מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Thesis for the degree

עבודת גמר (לתזה) לתואר

Doctor of Philosophy

דוקטור לפילוסופיה

Submitted to the Scientific Council of the Weizmann Institute of Science

מוגשת למועצה המדעית של מכון ויצמן

Rehovot, Israel

רחובות, ישראל

By:

Yael Feldman-Maggor

מאת:

יעל פלדמן-מגור

ארגון הוראה ולמידה אקדמית בסביבה טכנולוגית עתירת מידע: קורסי כימיה למורים וסטודנטים

Organizing learning and teaching in online information-rich environment: Higher education chemistry courses for teachers and students

Advisors:

Prof. Ron Blonder

Prof. Inbal Tuvi-Arad

מנחות:

פרופ' רון בלונדר

פרופ' ענבל טובי-ערד

January 2022

שבט תשפ"ב

תודות

לימודי הדוקטורט היו עבורי מסע מרתק, מאתגר, ומעניין הכרוך בלמידה מתמשכת. עבודת הדוקטורט היוותה נקודת מפנה בחיי בה חזרתי לארץ לאחר כמה שנים בהן התגוררתי עם משפחתי בארצות הברית. המחקר ליווה אותי בתקופה המרגשת של גידול ילדי הקטנים והאהובים, תמר ולביא. מרבית הזמן בשנים האחרונות הוקדש להם ולמחקר.

הגשת העבודה מאפשרת לי הזדמנות יוצאת דופן להודות לכל אלו שבלעדיהם, כתיבתה לא הייתה אפשרית.

ראשית, ברצוני להודות לשתי המנחות שלי, פרופ' ענבל טובי-ערד ופרופ' רון בלונדר על תמיכתן ואמונתן בי לאורך כל תהליך המחקר. פרופ' ענבל טובי-ערד מלווה אותי כבר מעל לעשור, עוד מתחילת החשיבה על תזה לתואר שני ויש לה השפעה רבה על התפתחות והמיומנות המקצועית שלי. למדתי ממנה על חשיבות הירידה לפרטים הקטנים ואיך להסתכל על התקדמות המחקר והאתגרים שבדרך באופן אופטימי ככל האפשר. לאחר סיום התואר השני היא הסכימה ללוות אותי גם בתהליך המשמעותי של מחקר הדוקטורט. במהלך תקופת התואר השני נחשפתי לראשונה למחקריה החשובים של פרופ' רון בלונדר והתרגשתי שהסכימה להנחות אותי בדוקטורט. הייתה לי הזכות לגלות ולהכיר אישיות מיוחדת במינה שהיוותה עבורי דוגמה אישית לאורך כל התקופה. הנחייתה של רון, יחד עם הזדמנויות שונות בהן שלחה אותי להציג את המחקר, העלו את תחושת המסוגלות העצמית שלי כחוקרת מתחילה. שתי המנחות נפגשו איתי ללא הגבלה, ייעצו, העירו, קידמו, עודדו ולימדו אותי רבות על תחום הוראת הכימיה, על ביצוע המחקר, ושלבי הכתיבה, והכל בסבלנות ונינוחות רבה. זכיתי לעבוד עם שתי מנחות שהיו שם בשבילי לא רק בהקשר המקצועי אלא גם בכל עניין אישי שעלה בו הצורך לתמיכה והקשבה. אני שמחה על הזכות שנפלה בחלקי ללמוד תחת הנחייתן, ובטוחה שההשפעה של שנים אלו תלווה אותי בהמשך החיים.

מחקר הדוקטורט כלל כמות נתונים מרובה וסבוכה ומי שהיה שותף לדרך המורכבת הכרוכה בהכנתם לניתוח הוא שגיב ברהום ועל כך אני מודה לו באופן מיוחד על הקדשת שעות מרובות לעבודה משותפת. תודה מיוחדת נוספת מגיעה גם לעדנה רולניק, הגרפיקאית המוכשרת במחלקה להוראת המדעים שעבדה איתי בסבלנות על איורים רבים להצגות בכנסים שרובם שולבו בעבודה זו.

תודה גדולה לשני חברי הוועדה המלווה: פרופ' מיכל ארמוני וד"ר ארנון הרשקוביץ על קריאה מעמיקה, שאלות חכמות, הארות, הערות ותובנות שעזרו לי לקדם ולשפר את העבודה במהלך שנות המחקר. עבודת הדוקטורט נשענה רבות על מחקרים קודמים של ד"ר ענת כהן ועמיתיה ואני מודה לה על קריאה מעמיקה בסוף התהליך.

תודה נוספת לד"ר דורותה צ'רקי וד"ר דינה עינות-יוגב שפתחו בפני את הקורסים שהן מרכזות באוניברסיטה הפתוחה למחקר וההערכה. תודה גם לכל המורים והסטודנטים שהסכימו להתראיין למטרת המחקר אך יאלצו להישאר בעילום שם משיקולי שמירה על פרטיות.

ברצוני להודות למדרשת פיינברג במכון ויצמן, קרן המחקר של האוניברסיטה, משרד הקליטה וללשכת המדען הראשי במשרד החינוך שתמכו במחקר זה באמצעות מלגות נדיבות שאפשרו לי להקדיש את זמני לכתיבה ומחקר במשך כחמש שנים. תודה גם לקרן פרויקט i-PEN (Innovative Photonics Education in Nanotechnology) על התמיכה במהלך המחקר. כמו-כן תודה לסטיב מאנץ' על העריכה הלשונית המעמיקה של עבודה זו.

תודה לחברי הסגל במחלקה להוראת המדעים ובמיוחד לד"ר גיורא אלכסנדרון שהוביל אותי לחשוב על מספר נקודות חשובות עוד בתחילת הדרך וגם בסוף תהליך המחקר. לפרופ' ניר אוריון ופרופ' דיויד פורטוס שהלמידה בקורסים שלהם השפיע רבות על האופן בו עוצב המחקר בעבודות דוקטורט.

