#191, F, A: 19, C: DE, UG, O)

Lack of motivation

These reports underline how poor motivation, due to various reasons, affected learners retention:

> "Lack of motivation at the end [...]." (P#176, F, A: 33, C: LV, PG, O)

> "Unmotivated, preoccupied, forgot." (P#242, F, A: 20, C: NO, SS, O)

> "Got too distracted." (P#246, M, A: 18, C: GB, PS, O)

To sum up the answer to the second research question in Chapter 3.3 of what the non-completers in IOL 2 saw as the main cause for leaving the course, the findings first shed light on various reasons as to why committed learners in such learning environments disengage before completing a course, and thus indicate that crucial factors for their leaving may be due to forces external to the course itself, such as insufficient time to focus on their studies. These findings support earlier findings that suggest that diverse external and personal factors may serve as essential barriers for adult learners in such open learning environments (Belanger & Thornton, 2013; Gimeno, 2020; Henderikx et al., 2018; Reich, 2014; Shapiro et al., 2017). In relation to the factor of lack of time, additional statements from the informants that they intended to complete the course in due time indicate the intention to continue the course later (P#10; P#172 above). Similarly, the theme of still working on the course showed that many of the participants who were identified as 'drop-outs' in the study had not left the course at all, but intended to complete it eventually. These remarks may suggest that they view learning in such open online learning environments as an opportunity for ongoing learning (Ingolfsdottir, 2014; Shapiro et al., 2017). Furthermore, responses demonstrating that learners' perception of unsuitable level of study causes drop-out may support previous evidence that lack of prior knowledge of a subject or inadequate background may explain student attrition in MOOCs (Belanger & Thornton, 2013; de Freitas et al., 2015; Shapiro et al., 2017), or merely lack of motivation as when push comes to shove, the learner turns out to be less interested in the course than he or she anticipated (Perna et al., 2014).

In summary, the data analyzed in this study provide valuable insight into various disengagement factors in such learning environments. Overall, these findings support the previous hypothesis presented in Chapter 1.4, that is, that learners who disengage from a MOOC before completing the course would have diverse reasons for leaving, and that these reasons might relate to factors outside of the learning context itself. When considering the findings of the study, it is noteworthy that the main themes discovered in the data did not involve any issues pertaining to course design or lack of communication with fellow students or a tutor, as other similar studies have identified in their research on disengagement factors of MOOCs (Henderikx et al., 2018; Hew & Cheung, 2014; Hone & El Said, 2016). The next section sums up the findings presented above.

3.3.3 Summary

The qualitative part of the study addressed learners who completed IOL 2 on the one hand and those who disengaged before completing on the other, and asked these groups to explain why they continued to the end or left earlier. First, concerning the learners who completed the course, the thematic analysis of their responses uncovered various motives for them to continue to the end. The major conclusions are that appealing and satisfying course content can be a crucial factor in terms of student retention. Many of the informants' comments indicated that IOL's instructional design and methodology may have played a part in their success. The findings also show that the desire to master a target language may drive many learners to the end, and that a genuine interest in the target language itself and its culture encourages many to complete a course. The data from the learners in the blended learning mode in the study, moreover, revealed that they were above all driven by the fact that they were enrolled in a university program which they needed to complete in order to receive the credentials.

Second, based on the responses of the non-completers in the study, the findings show that various individual and outside factors may explain why committed learners leave such courses before completion. The thematic analysis of their responses thus revealed that pressure of time was a crucial reason for leaving the course, while a perceived unsuitable level of study or mere lack of motivation caused attrition among others. The data furthermore showed that several learners who were identified as non-completers in the study had in fact not left the course, but chose to take the advantage of this open learning environment studying at their own pace. In conclusion, the overall findings that have now been revealed emphasize that LMOOC learners may enroll in such courses for many reasons, and are therefore driven by multiple motivations. Furthermore, the resuls showed that various external factors can explain student attrition in such courses.

4. SUMMARY AND CONCLUSIONS

The primary goal of this study was to identify significant determinants of student retention and efficient engagement strategies in open online second language learning courses, or LMOOCs. The investigation was grounded on the IOL program and used mixed research methods to explore the topic. The study first presents empirical evidence of student retention and engagement behavior across seven equivalent courses, of which two are delivered in three different modes: blended, distance and open self-directed. The study identified overall low completion rates in the program, that a blended mode has the highest retention rate in comparison to other modes, and a specific attrition pattern among non-completers, as well as a pattern of user engagements across all courses. Furthermore, based on survey data in correlation with tracking data, the study found that the content-specific and tutor-specific factors explored in one course seemed to have an impact on learner engagement, while only partial evidence of the impact of these factors was found on actual retention. The study moreover found that learners attended the courses with different goals for course engagement in mind, and that the factor of initial intent to complete a course had a significant impact on course completion. Additionally, the results showed that the factor of age has a negative predictive value on student retention, while gender was not found to influence retention. Finally, grounded on qualitative data analysis, learners' reflections on their progression in the program revealed that multiple course-related, motivational, and external factors affected their completion or non-completion of a course.

The overall aim of the research was met by achieving six objectives. The first objective was to reveal tracked student retention in IOL overall in the seven courses and across the three different modes of delivery in IOL 1 and 2, the potential impact of the delivery mode on retention, and to identify what the overall engagement pattern in the program suggests about retention. In order to address these questions, the study used data mining of a large set of tracking data collected through IOL's tracker and LA in order to visualize and understand learners' interactions with the materials and their overall online behavior. The results first provided valuable confirmation of present knowledge of relatively low completion rates in such courses, ranging from 2.4% to 18.2% in the seven courses, and across the different modes of delivery in the program, ranging from 2.9% to 14.2%. Furthermore, when comparing retention between the three different delivery modes of IOL 1 and 2, where each course provides the same online learning material in different