תודה לשותפים לדרך וחברי היקרים למשרד הסטודנטים: אהוד אבירן, אלה יונאי, אינאס עיסא, טל הירש שמח, עאישה סינדיאני בסול ואיציק ארוך. עם חלקם זכיתי לצבור חוויות בכנסים אקדמיים באיטליה ובגרמניה עוד לפני תחילת מגפת הקורונה וחלקם הצטרפו ללוות אותי בשלבי סיום הדוקטורט. כולם תרמו לאווירה טובה שהפכה את החוויה להרבה יותר נעימה. תודה גם על ההקשבה והדיון בממצאי המחקר כבר בשלבים מוקדמים.

תודה מיוחד לחברתי הטובה במחלקה, ד"ר שלי רפ, על שיחות מקצועיות ואישיות. תודה מקרב לב לקבוצת הכימיה המורחבת שתמכה והתעניינה לאורך כל הדרך. מאחר ותקופת הדוקטורט אפשרה לי גם ללמוד לתעודת הוראה בכימיה זכיתי להכיר את הקבוצה אף יותר לעומק וברצוני להודות במיוחד לד"ר רחל ממלוק-נעמן, ד"ר אינה שוורץ-סברו, שרה אקונס, ד"ר רות וולדמן, ד"ר מלכה יאון, ד"ר דבורה קטוביץ, ד"ר דבורה (דידי) מרצ'ק, ד"ר שלי ליבנה, זיוה בר דב, ענבר חיימוביץ, ד"ר שרון גלר, ורד אדלר וחגית לוי.

תודה לכל אותם חברים וקולגות שייעצו בנושאי הסטטיסטיקה ושיטות המחקר לאורך הדרך: יטי ורון, טניה נזרצקי, ד"ר סער קרפ גרשון, שי פרח, עמית לזרוס, ד"ר מוריה אריאלי, ד"ר עידית פסט וד"ר ארז מרנץ. תודה מיוחדת גם לד"ר רחל אידלמן, וד"ר גיל שוורץ על החברות שהתחילה בארגון כנס טכנולוגיות למידה ונמשכה במהלך הדרך.

תודות נוספות:
ליתר חברי מחלקת הגרפיקה במכון ויצמן בעבר ובהווה: ציפי עובדיה, אבי טל וזיו אריאלי.
לצוות שהם באוניברסיטה הפתוחה, במיוחד לאיתי הר-אבן, תום קלס וירון הלפר.
לצוות הטכנולוגי במכון ויצמן ובמיוחד למרינה ארמיאץ', סטלה חזינה, מתן ברקוביץ', עדי פויארקוב.
לצוות האדמיניסטרטיבי, ובמיוחד עדי שמרון ממכון ויצמן, ורונה אלון ועינב רוזנשטיין מהאוניברסיטה הפתוחה.

ברצוני להודות למשפחתי התומכת ובמיוחד לאמי האהובה חנה פלדמן שתמיד האמינה בי, עודדה אותי ואף סייעה רבות עם הילדים והבישולים. תודה גם לחמי, אלי מגור, שתמיד התעניין בשלבי העבודה השונים ועזר באופן קבוע עם הילדים. תודה גם לאחי היקר, יאיר פלדמן על תמיכה והתעניינות, לדודתי האהובה, אסתי כהן. תודה גם לדודי, ד"ר מוטי כהן, שהיה הראשון שלימד אותי מה הם הרגלי למידה נכונים עוד כשהייתי בתיכון. היות ועבודה זו עוסקת בדרכי למידה לעובדה זו השפעה משמעותית גם על המחקר הנוכחי. תודה גם לגיסיי, ד"ר נועם מגור, רבקה מגור, ושירה מגור ובנות דודי מירב כהן ולילך כהן. היכולת לחקור ולכתוב עבודה מסוג זה התאפשרה גם הודות לתמיכתם המשפחתית.

תודה נוספת לכמה מחברותי הטובות ששמרו על קשר לאורך תקופה עמוסה זו: רותם חסון, דנה מזרחי, אדוה בן-יעיש, שונמית חרות, יפעת וינשטיין, תמר וייס-ארגמן, גבי חרן וילנר, נורית לביא, ד"ר לינור מולר ודנה שדות.

חשוב לי להודות גם לאנשים שלא זכו הגיע לשלב זה, אך השפיעו על החלטתי לבחור במסלול האקדמי ולעסוק במחקר המשלב טכנולוגיה: אבי, קונרד פלדמן ז"ל, שאמנם לא זכה ללוות אותי בבגרותי אך עוד בילדות תמיד דאג להיות מראשוני משתמשי הטכנולוגיה וחשף אותי לתחום באופן עקיף. תודה לסבי היקר, צבי וייס ז"ל, שהעניק לי את התמיכה שאפשרה לי להגיע ללימודים גבוהים. עבודה זו מוקדשת לזכרו. תודה גם לסבתי, דבורה וייס ז"ל, סבתי, אווה פלדמן ז"ל, וסבתו של ארז בעלי, רחל גורנברג ז"ל שצפתה את הדרך הזו עוד בשיחות הראשונות שלנו.

יותר מכולם אני מוקירה תודה עמוקה לבן זוגי, בעלי היקר והאהוב והשותף לדרך, ארז מגור, שתמיד מאמין, תומך, מעודד ומאפשר. העשייה והכתיבה של המחקר הנוכחי התאפשר בזכות השלווה המשפחתית. במהלך השנים האחרונות ארז היה עסוק בעצמו בכתיבת הדוקטורט שלו ועדין תמיד ידע לשלב, להתפשר, להקשיב לרעיונות, לממצאים חדשים, לקרוא חלקים מרובים מהעבודה. הכל תמיד בנועם, חום, סבלנות, ובשילוב התלהבות המאפיינת את אישיותו המיוחדת.

Declaration

I hereby declare that this thesis summarizes my independent and original research

Yael Feldman-Maggor

Table of Contents

Abstract

This dissertation examines chemistry teachers' and undergraduate university students' learning patterns in online chemistry courses. Online learning is not a new phenomenon; however, it has gained momentum in the Internet age, which accelerated in the wake of the Covid-19 pandemic. On the one hand, the advantage of online courses is that students can learn from anywhere, at any time. On the other hand, studies indicate that students' completion rate is lower in online courses than in face-to-face learning. The main goals of this study are to identify learning patterns that can predict students' successful completion of online chemistry courses and develop tools for evaluating online courses, using the theoretical frameworks of engagement and self-regulated learning. Self-regulated learning can be defined as the learners' ability to act independently, be active, and manage their learning process. Self-regulated learning is essential in all forms of learning, but it is of even greater importance in online learning, given its flexibility.

The study was conducted in two stages: The first stage is descriptive; its purpose is to characterize the learners according to their learning patterns in the online learning environment. This stage relied on qualitative and quantitative research methods. The second stage of the study, the prediction stage, relied mainly on quantitative methods. Key study findings include two models designed to determine whether learners will complete the course.

The study was based on data collected from chemistry courses given at two different academic institutions: the Open University and the Weizmann Institute of Science. The Open University data spanned seven cycles of online chemistry courses over three academic years (2017-2020), in which a total of 954 students were enrolled. The Weizmann Institute data spanned three online course semester-long cycles over three academic years (2016-2019), in which 95 teachers were enrolled for professional development purposes. The descriptive stage, in which the learning patterns were characterized, was based on interviews with participants registered in any one of those courses. In addition to the interviews, the teachers' learning patterns were also analyzed based on the reflective summary and their final course assignments.

The analysis was underpinned by several characteristics of self-regulated learning theory: goal setting, the learning environment, learning strategies, time management, help-seeking, and self-evaluation. From the analysis of the Weizmann Institute course,

we learned about new learning patterns that we presented through five case studies. These patterns include, for example, continuous learning from week to week, completing the course in intervals, or completing the course all at once: "binge-watching" the course.

Next, the qualitative analysis learning patterns in several chemistry courses were analyzed using log files extracted from the Moodle learning management system. These log files are reports detailing learners' various actions on the course website without compromising their privacy. The data in these reports include (but are not limited to) dates on which each learner was active, the learners' number of visits to the website, and whether, when, and how many times they accessed the course activities. In addition, the research dataset includes demographic information and data on the learners' achievements, which, together with their online activity data, provide a holistic picture of the learners' characteristics.

Quantitative data analysis using EDM methods is a complex process. When we began to receive the log files and combine them with the demographic and academic achievement data, we found that the raw data were not suitable for direct analysis. Rather, they required preliminary processing and testing. The methodology chapter describes the method we developed to manage and undertake the initial processing of the data collected. This method includes four main stages: data gathering, data interpretation, database creation, and data organization – where each stage consists of several sub-stages. The development and use of this method revealed that early pre-processing of the data could prevent considerable inaccuracies in the research findings and significantly strengthen the reliability of the resulting conclusions.

The descriptive (first) stage laid the foundation for the analysis stage, in which we identified various parameters that contribute either to successful completion or to non-completion of the course. In the second stage of the research, from the analysis of the Open University courses' log files, we constructed two logistic regression models aimed to identify unique variables that can predict whether the course will be successfully completed. The models indicate that two factors are strong predictors for completing the course: (i) the submission status of the first optional assignment in week 5 of the course; and (ii) the students' cumulative video opening pattern (SCOP) by week 7. The logistic regression model we applied in the analysis relating to the Weizmann Institute's courses indicates that students' cumulative video opening pattern by week 5 is a strong predictor of course completion.

At the Weizmann Institute, where we studied the “Introduction to Materials and Nanotechnology” online course for teachers PD, we also evaluated the learning outcomes and difficulties. These factors facilitated meeting another goal of the research: developing evaluation tools for online courses for teachers’ professional development. To this end, we developed a framework that combines educational data mining methods with traditional evaluation tools. This grouping leads to a multi-dimension evaluation framework that considers: 1) knowledge, 2) the complexity of learners’ understanding, and 3) identifying learners’ struggles. The first facet was evaluated using the pre-post knowledge questionnaires, the second using the structure of observed learning outcomes (SOLO) taxonomy to analyze the course assignments, and the third by analyzing the online Moodle log files, together with semi-structured interviews. This multi-dimension evaluation tool allowed us to assess how teachers have expanded their knowledge and skills in subjects that are not part of the high-school science curriculum. By examining the teachers’ learning patterns in the online video lessons, we identified the more challenging topics resulting in course non-completion.

This study has potential applications for researchers, lecturers, and learners. Our qualitative analysis can be used to develop and update existing SRL questionnaires to make them more relevant for evaluating learning in online settings. Our quantitative analysis, in particular, the models we developed, can improve learning evaluation already in the middle of the course rather than only at the end. These models also make it possible to design future intervention research strategies. As for learners, we wish to emphasize the importance of developing their self-regulated learning and to show how their learning process choices affect their potential to successfully complete the course.

Abstract In Hebrew

תקציר

מחקר זה עוסק בלומדים המשתתפים בקורסים מקוונים בכימיה. למידה מרחוק קיימת כבר זמן רב, אך צברה תאוצה עם התפתחות רשת האינטרנט ואף ביתר שאת בעקבות התפרצות מגפת הקורונה. מחד, יתרונם של קורסים מקוונים הוא האפשרות ללמוד מכל מקום ובכל זמן. מאידך, מחקרים עכשוויים מראים שאחוז ההצלחה בקורסים מקוונים נמוך יותר ביחס לקורסים המעוברים פנים מול פנים. מטרותיו העיקריות של מחקר זה הן זיהוי דפוסי למידה שיאפשרו לחזות סיום בהצלחה של קורסי כימיה מקוונים ופיתוח כלים להערכת קורסים מקוונים תוך שימוש במסגרת התאורטית של הכוונה עצמית בלמידה. הכוונה עצמית היא היכולת של הלומד לפעול באופן עצמאי להיות פעיל ולנהל את תהליך הלמידה שלו. הכוונה עצמית חשובה בכל סוג של למידה אך בלמידה המקוונת לאור הגמישות המתאפשרת ללומדים היא חשובה על אחת כמה וכמה.