modes, the study found that the blended learning mode of both courses was more effective in keeping learners engaged to the end as compared to the two other modes of the same course. However, relatively little difference in the completion rates was found between the distance and the open self-directed modes of the courses. Third, the data analysis of the online behavior and progress of non-completers in the IOL 1 and 2 courses illuminated a pattern of regular attrition across the courses with sharp initial attrition, particularly in the open self-directed modes, and relatively high attrition peaks at certain junctures throughout the courses, some of which occurred very late, specifically in the distance and blended modes of IOL 2. Furthermore, with the analysis of data on patterns of user engagement across all the seven courses and the three different modes of delivery in IOL 2, which used different parameters for coverage of course content, the study identified various engagement patterns overall throughout the courses. The analysis thus demonstrated that many of the learners in the study who did not finish a course to the very last content page did, however, complete the majority of the content. The nuanced picture of learner retention and progression that was illuminated in the study and the overall engagement patterns throughout the whole program, therefore, called into question what course completion actually means in such courses, and justified a reevaluation of earlier parameters by which retention is measured in the two follow-up studies. Accordingly, by defining course completers as those who completed 80% to 100% of a course's content and by excluding data from those who left a course at the very beginning and thus had little experience using the learning material, this study provides a new perspective and definition of 'course completion' and 'experienced learners'. These findings of the study, therefore, highlight the value of exploring learners' tracked progress and behavior in detail within the context of their learning materials, in order to gain further understanding of student retention in MOOCs, with consideration to both those who cover course content to the end and those who do not.

The second objective of the study was to illuminate whether learners consider specific factors assigned to the structure and organization in IOL, and the design and pedagogical principles important for their motivation to engage with the course, and to reveal a potential impact of these content-specific elements on student retention. This was executed by collecting and analyzing survey data from learners in IOL 2, which were then measured against tracked retention data from the same sample. The findings revealed first that the majority of the learners considered all six content-specific factors in the study to be

important motivators for them to continue with the course. Factors related to gradual and scaffolded presentation of input in the course and the variety in types of learning objects stood out in particular as the most important motivators for learners in the study. The study thus found that the instructional methodology and engagement strategies applied in IOL may be of benefit for the language learner in order to engage him or her with the learning material, and in this way underlines the value of using CALL design within the LMOOC learning environment. Furthermore, the investigation of the potential impact of these content factors on student retention found that three of these factors seemed to have a positive impact on course completion; namely, the elements of curated and sequenced course structure, the gradual and scaffolded presentation of input, and the variety in types of learning objects provided in the course. The other course-related factors investigated in the study, that is, clear and salient learning objectives, form-focused and scaffolded presentation of grammar, and continuing storylines, were not found to have a positive impact on course completion in the study. The additional sub-questions addressed in the survey on learners' usage of other features pertaining to the course content indicated that most of them made use of the available resources provided in the program, such as automated feedback on assignments and the grammar resource.

The third objective of the study was to reveal whether learners in the blended and distance modes of IOL believed that the specific factors related to tutor support and guidance were important to engage them with the course, and whether these factors have an impact on retention. This was performed by collecting and analyzing survey data from learners in the blended and distance modes of IOL 2, which were then correlated with the tracked retention data from the same sample. The study revealed first that the majority of the learners in both modes considered the four tutor-specific factors in the study to be important for their motivation to engage with the course. The factors of set syllabus and private interaction with the tutor seemed to be particularly highly appreciated by the participants in both modes. These findings underscore the potential benefit of the presence and guidance of a tutor in the MOOC learning environment on learner engagement with the material. However, the study found mixed results when the potential impact of these tutor-related factors on retention was observed. All of these factors were found to have a positive impact on retention for the blended learners, while the same factors did not seem to influence learner retention in the distance course. The special position of the distance learners that was discovered in this context, therefore, raised further questions. As became

evident in the investigation of the tracking data on retention in the IOL program both overall and across the different modes of delivery, little difference was found in the completion rates between the tutorial distance mode and the non-tutorial open self-directed mode of IOL 2. These overall findings thus raised the question of whether other elements than the tutor-related factors in these modes play a decisive role in student retention, and might therefore explain why the blended mode is more effective in retaining learners than the distance mode.

The fourth objective of the study was to illuminate whether learners overall in IOL 2, including the three different modes of delivery, had an initial intent of completing the course once they started, and whether such a goal had an impact on student retention. The factor of the initial goal to complete the course was specifically considered in order to shed some light on possible reasons why the blended mode was more efficient in retaining learners than the other modes in the study. This was executed by collecting and analyzing survey data in IOL 2, both overall and by mode of delivery, which were then correlated with the tracked retention data from the same sample. The findings revealed first that diverse learners attend the course with a variety of goals in mind. While some seemed determined to complete the course when they entered, others seemed to have planned to cover only a part of it or were unsure about their intentions. This result stresses the importance of considering learners' initial intended participation as relates to MOOC retention. Furthermore, the results showed that the majority of learners in IOL 2 overall, regardless of mode, had the stated initial goal of completing the course, and that this factor had a statistically significant impact on student retention overall in the study. On the other hand, when focusing on this factor with respect to the three different modes of delivery in the course, the study revealed that most of the blended learners and all of the distance learners had the stated initial intent of completing the course, while less than half of the open self-directed learners had this goal. The intial goal factor was found to have a statistically significant impact on retention in the blended and the open self-directed modes, while the distance mode did not allow for comparison. Based on the data that showed that the majority of blended learners in IOL 2 intended to complete the course and were successful in doing so, these findings were considered to partially explain why the blended learning mode was more efficient in retaining learners than the other modes in the study.