המחקר בוצע בשני שלבים. השלב הראשון הינו שלב תיאורי ומטרתו לאפיין את הלומדים מבחינת דפוסי הלמידה באתר הקורס. שלב זה התבסס על מתודולוגיה איכותנית וכמותנית. השלב השני של המחקר, שלב הניבוי הסתמך בעיקרו על שיטה כמותנית. ממצאים מרכזיים של שלב הניבוי הינם שני מודלים שפותחו במטרה לחזות את סטאטוס סיום הקורס בקרב הלומדים בשלבים מוקדמים של הקורס.

המחקר התבסס על נתונים שנאספו במסגרת קורסי כימיה שניתנו בשני מוסדות לימוד: האוניברסיטה הפתוחה (שמונה מחזורים של שלושה קורסי כימיה מקוונים מהשנים האקדמיות 2017-2020 שכללו 954 סטודנטים בסה"כ) ומכון ויצמן למדע (שלושה מחזורי קורס מקוון מהשנים האקדמיות 2016-2019 שכללו 95 מורים בסה"כ). השלב התיאורי בו אופיינו דפוסי הלמידה התבסס על ראיונות עם סטודנטים הלומדים בקורסי האוניברסיטה הפתוחה ומורים שהשתתפו בקורס לפיתוח מקצועי שהועבר במכון ויצמן. ניתוח הראיונות התבסס על מספר מאפיינים של תאוריית ההכוונה העצמית בלמידה: הצבת מטרות, סביבת הלמידה, אסטרטגיות למידה, ניהול זמן, פנייה לעזרה והערכה עצמית. בנוסף לראיונות, ניתוח דפוסי הלמידה של המורים במכון ויצמן התבסס גם על סיכום אישי רפלקטיבי, והעבודה המסכמת של הקורס. מהניתוח האיכותני שנערך בהקשר לקורס לפיתוח מקצועי שניתן במכון ויצמן למדע למדנו על דפוסי למידה חדשים אותם הצגנו באמצעות חמישה תיאורי מקרה. דפוסים אלו כללו לדוגמה, למידה רציפה משבוע לשבוע, השלמת הקורס בחלקים, או השלמת הקורס בבת אחת במסגרת צפיית בינג'.

השלב הראשון כלל גם ניתוח כמותני בו המחקר התבסס בעיקר על שיטת מחקר של כריית נתונים. במסגרת שיטה זו נותחו דפוסי הלמידה של לומדים במספר קורסים בכימיה באמצעות אנליזה של קבצי יומן שהופקו ממערכת ניהול למידה מסוג מודל (Moodle). קבצי יומן הם דוחות המפרטים את הפעולות השונות שעשו הלומדים באתר הקורס תוך שמירה על פרטיות הלומד. הם כוללים, בין היתר, נתונים על מועד הפעילות, זמן צפייה בווידאו, מספר כניסות לאתר. נתונים אלו מאפשרים לנו להבין למשל את

תדירות השימוש במשאבי הלמידה השונים. בנוסף לנתוני הפעילות המתוקשבת נתוני המחקר כוללים נתונים דמוגרפיים והישגיים המאפשרים לקבל תמונה כוללת על מאפייני הלומדים.

ניתוח נתונים כמותניים באמצעות כריית נתונים הינו תהליך מורכב. עם התחלת קבלת הנתונים מקובץ היומן ושילובם עם נתוני הדמוגרפיה והישגים הלימודיים, נמצא כי הנתונים הגולמיים המתקבלים אינם מתאימים לניתוח ישיר ומצריכים עיבוד ובדיקות מקדימות. בפרק המתודולוגיה מתוארים שלבי העבודה שפותחו במהלך הדוקטורט לצורך ניהול וטיפול מקדים במידע הנאסף. אלו כוללים ארבעה שלבים עיקריים: איסוף הנתונים, פרשנות הנתונים, בניית מסד הנתונים וארגון הנתונים כאשר כל שלב מורכב ממספר תתי שלבים. מפיתוח שלבי עבודה אלו והשימוש בהם נמצא כי עיבוד מוקדם של הנתונים יכול למנוע אי-דיוקים גדולים בממצאי המחקר, ולחזק באופן משמעותי את מהימנות המסקנות.

כאמור, השלב התיאורי אפשר לאפיין דפוסי למידה. בהתבסס על השלב התיאורי מיקדנו את השלב השני של המחקר. באוניברסיטה הפתוחה הגדרנו בשלב השני שני פרמטרים עיקריים: הגשת מטלות בחירה ודפוסי פתיחה מצטברים של מפגשים/הקלטות הוידאו. דפוס זה הינו משתנה שפותח במסגרת המחקר במטרה להעריך את קצב ההתקדמות של לומדים בקורס מקוון. פרמטרים אלו שימשו לבניית שני מודלים של רגרסיה לוגיסטית שמטרתם לזהות משתנים ייחודיים המאפשרים לחזות את סיום הקורס בהצלחה. המודלים מראים כי הן סטטוס ההגשה של מטלת הבחירה בשבוע החמישי והן דפוס פתיחת הווידאו המצטבר של הסטודנטים בשבוע השביעי משמשים כמנבאים חזקים לסיום הקורס בהצלחה. גם בקורס מבוא לחומרים וננוטכנולוגיה במכון ויצמן למדע נעשה שימוש ברגרסיה לוגיסטית שמראה כי דפוס פתיחת הווידאו המצטבר של המורים בשבוע החמישי מהווה מנבא חזק לסיום הקורס בהצלחה.

בקורס מבוא לחומרים וננוטכנולוגיה במכון ויצמן למדע התמקדנו גם בהערכת הקורס מבחינת למידת התוכן הכימי וההתקדמות ברכישת הלמידה לאורך הקורס. הלומדים ענו על שאלון ידע לפני ואחרי הקורס, התבקשו לקשר את הנלמד לתוכנית הלימודים והגישו משימה מסכמת. ההערכה זו אפשרה לענות על מטרה נוספת של המחקר והיא פיתוח כלי הערכה של קורסים מקוונים. המסגרת שפותחה משלבת את הערכת הפעילויות המקוונות עם כלי הערכה מסורתיים. שילוב זה אפשר לבנות כלי להערכה רב ממדית הכוללת: 1) ידע; 2) מורכבות ההבנה של הלומדים; ו-3) זיהוי קשיי הלומדים. הערכת ההיבט הראשון נעשתה באמצעות שאלון ידע לפני-אחרי, השני באמצעות יישום של טקסונומית הסולו (SOLO- Structure of Observed Learning Outcomes) והשלישי על ידי ראיונות מובנים למחצה וניתוח דוחות פעילות מקוונים. באמצעות כלי הערכה הרב-ממדי ניתן להראות כיצד מורים הרחיבו את הידע והכישורים שלהם בנושא שאינו חלק מתוכנית הלימודים למדעים בבית הספר התיכון. משימות הקורס מנחות את המורים להציע דרך משלהם לחבר את התכנים המתקדמים שהם לומדים בקורס לתוכנית הלימודים בכימיה במהלך הלמידה בקורס. מממצאים אלו ניתן ללמוד על התרומה הפוטנציאלית של טקסונומיית הסולו ככלי עיצוב קורס המאפשר לספק ללומדים הכוונה שתסייע להם

להשיג רמה גבוהה יותר של מורכבות בלמידה. על ידי בחינת דפוסי פתיחת שיעורי הווידאו המקוונים של המורים, זוהו נושאים מאתגרים יותר שעלולים למנוע מהלומדים להשלים את הקורס.

למחקר זה יש יישומים פוטנציאליים עבור חוקרים, מרצים ולומדים. ניתן להשתמש בניתוח האיכותני שלנו כדי לפתח ולעדכן שאלוני הכוונה עצמית קיימים כדי להפוך אותם לרלוונטיים יותר ללמידה בסביבה מקוונות. הניתוח הכמותי, בפרט, המודלים שפותחו, יכולים לשפר את הערכת הלמידה כבר באמצע הקורס ולא רק בסופו. מודלים אלו מאפשרים גם לעצב אסטרטגיות מחקר התערבות עתידיות. באשר ללומדים, אנו רוצים להדגיש את החשיבות של פיתוח הלמידה בהכוונה עצמית שלהם ולהראות כיצד הבחירות של הלומדים בתהליך הלמידה משפיעות על הפוטנציאל שלהם לסיים את קורס המקוון בהצלחה.

List of Abbreviations

- **AUC** - Area Under the Curve
- **EDM** - Educational Data Mining
- **EM** - Electron Microscopy
- **GC** - General Chemistry
- **GCA** - General Chemistry A
- **GDPR** - General Data Protection Regulation
- **IRB** - Institutional Review Board
- **LA** - Learning Analytics
- **LMS** - Learning Management System
- **MOOC** - Massive Open Online Course
- **NLP** - Natural Language Processing
- **NST** - Nanoscale Science and Technology
- **OSLQ** - Online Self-Regulated Learning Questionnaire
- **OUI** - Open University of Israel
- **PD** - Professional Development
- **ROC** - Receiver Operating Characteristics
- **SAT** - Scholastic Aptitude Test
- **SCOP** - Student Cumulative Opening Pattern
- **SA/V** - Surface Area to Volume
- **SEM** - Scanning Electron Microscopy
- **SES** - Socio-Economic Status
- **SOLO** - Structure of the Observed Learning Outcome
- **SPM** - Scanning Probe Microscopy
- **SPOC** - Small Private Online Course
- **SRL -** Self-Regulated Learning
- **TEM** - Transmission Electron Microscopy
- **WOC** - The World of Chemistry

List of Tables

List of Figures

1. Introduction

1.1. Rationale

This study focuses on online chemistry courses for teachers' professional development and for undergraduate students. In recent years, online learning has become a widely popular educational platform. The online environment provides a convenient format for adult learners because its time flexibility and accessibility match well with adults' preferences for independent learning. Despite their many advantages, online learning courses also pose several challenges. For instance, the loss of face-to-face interaction creates obstacles for effective learning.

Furthermore, completion rates are notably lower in comparison to more traditional face-to-face courses (Levy., 2007; Onah et al., 2014; Shea, & Bidjerano, 2019; Cohen et al., 2019). The low completion rate results for several reasons, for example, the low level of self-efficacy, the low level of motivation, and the lack of self-regulated learning skills (Cohen et al., 2019; Davis et al, 2018; Kizilcec et al., 2017; Watted & Barak, 2018). The main goals of this study are to advance the development of evaluations tools and to identify learning patterns that can predict success in an online chemistry course. In this regard, we focus mainly on the importance of self-regulated learning, using a mixed-method approach that combines both qualitative and quantitative tools.

1.2 Introduction and Literature Review

The following literature review elaborates on the main topics relevant to this study: online learning, educational data mining, and self-regulated learning. More specific topics are discussed within the relevant chapters (chapters 3-5).

1.1.1 Online Learning and Educational Data Mining

In the early 2000s, when online videos first emerged, and even more so around 2005, when the shared video website YouTube emerged, online videos became easily accessible and available to all (Kay, 2012; Read & Lancaster, 2012). Consequently, online learning has become a conventional mode of learning in higher education (You, 2016). This, in turn, has led to a significant rise in online distance learning offered by universities, such as online lectures and homework assignments with integrated video footage, quizzes, and social-network discussion forums (Hershkovitz & Nachmias, 2011; Johnson et al., 2014).