The fifth objective of the study was to uncover learners' own views as to what they see as critical factors that either contributed to their completion of IOL 2 or prevented them from completing. This was carried out by eliciting qualitative data through open question sections in the survey, where those who had stated course completion were asked to reflect on the reason why they completed, and those who had the stated initial intent of completing and had reported non-completion of the course were asked to give thought to why they did not complete the course after all. The analysis of the text data discovered several themes across the two sets of data. Primarily, regarding the informants who completed the course, the study found that course completers in such learning context seemed to be driven by multiple motivations. First, several responses from these informants indicated that factors involved with the course content of IOL 2 and the engagement strategies used may be crucial factors of engaging learners – where appealing and satisfying content which is clearly structured and organized, and the input is gradually presented and scaffolded – seem to play an important role for the learners to persist in the course. Furthermore, the study discovered that the instructional strategies applied in IOL 2 to present and explain grammar in context, and the variety in the tasks and topics provided in the course may also be significant determinants of student retention in the course. Moreover, the data illuminated that many learners may be driven towards the end of a course by the desire to develop proficiency in the target language, while others are motivated by a genuine interest in the topic in hand. However, when this dataset was analyzed specifically according to mode of delivery in IOL 2, the study found that the promise of obtaining a credential in the blended mode was a notable motivating factor in driving learners in this mode towards course completion. These findings thus suggest that the option of receiving credentials may be one of the major factors that have an impact on retention in the blended mode, and may thus explain the unique position of blended learners in the study.

Second, regarding the analysis of the text data on the reasons why learners in IOL 2 who intended to complete the course did not succeed in doing so, the analysis of their responses found that critical factors for learners to leave a course may be mostly attributable to elements outside of the course itself. Hence, the results revealed that time constraints were a crucial reason for leaving for many learners, and that a perceived unsuitable level of study caused others to drop out. In yet other cases, learners may have dropped out due simply to lack of incentive. The data also indicate that many of these

learners had not left the program at all, but choose instead to take the time needed to complete the course or intend to come back to it and complete it in due time. All in all, when observing the qualitative data provided in this study, the findings highlight that diverse LMOOC learners attend such courses with various motives, objectives and needs which must be considered in the discussion on student retention.

Finally, the sixth objective of the study was to identify whether the age and gender of the learners had an impact on student retention in IOL 2. This was done by analyzing and correlating the registration data and retention data from the learners in this course. Regression analysis thus showed a statistically significant relationship between age and retention and that age had a negative impact on student retention, while no significant relationship was found between gender and retention. The study also revealed that the group of learners in the blended mode had the lowest mean age, while the highest mean age was found in the distance mode. These findings therefore suggest that the factor of age may somewhat explain the difference in retention rates between the groups of blended learners and distance learners in the study.

While the first part of the study, involving a large set of tracking data, contributes new empirical data on overall retention and engagement behavior in seven self-directed open online language courses and in different modes of delivery, and thus provides a better understanding of student retention and engagement in such learning environments, there are some limitations that must be acknowledged. First, as previously mentioned, the IOL tracking system used in the study does not distinguish registration data from user data. This means that once a user enrolls in IOL he or she is automatically positioned as a learner on the first content page (# 111) of all courses, regardless of which course or mode the user may interact with, and regardless of whether the student performed any action on the page. Because of this element of uncertainty, page 111 was excluded from data analysis for all the courses, including different delivery modes, explaining missing data n = 96,473. Another shortcoming to this part of the study is that the tracker notifies a user's 'last position', which reveals where the user was situated when he or she left the course. This may result in some uncertainty, for example if a student were to complete a course but then re-entered it later, possibly in the middle of the course, in order to review material. In such cases, his or her status as a completer of the course would be lost, and the learner's 'last position' would be registered as in the middle of the course, leaving the user classified in the data as a non-completer. The new and upgraded version of the IOL system that has

now been implemented for use on mobile devices with a different tracking system opens up an opportunity for future studies in this area; Not only does it distinguish registration data from user data, it also tracks learners' engagement with each learning object in a course rather than each content page, and therefore records students' behaviors more precisely. It would thus be necessary to apply the new tracking system for further investigation on the attrition patterns revealed in this study in close connection with the course content, in order to develop further understanding of engaging or disengaging factors in such a learning environment. The new system also has the advantage that it enables graphs and visualizations for research to be performed by non-data specialists, which is not the case for the tracking system used in the present study.

Second, while the survey study contributes new self-reported data in correlation with tracking data on the impact of content-specific factors, tutor-specific factors, the factor of an initial intent of completing a course, and other motivational and external factors on student retention in a beginner course, there are some limitations that must be mentioned. The research sample reflected a limited population of enrollees in IOL 2, representing 15.4% of participants in the period under investigation. The low rate of response may be due to various causes, such as the diverse sample and different backgrounds of the participants, the fact that they may not have been engaged in the course for a while, and the possibility that some learners may have been unable to take part because the survey was in English. The fact that this study used a post-course survey may also have affected the response rate. In that regard, however, it must be noted that not only the most engaged and successful students completed the questionnaire; 44% of respondents completed IOL 2 while 56% did not, and proportionately more of the respondents took IOL 2 recently (2016–2017). In the future it may be advisable to embed the questionnaires at different levels of the course, and even consider surveys with fewer questions, to promote survey participation. Furthermore, since participants in the study did not report their initial intent in terms of course engagement prior to the course, it cannot be ruled out that some were confirming or describing the behavior they had already displayed in the course. Another significant limitation of the study, which concerns the investigation of the tutor-specific factors in the three different modes of delivery in IOL 2, is that a total of 58 learners were discovered who showed incorrect responses when asked to mark which mode of delivery they studied. Due to this element of uncertainty, this group of learners was excluded from the data analysis, which led to a small sample size for the blended and

distance learners when analyzing data on tutor-specific factors in the survey study. These limitations consequently affected the qualitative data analysis concerning the informants who reported on the reason why they completed IOL 2, since these data were primarily analyzed by mode of delivery in Article II. Hence, comments from 28 informants across the three different delivery modes were excluded in that part of the study. It may be essential in future studies to send out different questionnaires aimed at learners in each mode of delivery. Finally, regarding the findings on the influence of the content-specific factors in IOL 2, it should be mentioned that the results may reflect an age bias, since participants in the study represent perspectives of various age groups where instructional needs may differ according to age. This study provides ground for further investigation of the content-specific factors in the IOL program in terms of different age groups and of engagement strategies in CALL in general, since the new IOL tracking system gives an opportunity to measure the effect of different types of learning objects on student engagement and retention in more detail, such as what types learners engage with the most, or whether they seem to be struggling with certain types of learning objects, which may influence student retention in such courses. When considering the survey study overall, it must be acknowledged that the selection of the study's sample may have affected the results since the answers to questions of experience and meaning also relate to users' social affiliations. In the future it would be advisable to explore a broader sample in the attempt to capture the many facets of the phenomenon. It must also be kept in mind that the overall findings presented here on the impact of modes of delivery and of the tutorial factors on retention may be somewhat biased, since the blended mode of IOL 1 and 2 is the only mode in the study that provides credits for completion. A possible area for future research could be to measure the level of retention based on content of the three different modes of delivery. In relation to this, it also must be considered that the participants' self-selected choice of mode could have reflected various initial motives, and thus affected the results of the study.