Many online courses are designed and built using a learning management system (LMS) that functions as the course's learning website. Although the LMS has become an essential part of any online environment, it also complements traditional learning. Some LMSs are closed systems that are marketed to academic institutions and schools without the ability to make any changes. Other LMSs function as open sources and enable institutions to make changes according to their specific needs (Islam, 2014). An LMS facilitates the delivery of highly informative courses that include diverse learning resources such as presentations, quizzes, videos, and an open online forum for student discussions. Therefore, an LMS helps encourage lecturers to enrich their courses with digital resources (Islam, 2012).

When students interact with an LMS, many parameters about their activities are automatically gathered and stored in log-file data (Baker & Inventado, 2014). This includes, for example, timestamps of each activity, counts of entrances to the website or specific files, and content added by users (e.g., in a forum). This process produces a massive amount of data that is impossible to analyze manually (Romero & Ventura, 2020). Two central research communities have developed with a joint interest in how such educational data can be exploited to contribute to the education system and to learning science. These communities developed methods known as "Educational Data Mining" (EDM) and "Learning Analytics" (LA) (Romero & Ventura, 2020).

EDM is a method for exploring unique types of data that originate from an educational setting (Luna et al., 2017). LA is a method for data measurement, collection, and analysis about learners and their learning context (Siemens and Long 2011). Both EDM and LA share the goal of improving the quality of educational data (Luna et al., 2017). However, several key differences exist between the two communities of researchers. EDM is an emerging interdisciplinary research field that leverages new computational approaches and explores data originating in educational contexts to address academic questions (Romero & Ventura, 2010). Siemens & Baker (2012) explained that researchers in the EDM community focus on automated methods to discover specific elements within educational data and that they aim to model specific constructs and the relationships between them.

On the other hand, researchers in the LA community typically emphasize attempting to understand systems as a whole in their full complexity. Gibson and Ifenthaler (2017) claim that researchers from both communities need to be equipped with a new set of fundamental competencies required for computationally intensive research, such as data management techniques and working with interdisciplinary teams that understand programming languages and learning theory. Regardless of the differences between the EDM and the LA research communities, the two significantly overlap regarding the investigators' objectives as well as the methods and techniques used in the study (Romero & Ventura, 2020). For convenience, we will mainly use the term EDM in this dissertation when discussing our methods and analysis.

Analysis of learning behavior using EDM methods can provide significant insights into the design of learning environments and can support decisions about the future management of educational resources (Hershkovitz & Nachmias, 2011; Gibson & Ifenthaler, 2017; Miah et al., 2020). However, this type of exploration requires complex strategies that begin with pre-processing the raw data into a suitable format ready for analysis (Angeli et al., 2017; Kapusta et al., 2019; Liñán & Pérez, 2015; Zacharis, 2015). An analytical project will require accessing, cleaning, integrating, analyzing, and visualizing data before attempting to make sense of the situation (Siemens, 2013).

In the pre-processing data phase, researchers need to consider the complexity of the educational dataset collected. Then, they should be able to identify those attributes that have missing values (Dutt et al., 2017; Gupta & Sabitha, 2019). Romero et al. (2014) noted various technical challenges in data gathering and defined several procedures for collecting and integrating data from multiple sources. They also explicated data cleaning and removing outliers, dimensionality reduction, and finally filtering and transformation. However, most published studies do not include a detailed description of these stages of data pre-processing and usually only list the variables on which they focus without providing further details (Romero et al., 2014). This occurs even though researchers often devote 60%-90% of their time to pre-processing the data (Ramírez-Gallego et al., 2017; Romero et al., 2014). Since LMSs are constantly updated, pre-processing may eventually become part of their default setup. However, the time gap for their adoption by academic institutions often leaves many researchers with partially processed or unprocessed datasets. Thus, it is essential for researchers to fully understand the details of the data at their disposal and report the pre-processing procedures used in their scientific publications. Striving to do so will enable one to build models for predicting students' behavior, increase the reliability, authenticity, and reproducibility of this type of research, and provide the means for a trustworthy comparison of different studies in meta-analysis studies (Holme, 2019; Pelánek et al., 2016).

Data collection, cleaning, and filtering are likely to have a great impact on interpreting the data. For example, Alexandron and colleagues showed that about 15% of the students in a physics Massive Open Online Course (MOOC) course at the Massachusetts Institute of Technology created two user accounts: One account was used to solicit and experience tasks (“fake learners”), and another account was used to submit assignments and receive recognition for completing the course. Their results showed that fake users attempted fewer questions and showed minimal interest in the instructional materials, in contrast to the actual learners. On the other hand, their time-on-task was significantly faster. Such accounts can bias the research results; however, this can be bypassed by careful cleaning and interpretation processes before the analysis (Alexandron et al., 2019).

Only after this stage has been completed can the researchers advance to a reliable quantitative analysis. The EDM analysis enables one to trace students’ activities regarding engagement at any stage of a course’s progression. It also allows the learning processes to be evaluated through formative and summative assessment (Cohen; 2017; Soffer & Cohen, 2019).

1.1.2 Evaluation of Online Courses

The online learning format has several advantages, including time flexibility, accessibility, and visibility (Bouhnik & Marcus, 2006). Along with its many advantages, online learning presents several challenges. These include the loss of face-to-face interactions (Shaked et al., 2020), reduced learner engagement due to passive learning from video lessons (Brame, 2016), and high drop-out rates (Stephens & Jones, 2014; Tømte, 2019, Shea & Bidjerano, 2019; Arora et al., 2014; Soffer et al., 2017). These unique challenges necessitate the creation of new evaluation frameworks (Alturkistani et al., 2020; Rodrigues et al., 2018). These kinds of frameworks can assist researchers and course designers in assessing these courses, both in terms of the content studied and their online format. Learning outcomes can be evaluated by applying traditional course evaluation tools. However, the significant volume and the wide variety of digital data generated by online course environments have opened up new opportunities for evaluation using EDM techniques (Alturkistani et al., 2020; Bozkurt et al., 2017; Peña-Ayala, 2014 Romero & Ventura, 2010).

With the growing interest in online learning, many educators and researchers are increasingly concerned about the quality of the courses (Baldwin & Ching. 2021). Existing evaluation research of online courses has focused on the completion rate, learning patterns, collaboration, interaction, and access to online materials (Rodrigues et al., 2018). However, further research is required regarding the methodological approaches suitable to evaluate online courses and learning outcomes (Martin et al., 2019; Rodrigues et al., 2018). In this research, we combine the EDM technique for course evaluation with a more traditional evaluation method. Our assessment of learning outcomes relies on the Structure of the Observed Learning Outcome (SOLO) taxonomy (Biggs & Collis,1982), which classifies learning outcomes in terms of their complexity. The SOLO taxonomy is presented in greater detail in chapter 4.

One central aspect often identified in the course evaluation, which has received increasing attention in the literature, is online courses' high dropout rate (Shea, & Bidjerano, 2019; Arora et al., 2014; Soffer et al., 2017). Dropout rates vary according to the course setting: academic, open online, or professional development (PD) (Levy., 2007; Onah et al., 2014). This low completion rate is usually noticed at the end of the course during the summative evaluation phase. However, this is often too late for intervention (Soffer & Cohen., 2019). This can be addressed by creating early prediction models that can design early interventions in future courses.

The following section will elaborate on the challenges of developing early prediction models that measure persistence in online courses. Here we focus mainly on the theoretical frameworks of learner engagement and self-regulated learning.

1.1.2.1 Persistence in online learning, learners' engagement, and self-regulated learning

Many researchers have attempted to predict success in online courses, and some have developed models that can detect student dropout (Arora et al., 2014; Soffer et al., 2017; You 2016). According to Costa and colleagues (2017), successful models for predicting persistence and learning success are not based solely on computerized log files. Instead, they involve a combination of log-file data and additional information such as grades and demographic data. This approach was applied by Shelton, Hang, and Bugman (2016), who developed a model based on demographic data and students' website activity that successfully predicted 78% of students at risk of dropping out of the course by the tenth week (out of 16 weeks) of the course.

Dalipi et al. (2018) stressed that most of the research on dropout prediction is based on MOOCs. Such courses have become an integral part of the higher education system. However, because MOOCs can significantly differ from other online academic or PD courses (Watted, A., & Barak., 2018), we know less about learning patterns in non-MOOC online learning. Our research addresses this gap by examining course completion in more traditional online academic and professional development courses.

Existing studies have identified several factors affecting persistence and success in online courses. These factors include learner engagement, self-regulated learning, course design, and modes of interaction between the lecturers and the learners (Cohen et al., 2019; Kizilcec et al., 2017). In this dissertation, we focus on learners' engagement characteristics and the role of SRL in determining students' persistence in online learning (Li et al., 2020; Soffer & Cohen, 2019). These factors represent research and theoretical frameworks applied to assess students' functioning and performance in academic contexts (Wolters & Taylor, 2012).

Broadly defined, student engagement is viewed as a person's active participation in school-related activities (Wolters & Taylor, 2012). Academic engagement is characterized by behaviors that aim at high-quality accomplishments; it can be determined by asking questions on content in class, completing assigned classwork, and accruing credits toward graduation (Appleton, 2012). Learner engagement in the context of online learning is a multifaceted concept; it can be measured differently, depending on the learning contexts and objectives (Trowler, 2010). For example, if learners are placed in a collaborative learning environment, their engagement with their team would be of primary interest. In contrast, if learners are supposed to perform independent online learning, their engagement with online content should be essential for their learning.

Given that there is no real-time guidance from an instructor who can ensure their timely progress in such an environment, learners' engagement with the course content is critical for them to achieve independent learning (Hampton & Pearce, 2016). In that regard, Angrave et al. (2020) highlighted the need to identify reliable measures representing different aspects of learner engagement in a video-based learning environment. Soffer & Cohen (2019) suggested measuring engagement by assessing students' activities in the online course, learning materials, interpersonal interaction, and learning outcomes. They explored engagement by analyzing the LMS log files using EDM methods to predict success and course completion. Soffer & Cohen (2019) also distinguished between course completers and non-completers using the engagement characteristics. They found that engagement with the course materials (i.e., the average unit page entries, course homepage entries, and total entries) and engagement in the online forums and assignment submissions were significant predictors of course completion. They thus emphasize the importance of engagement in the online course's various activities.

The central variables used as engagement indicators in prediction models often involve video activity (Kovacs 2016). These include click sequences (e.g., re-watching a video, fast-forwarding, pausing, fractional, and the total amount played) or the number of videos viewed per week (Lemay & Doleck; 2020). As to predicting course success, the existing research is ambiguous. Soffer & Cohen (2019) found that video lecture variables such as video views, video activity in days, and minutes of video viewed were not found to be significant predictors of course completion and success on the final exam. Other studies have found video activities to be a significant predictor of course success (Lemay & Doleck 2020; Lu et al. 2018). Although most of these studies used video views per week, this dissertation will present an accumulative variable for assessing the video opening pattern. This provides a valuable indicator of how learners are progressing in the course from week to week.

Wolters & Taylor (2012) suggested that engaged learners exhibit behavior reflective of self-regulated learning (SRL). SRL is defined as setting one's goals and managing one's own learning and performance (Zimmerman, 2000). Birenbaum (1997) suggested three learning-strategy categories: cognitive, metacognitive, and resource management learning. Cognitive learning strategies include problem-solving abilities, critical thinking, database use, and selecting and processing relevant information. Meta-cognitive skills include applying learning strategies, self-esteem, and reflection. Resource management proficiency includes managing the time and the learning environment. According to Pintrich (2004), the behavior of learners matches their self-regulation capability; one manifestation is learning persistence. Studies that have used this concept in online courses found that learners with high self-regulation skills have better chances of completing online courses than those who lack such skills (Rakes & Dunn, 2010). This is likely because online learners are responsible for initiating, planning, and conducting their learning. Indeed, many online learners have expressed how difficult it is to maintain their motivation and persistence throughout the course (Michinov et al., 2011).