Third, while this study applied a common method of qualitative data elicitation in SLA research by using an open-ended questionnaire, and thus provided valuable insight into what the language learners in IOL 2 see as critical factors of retention or attrition in such learning environments, there are some limitations that need to be acknowledged. In future studies, it would be necessary to draw upon multiple methods and sources of qualitative data, such as interviews, in order to provide a more detailed and nuanced

picture of the setting and thus strengthen the validity of interpretations in this study. As concerns the text data from the informants in the study who described the main reasons for their course completion, it should also be mentioned that the findings may be somewhat biased. Therefore, it cannot be ruled out that some of the closed-ended questions in the survey, which addressed questions on specific course-related elements, may in some way have affected responses in the survey's free writing form. Finally this research project offers the possibility of replication studies in different settings in the future, as the IOL program has been reproduced and made available for other languages.

Collectively, this research study contributes wide-ranging data from different sources on the engagement behavior and perspectives of LMOOC learners in the context of their learning materials. First, with the investigation of large-scale tracking data collected from a large number of students in several equivalent courses, of which some are delivered in different delivery modes, the study provides new empirical evidence of engagement behaviors and progress throughout an entire program, from the time students start a course until they leave. Through the mining of these data and the use of LA, the study also contributes unique evidence of users' involvement with every single content page in a course, and thus provides a detailed picture of users' engagement with the learning materials. Furthermore, while the retention data were broken down in the attempt to reveal when the cohort of learners who did not fully cover the course content disengaged from the program, the study offers valuble evidence of patterns of student attrition. The study therefore provides evidence of when during a course learners may be at risk of disengaging from the program. By providing such data on the timing of drop-outs, the study not only confirms existing knowledge that students in MOOCs commonly drop out early on in these courses, but also that learners may disengage towards the end of a course. Moreover, by investigating the overall engagement patterns of students in a program with the use of different parameters for course completion, focusing specifically on those who do not complete a course to the end, this study contributes more detailed data on how much of a course's content these learners do cover. Accordingly, course completion was defined in the follow-up studies as 80% to 100% completion of course content. Based on these findings, this study reevaluated the previous frameworks that measure students' attendance in MOOCs. Overall, considering the analysis of the tracking data in this research, one of the main implications of the study is that LMOOC developers and instructors may benefit greatly from a built-in tracking device that allows the recording and analysis of important

information about learners' behavior in such courses. Through data visualization and the identification of attrition patterns among users, which show that attrition happens early and late in courses, this study also suggests that the use of a tracking system may benefit instructors in LMOOCs in the sense that it allows them to intervene and make an effort to increase the persistence of their students. Based on the results of the tracking data research, this study moreoever suggests new ways of measuring retention in MOOCs, and highlights that completion rate alone may not be an appropriate metric for student engagement and retention in MOOCs.

Second, through the use of a follow-up survey study on learning content and tutor support whose data correlated with the tracking data, this study not only contributes new data on language learners' perspectives and their experiences of the effectiveness of such learning environments for their engagement, but also provides insights into the actual influence of these factors on retention. As the results highlight the potential crucial importance of engaging design strategies and pedagogy on encouraging the learning process, the study indicates that LMOOCs require a platform specifically aimed at teaching and learning a language. This study offers a new framework for promoting student engagement and suggests specific strategies that involve thoroughly structured and organized content, scaffolding of input and instruction, and multiform learning context for educators and LMOOC developers. Considering, on the other hand, the results of the data on learners' experiences with tutored factors, the findings suggest that supervision throughout the course and the opportunity to seek help and clairification on a topic may be crucial to encourage LMOOC learners in their learning. The framework that this study offers, including engaging instructional strategies and the availability of a tutor may, therefore, provide a useful guide for educators and developers of LMOOC courses. This framework may also direct them in finding additional solutions to further support the LMOOC learners in order to improve their experience and persistence in such courses. While the current study investigated the impact of content factors and tutor factors individually, future research could also entail a regression analysis to describe the relationships between these variables and retention, in order to make predictions about significant elements of retention in LMOOCs. Furthermore, as this study contributes new information about the value for the autonomous beginner language learner of using instructional resources in CALL involving SLA pedagogical approaches, and of receiving support from a tutor, the study suggests that future research should consider whether

advanced LMOOC learners, who may also represent users of various age groups and geographic location, benefit from such framework of instruction.

Finally, by capturing learners' own thoughts on the reasons why they completed or did not complete a course, the study has contributed broad individual perspectives on critical factors of LMOOC retention. While the text data analysis confirms existing knowledge of the diversity of the MOOC learners who attend these courses with diverse motives and different goals of participation, it has offered in-depth insight into the experiences of LMOOC learners who have completed a course to the end, and sheds light on how various intrinsic and extrinsic motivations may promote student engagement and completion of a course. Based on the findings of the study that indicate that many learners seem to be mainly driven towards course completion by the credentials offered in some courses, this study suggests that using a LMOOC as a learning opportunity within the context of a university-accredited program would increase the likelihood of completion. Furthermore, as the study provides clarity to the present understanding of how diverse individual and circumstantial obstacles may prevent MOOC learners from using a course, this study has also emphasized the importance of paying attention to the resourceful informants in such environment who do not fully complete a course. Even though non-course-related factors may play a significant role in learners' withdrawal from MOOCs, these findings may guide course developers in finding suitable redesign solutions to better engage learners. Since this study has explored the perspectives of those who fulfilled their goal of completing a course, as well as those who had the initial goal of completing a course but did not succeed in doing so, this study suggests that future research should consider the perspectives of learners who had no intention of completing a course when they enrolled but ended up doing so, which may further illuminate engaging factors in LMOOCs.

The research project presented has identified a range of significant determinants of student retention, and has broadened and extended related literature in the emerging LMOOC field. The study provides a new framework for how to promote student engagement in LMOOCs and informs design directions for course developers and online instructors in higher education.

References

Adamopoulos, P. (2013). What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. *Proceedings of the 34th international conference on information systems*. https://aisel.aisnet.org/icis2013/proceedings/BreakthroughIdeas/13

Aebersold, J. A., & Field, M. L. (1997). *From reader to reading teacher*: *Issues and strategies for second language classrooms*. Cambridge: Cambridge University Press.

Aldowah, H., Al-Samarraie, H., Alzahrani, A. I., & Alalwan, N. (2020). Factors affecting student dropout in MOOCs: a cause and effect decision-making model. *Journal of Computing in Higher Education*, *32*, 429–454. https://doi.org/10.1007/s12528-019-09241-y

Arnbjörnsdóttir, B. (2004). Teaching morphologically complex languages online: Theoretical questions and practical answers. In P. J. Henrichsen (Ed.), *CALL for the Nordic languages. Tools and methods for computer assisted language learning* (pp. 79–94). Copenhagen: Samfundslitteratur.

Arnbjörnsdóttir, B. (2007). Kenningar um tileinkun og nám annars máls og erlendra mála [Theories on second language acquisition and learning]. In A. Hauksdóttir, & B. Arnbjörnsdóttir (Eds.), *Mál málanna. Um nám og kennslu erlendra mála* (pp. 13–47). Reykjavík: Vigdís Finnbogadóttir Institute of Foreign Languages.

Arnbjörnsdóttir, B. (2008). Kennsla tungumála á netinu: Hugmyndafræði og þróun Icelandic Online [Teaching languages online: The methodology and development of Icelandic Online]. In A. S. Þráinsdóttir, B. Hafstað, & B. Sigurðsson (Eds.), *Hrafnaþing, 5* (pp. 7–31). Reykjavík: University of Iceland.

Arnbjörnsdóttir, B., Friðriksdóttir, K., & Bédi, B. (2020). Icelandic Online: twenty years of development, evaluation and expansion of an LMOOC. In K.-M. Frederiksen, S. Larsen, L. Bradley, & S. Thouësny (Eds.), *CALL for widening participation: short papers from EUROCALL 2020* (pp. 13–19). Research-publishing.net. https://doi.org/10.14705/rpnet.2020.48.1158

Bárcena, E., & Martín-Monje, E. (2014). Introduction. Language MOOCs: an emerging field. In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs. Providing learning, transcending boundaries* (pp. 1–15). Warsaw: De Gruyter Open. https://doi.org/10.2478/9783110420067.1

Beaven, T., Codreanu, T., & Creuzé, A. (2014). Motivation in a language MOOC: Issues for course designers. In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs. Providing learning, transcending boundaries* (pp. 48–66). https://doi.org/10.2478/9783110420067.4

Belanger, Y., & Thornton, J. (2013). *Bioelectricity: A quantitative approach*. *Duke University's first MOOC*, 1–21. https://hdl.handle.net/10161/6216

Bettinger, E. P., Fox, L., Loeb, S., &, Taylor, E. S. (2017). Virtual classrooms: How online college courses affect student success. *American Economic Review*, *107*(9), 2855–2875. https://doi.org/10.1257/aer.20151193

Blin, F. (2016). The theory of affordances. In C. Caws, & M-J Hamel (Eds.), *Language-learner computer interactions. Theory, methodology and CALL applications* (pp. 41–64). Amsterdam/Philadelphia: John Benjamins.

Bruner, J. S. (1974). *Beyond the information given: Studies in the psychology of knowing*. New York: Norton.

Castrillo, M. D. (2014). Language teaching in MOOCs: the integral role of the instuctor. In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs. Providing learning, transcending boundaries* (pp. 67–90). Warsaw: De Gruyter Open. https://doi.org/10.2478/9783110420067.5

Chapelle, C. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning & Technology*, *2*(1), 22–34. https://eric.ed.gov/?id=EJ609869

Chapelle, C. A. (2003). *English language learning and technology*: *Lectures on applied linguistics in the age of information and communication technology*. Amsterdam: John Benjamins.

Chapelle, C. A. (2009). The relationship between Second Language Acquisition Theory and computer-assisted language learning. *The Modern Language Journal*, *93*. https://doi.org/10.1111/j.1540-4781.2009.00970.x

Chen, C., Sonnert, G., Sadler, P. M., Sasselov, D. D., Fredericks, C., & Malan, D. J. (2020). Going over the cliff: MOOC dropout behavior at chapter transition. *Distance Education*, *41*(1), 6–25. https://doi.org/10.1080/01587919.2020.1724772

Chun, D. M. (2001). L2 reading on the Web: Strategies for accessing information in hypermedia. *Computer Assisted Language Learning*, *14*(5), 367–403.

Chun, D. M. (2012). Replication studies in CALL research, *CALICO Journal*, 29(4), 591–600. https://doi.org/10.1558/cj.29.4.591-600

Chun, D. M. (2013). Contributions of tracking user behavior to SLA research. *CALICO Journal*, *30*(2), 256–262. https://doi.org/10.1558/cj.v30i0.256-262

Chun, D. M. (2016). The role of technology in SLA research. *Language Learning & Technology*, *20*(2), 98–115. http://dx.doi.org/10125/44463

Colpaert, J. (2006). Pedagogy-driven design for online language teaching and learning. *CALICO Journal*, *23*(3), 477–497.

Colpaert, J. (2010). Elicitation of language learners' personal goals as design concepts. *Innovation in Language Learning and Teaching*, *4*(3), 259–274. https://doi.org/10.1080/17501229.2010.513447

Colpaert, J. (2014). Conclusion. Reflections on present and future: towards an ontological approach to LMOOCs. In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs. Providing learning, transcending boundaries* (pp. 161–172). Warsaw: De Gruyter Open. https://doi.org/10.2478/9783110420067.10

Colpaert, J. (2018). Transdisciplinarity revisited. *Computer Assisted Language Learning*, *31*(5–6), 483–489. https://doi.org/10.1080/09588221.2018.1437111

Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment*. Strasbourg: Language Policy Unit.

Creswell, J. W. (2006). Understanding mixed methods research. http://www.sagepub.com/upm-data/10981_Chapter_1.pdf

Creswell, J. W. (2013). *Qualitative inquiry and research design. Choosing among five approaches* (3rd ed.) Thousand Oaks: SAGE.

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning, 32*(3), 218–231. https://doi.org/10.1111/jcal.12130

de Freitas, S. I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology, 46*(3), 455–471. https://doi.org/10.1111/bjet.12268

Doiz, A., Lasagabaster, D., & Sierra, J. M. (2014). Giving voice to the students. What (de)motivates them in CLIL classes? In D. Lasagabaster, A. Doiz, & J. M. Sierra (Eds.), *Motivation and foreign language learning: From theory to practice* (pp. 117–138). Amsterdam: John Benjamins.

Doughty, C. & Williams, J. (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Duff, P. A. (2012). How to carry out case study research. In A. Mackey, & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 95–116). New York: Wiley-Blackwell.

Durksen, T. L., Chu, M-W., Ahmad, Z. F., Radil, A. I., & Daniels, L. M. (2016). Motivation in a MOOC: a probabilistic analysis of online learners' basic psychological needs. *Social Psychology of Education*, *19*, 241–260. https://link.springer.com/article/10.1007/s11218-015-9331-9

Dörnyei, Z., & Csizér, K. (2012). How to design and analyze surveys in second language acquisition research. In A. Mackey, & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 74–94). New York: Wiley-Blackwell.

Dörnyei, Z., Muir, C., & Ibrahim, Z. (2014). Directed motivational currents. Energising language learning by creating intense motivational pathways. In D. Lasagabaster, A. Doiz, & J. M. Sierra (Eds.), *Motivation and foreign language learning: From theory to practice* (pp. 9–29). Amsterdam: John Benjamins.

Ebben, M., & Murphy, J. S. (2014). Unpacking MOOC scholarly discourse: a review of nascent MOOC scholarship. *Learning, Media and Technology*, *39*(3), 328–345. https://doi.org/10.1080/17439884.2013.878352

El Said, G. R. (2017). Understanding how learners use massive open online courses and why they drop out: Thematic analysis of an interview study in a developing country. *Journal of Educational Computing Research, 55*(5), 724–752. https://doi.org/10.1177/0735633116681302

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, *35*(3), 407–432. https://doi.org/10.2307/3588029

Fischer, R. (2007). How do we know what students are actually doing? Monitoring students' behavior in CALL. *Computer Assisted Language Learning*, *20*(5), 409–442. https://doi.org/10.1080/09588220701746013

Friðriksdóttir, K. (2008). Íslenska sem annað mál: Hvernig lærist fallbeyging nafnorða í íslensku? [The acquisition of noun case in L2 Icelandic]. In A. S. Þráinsdóttir, B. Hafstað, & B. Sigurðsson (Eds.), *Hrafnaþing*, *5* (pp. 33–53). Reykjavík: University of Iceland.

Friðriksdóttir, K. (2015). Styðjandi námsumhverfi Icelandic Online [Icelandic Online – enhancing vocabulary L2 learning and reading comprehension]. In T. Sigurðardóttir, & M. Garðarsdóttir (Eds.), *Frændafundur*, *8* (pp. 45–74). Tórshavn: University of the Faroe Islands.

Friðriksdóttir, K. (2018). The impact of different modalities on student retention and overall engagement patterns in open online courses. *Computer Assisted Language Learning*, *31*(1–2), 53–71. https://doi.org/10.1080/09588221.2017.1381129

Friðriksdóttir, K. (2019). The effect of tutor-specific and other motivational factors on student retention on Icelandic Online. *Computer Assisted Language Learning*, 1–22. Advance online publication. https://doi.org/10.1080/09588221.2019.1633357

Friðriksdóttir, K. (2021). The effect of content-related and external factors on student retention in LMOOCs. *ReCALL*, *33* (2), 128–142. https://doi.org/10.1017/S0958344021000069

Friedman, D. A. (2012). How to collect and analyze qualitative data. In A. Mackey, & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 180–200). New York: Wiley-Blackwell.

Frydenberg, J. (2007). Persistence in university continuing education online classes. *International Review of Research in Open and Distance Learning*, *8*(3), 1–15. https://doi.org/10.19173/irrodl.v8i3.375

Gaebel, M. (2014). MOOCs: Massive open online courses. An update of EUA's first paper (January 2013). EUA occasional papers. *European University Association*, 1–35. https://eric.ed.gov/?id=ED571272

Garrett, N. (1991). Technology in the service of language learning: Trends and issues. *The Modern Language Journal*, *75*(1), 74–101. https://doi.org/10.1111/j.1540-4781.1991.tb01085.x

Garrett, N (2009). Computer-assisted language learning trends and issues revisited: Integrating innovation. *The Modern Language Journal*, *93*(s1), 719–740. https://doi.org/10.1111/j.1540-4781.2009.00969.x

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education, 7*(2), 95–105. https://doi.org/10.1016/j.iheduc.2004.02.001

Gelan, A., Fastré, G., Verjans, M., Martin, N., Janssenswillen, G., Creemers, M., ... Lieben, J. (2018). Affordances and limitations of learning analytics for computer-assisted language learning: a case study of the VITAL project. *Computer Assisted Language Learning*, *31*(3), 294–319. https://doi.org/10.1080/09588221.2017.1418382

Gillespie, J. (2020). CALL research: Where are we now? *ReCALL*, *32*(2), 127–144. https://doi.org/10.1017/S0958344020000051

Gimeno, A. (2020, October). *Analysing Learner Motivation.* Paper presented at The Future of Education International Conference, 1–5. https://www.researchgate.net/publication/351558903_Analysing_Learner_Motivation

Godwin-Jones, R. (2011). Emerging technologies. Autonomous language learning. *Language Learning & Technology*, *15*(3), 4–11. http://llt.msu.edu/issues/october2011/emerging.pdf

Godwin-Jones, R. (2017). Emerging technologies. Scaling up and zooming in: Big data and personalization in language learning. *Language Learning & Technology*, *21*(1), 4–15. https://core.ac.uk/reader/211322178

Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015). Predictors of retention and achievement in a massive open online course. *American Educational Research Journal*, *52*(5), 925–955. https://journals.sagepub.com/doi/10.3102/0002831215584621

Guest, G. (2012). Describing mixed methods research: An alternative to typologies. *Journal of Mixed Methods Research*, *7*(2), 141–151. https://doi.org/10.1177/1558689812461179

Harker, M., & Koutsantoni, D. (2005). Can it be as effective? Distance versus blended learning in a web-based EAP programme. *ReCALL 17*(2), 197–216. https://doi.org/10.1017/S095834400500042X

Heift, T. (2016). Learner personas and the effects of instructional scaffolding on working behaviour and linguistic performance. In C. Caws, & M-J Hamel (Eds.), *Language-learner computer interactions: Theory, methodology and CALL applications* (pp. 117–136). Amsterdam/Philadelphia: John Benjamins.

Henderikx, M., Kreijns, K., & Kalz, M. (2017). Refining success and dropout in massive open online courses based on the intention-behavior gap. *Distance Education*, *38*(3), 353–368. https://doi.org/10.1080/01587919.2017.1369006

Henderikx, M., Kreijns, K., & Kalz, M. (2018). A classification of barriers that influence intention achievement in MOOCs. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *Lifelong technology-enhanced learning. Proceedings of the 13th European conference on technology enhanced learning 2018*, 3–15. https://doi.org/10.1007/978-3-319-98572-5_1

Henderikx, M., Kreijns, K., Muñoz, J. C., & Kalz, M. (2019). Factors influencing the pursuit of personal learning goals in MOOCs. *Distance Education*, *40*(2), 187–204. https://doi.org/10.1080/01587919.2019.1600364

Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCS. *British Journal of Educational Technology, 47*(2), 320–341. https://doi.org/10.1111/bjet.12235

Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review*, *12*, 45–58. https://doi.org/10.1016/j.edurev.2014.05.001

Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education, 98*, 157–168. https://doi.org/10.1016/j.compedu.2016.03.016

Hsieh, H-F, & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Hubbard, P. (2012). *Curation for systematization of authentic content for autonomous learning*. Talk delivererd at EUROCALL 2012 conference, Gothenburg. http://www.stanford.edu/~efs/eurocall2012.pdf

Hubbard, P. (2013). Digital content curation for CALL. *Glasgow World CALL 2013: Global perspectives on computer-assisted language learning.* Glasgow, 10–13 July.

Höfler, E., Zimmermann, C., & Ebner, M. (2017). A case study on narrative structures in instructional MOOC designs. *Journal of Research in Innovative Teaching & Learning*, *10*(1), 48–62. https://doi.org/10.1108/JRIT-09-2016-0005

IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp.

Ihantola, P., Fronza, I., Mikkonen, T., Noponen, M., & Hellas, A. (2020). Deadlines of MOOCs: How do students behave in MOOCs with and without deadlines. *IEEE Frontiers in Education Conference*, 1–9. https://doi.org/10.1109/FIE44824.2020.9274023

Ingolfsdottir, K. (2014). Impact of MOOCs and other forms of online education. Point of view. *Proceedings of the Institute of Electrical and Electronics Engineers* (IEEE), *102*(11), 1639–1643. https://doi.org/10.1109/JPROC.2014.2360025

Jeong, H. J. & Lee, W. C. (2016). The level of collapse we are allowed: comparison of different response scales in safety attitudes questionnaire. *Biometrics & Biostatistics International Journal*, *4*(4): 128–134. https://doi.org/10.15406/bbij.2016.04.00100

Joo, Y. J., So, H-J, & Kim, N. H. (2018). Examination of relationships among students' self-determination, technology acceptance, satisfaction, and continuance intention to use K-MOOCs. *Computers & Education*, *122*, 260–272. https://doi.org/10.1016/j.compedu.2018.01.003

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *International Review of Research in Open and Distributed Learning*, *15*(1), 133–160. http://www.irrodl.org/index.php/irrodl/article/view/1651

Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distributed Learning*, *16*(3), 341–358. https://doi.org/10.19173/irrodl.v16i3.2112

Kern, R., Ware, P., & Warschauer, M. (2004). Crossing frontiers: New directions in online pedagogy and research. *Annual Review of Applied Linguistics*, *24*, 243–260.

Khechine, H., Lakhal, S., Pascot, D., & Bytha, A. (2014). UTAUT model for blended learning: The role of gender and age in the intention to use webinars. *Interdisciplinary Journal of E-Learning and Learning Objects*, *10*, 33–52. https://doi.org/10.28945/1994

Kim, T., Yang, M., Bae, J., Min, B., Lee, I., & Kim, J. (2017). Escape from infinite freedom: Effects of constraining user freedom on the prevention of dropout in an online learning context. *Computers in Human Behavior*, *66*, 217–231. https://doi.org/10.1016/j.chb.2016.09.019

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. *Proceedings of the third international conference on learning analytics and knowledge*, pp. 170–179. https://doi.org/10.1145/2460296.2460330

Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction*, *22*(2), 1–24. https://doi.org/10.1145/2699735

Koller, D., Ng, A., & Chen, Z. (2013). Retention and intention in massive open online courses: In depth. *EDUCAUSE Review*. https://er.educause.edu/articles/2013/6/retention-and-intention-in-massive-open-online-courses-in-depth

Lantolf, J. (2000). *Sociocultural theory and second language learning*. New York: Oxford University Press.

Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education*, *29*, 40–48. https://doi.org/10.1016/j.iheduc.2015.12.003

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity. International Journal of Methodology*, *43*, 265–275. https://doi.org/10.1007/s11135-007-9105-3

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, *48*(2), 185–204. https://doi.org/10.1016/j.compedu.2004.12.004

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 31–40. http://er.educause.edu/~/media/files/article-downloads/erm1151.pdf

Martín-Monje, E., Castrillo, M. D., & Mañana-Rodríguez, J. (2018). Understanding online interaction in language MOOCs through learning analytics. *Computer Assisted Language Learning*, *31*(3), 251–272. https://doi.org/10.1080/09588221.2017.1378237

Namey, E., Greg, G., Thairu, L., & Johnson, L. (2008). Data reduction techniques for large qualitative data sets. In G. Guest, & K. M. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 137–161). Lanham, MA: Altamira Press.

Oshlyansky, L., Cairns, P., & Thimbleby, H. (2007). Validating the Unified Theory of Acceptance and Use of Technology (UTAUT) tool cross-culturally. In D. Ramduny-Ellis, & D. Rachovides (Eds.), *British Computer Society*, *2*, 83–86. Proceedings of the 21st British HCI Group annual conference. Lancaster University. https://doi.org/10.14236/ewic/HCI2007.67

Patterson, B., & McFadden, C. (2009). Attrition in online and campus degree programs. *Online Journal of Distance Learning Administration*, *12*(2), 1–11. https://www.learntechlib.org/j/OJDLA/v/12/n/2/

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in massive open online courses. *Educational Researcher*, *43*(9), 421–432. https://doi.org/10.3102/0013189X14562423

Qualtrics Labs (2012). *Qualtrics survey software: Handbook for research professionals* (2nd ed.). Qualtrics Labs.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Reich, J. (2014). MOOC completion and retention in the context of student intent. *EDUCAUSE Review*, *49*(6). https://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent

Rodriguez, C. O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance and E-Learning*, 1–13. https://eric.ed.gov/?id=ej982976

Rosenshine, B., & Meister, C. (1992). The use of scaffolds for teaching higher-level cognitive strategies. *Educational Leadership*, *49*, 26–33. http://www.formapex.com/telechargementpublic/rosenshine1992a.pdf

Ross, J., Sinclair, C., Knox, J., Bayne, S., & Macleod, H. (2014). Teacher experiences and academic identity: The missing components of MOOC pedagogy. *Journal of Online Learning and Teaching*, *10*(1), 57–69. https://jolt.merlot.org/vol10no1/ross_0314.pdf

Rubio, F., Thomas, J. M., & Li, Q. (2018). The role of teaching presence and student participation in Spanish blended courses. *Computer Assisted Language Learning*, *31*(3), 226–250. https://doi.org/10.1080/09588221.2017.1372481

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Salmon, G., Pechenkina, E., Chase, A-M., & Ross, B. (2017). Designing massive open online courses to take account of participant motivations and expectations. *British Journal of Educational Technology*, *48*(6), 1284–1294. https://doi.org/10.1111/bjet.12497

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226.

Schulze, M., & and Scholz, K. (2016). CALL theory: Complex adaptive systems. In C. Caws, & M-J Hamel (Eds.), *Language-learner computer interactions. Theory, methodology and CALL applications* (pp. 65–87). Amsterdam/Philadelphia: John Benjamins.

Shapiro, H. B., Lee, C. H., Roth, N. E. W., Li, K., Çetinkaya-Rundel, M., & Canelas, D. A. (2017). Understanding the massive open online course (MOOC) student experience: An examination of attitudes, motivations, and barriers. *Computers & Education*, *110*, 35–50. https://doi.org/10.1016/j.compedu.2017.03.003

Sokolik, M. (2014). What constitutes an effective language MOOC? In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs. Providing learning, transcending boundaries* (pp. 16-32). Warsaw: De Gruyter Open. https://doi.org/10.2478/9783110420067.2

Taylor, S. J., & Bogdan, R. (1998). *Introduction to qualitative research methods. A guidebook and resource* (3rd ed.) (pp. 134–163). New York: Wiley.

Teixeira, A. M., & Mota, J. (2014). A proposal for the methodological design of collaborative language MOOCs. In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs. Providing learning, transcending boundaries* (pp. 33–47). Warsaw: De Gruyter Open. https://doi.org/10.2478/9783110420067.3

Thomas, M., & Gelan, A. (2018). Special edition on language learning and learning analytics. *Computer Assisted Language Learning*, *31*(3), 181–184. https://doi.org/10.1080/09588221.2018.1447723

Vorobel, O., & Kim, D. (2012). Language teaching at a distance: An overview of research. *CALICO Journal*, *29*(3), 548–562. https://doi.org/10.11139/cj.29.3.548-562

Watson, W. R., Yu, J. H., & Watson, S. L. (2018). Perceived attitudinal learning in a self-paced versus fixed-schedule MOOC. *Educational Media International*, *55*(2), 170–181. https://doi.org/10.1080/09523987.2018.1484044

Wang, Y., & Baker, R. (2015). Content or platform: Why do students complete MOOCs? *Journal of Online Learning and Teaching*, *11*(1), 17–30.

Wang, Y-S., Wu, M-C., & Wang, H-Y. (2009). Investigating the determinants and age and gender differences in the acceptance of mobile learning. *British Journal of Educational Technology*, *40*(1), 92–118. https://doi.org/10.1111/j.1467-8535.2007.00809.x

Wiebe, E., Thompson, I., & Behrend, T. (2015). MOOCs from the viewpoint of the learner: A response to Perna et al. (2014). *Educational Researcher*, *44*(4), 252–254. https://doi.org/10.3102/0013189X15584774

Þorvaldsdóttir, S. & Garðarsdóttir, M. (2015). Fall eða setningarstaða? Þróun falla í máltileinkun. [The development of case assignment in Icelandic as a second language.] In Sigurðardóttir, T. & Garðarsdóttir, M. (eds.), *Frændafundur, 8* (pp. 135–158). Tórshavn: University of the Faroe Islands.

Original Publications