Nawrot & Doucet (2014) studied time management, a central element of SRL, and found that inadequate time management was responsible for 51% of the dropouts in a MOOC. Accordingly, they recommend encouraging learners to acquire time-management skills. However, simply providing general information concerning time management and self-regulation is not enough to promote persistence (Kizilcec et al., 2017). Kizilcec suggested ongoing training that makes SRL an integral part of the learning resources; proper course design could be a more effective way to integrate these skills into online courses (Kizilcec et al., 2017).

Most studies on SRL implement a self-report questionnaire to measure students' level of self-regulation (Hadwin et al., 2007; Barnard et al., 2009). The Online Self-Regulated Learning Questionnaire (OSLQ) of Barnard et al. (2009) is a specific means of assessing SRL in the context of online learning. It addresses several categories that characterize SRL in an online environment: goal setting, learning environment, task strategies, time management, help-seeking, and self-evaluation. However, Baker et al. (2020) pointed out that the main disadvantage of relying on self-reported questionnaires is that many individuals suffer from self-report bias, and students' memories are often insufficient for them to accurately recall past behavior or predict future events.

SRL can also be assessed by analyzing data produced through LMSs (Eidelman et al., 2019, You, 2015). To take advantage of the event-based data produced by LMSs, it is necessary to interpret the data in terms of SRL processes. One way to do so is to count specific types of observable actions supported by the online environment; this may directly reflect specific SRL strategies such as help-seeking, note-taking tools, and so on (Aleven et al., 2010). Such an analysis can be carried out using the LA approach, which explores the unique and increasingly large-scale data originating from educational settings (Luna et al., 2017).

Data collected with this system tend to be fine-grained event data and thereby help support a view of SRL as a sequence of events (Aleven et al., 2010). Analyzing these data allows instructors to discover meaningful patterns (Gašević et al., 2015). It also helps identify students who are highly likely to complete the course instead of those who might need help at an early stage. Identifying these students early enables proactive feedback and the ability to adjust and adapt instructional strategies (Dietz-Uhler & Hurn, 2013).

However, it is impossible to receive a complete picture of students' SRL solely from log file data. For instance, Baker et al. (2020) show that researchers could not capture the student activity in web pages outside of the LMS; therefore, their data on online course-related activity were incomplete. According to Li et al. (2020), time management is the central SRL dimension that can be assessed using log file data. Examples of such variables include meeting the assignment submission deadline and the time of the activity in the course (Cerezo et al., 2016; You, 2016; Cormack et al., 2020). Unlike self-reported measures usually collected once or a limited number of times, these measures help researchers investigate how students manage their time during the course (Baker et al., 2020).

The gap between the known importance of SRL and the existing predictive models calls for further development of these tools and in-depth analysis (Li et al., 2020; You, 2016). SRL theories help analyze how students manage their learning and evaluate the actions that they choose to perform (Roll & Winne, 2015). To this end, this research focuses on students' (and teacher-learners) SRL in online chemistry courses.

Existing research on predicting persistence in chemistry courses has focused on background indicators such as high-school achievement and scholastic aptitude test (SAT) scores (Lewis & Lewis, 2007). The increase in online platforms (Amaral et al., 2013) provides new opportunities to focus on more proximate indicators to predict student performance in each course. With data generated from these platforms, we can consider students' past achievements, academic backgrounds and present different learning behavior patterns.

1.1.3 Chemistry Education and Online Learning

The integration of technology into chemistry education positively contributes to teaching and learning chemistry (Barak, 2007; Barnea & Dori, 1999; Battle et al., 2010; Feldman-Maggor et al., 2016; Tuvi-Arad & Blonder., 2019). This is true in terms of adaptation to chemical content (Clark & Chamberlain, 2014) and the possibility of creating interactivity and dialogue between learners beyond class time (Rap & Blonder, 2016). A significant resource in online courses is videos, which enable the presentation of microscopic processes and experimental techniques involving expensive instruments that are not usually available in the classroom (Blonder et al., 2013). Watching videos also make it possible to address difficult topics and complex concepts (Johnson et al., 2014; Read & Lanscate, 2012).

Yet despite these advantages, online chemistry courses often have a low completion rate of online learning (Eitemüller et al., 2020; Gregori et al., 2018). To better understand this phenomenon in the context of online chemistry courses, this research focuses on the role of engagement and SRL. We evaluated learners' existing SRL skills that help them successfully complete the online course. Studying SRL in the context of chemistry will help to identify which specific SRL skills are essential while learning chemistry online.

2. Methodology

2.1 Research Questions

The main research questions are as follows:

Q1. What characterizes learners who are likely to complete online chemistry courses and those that are less likely to do so?

Q2. How can we evaluate learning outcomes in the context of online learning?

Q3. How can we identify learners' difficulties in the online course?

Q4. What is the earliest stage in the online course in which one can predict course completion, and which indicators are required to make these predictions?

To address these questions, we divided the research into two stages:

1. Characterization:
    - We characterized students' and teachers' learning patterns and difficulties in online chemistry courses.
    - We identified the learning patterns in online chemistry courses, which led to completing the course successfully or unsuccessfully.
2. Prediction:
    - We built two statistical models that help predict whether learners are likely to complete the online course successfully or not.

In each of the following chapters, we provide more specific research questions derived from the items presented above.

2.2 Research Set-up

This study focused on courses taught at two different institutions in Israel: The Open University of Israel (OUI) and the Weizmann Institute of Science. Below we discuss these institutions and their unique characteristics. Our decision to analyze online courses from two educational institutions stems from two reasons. First, at the Open University, we had a large sample of students, which could be used to build statistical models. Still, due to data limitations, we were unable to analyze the course content. Although the population was smaller at the Weizmann Institute, it provided an opportunity to closely analyze students' learning outcomes and assess their ability to acquire scientific content. Second, an analysis of data from two institutions made it possible to draw more general and reliable conclusions that do not necessarily depend on the nature of specific learners.

2.2.1 The Open University of Israel

The OUI aims to make higher education accessible; therefore, it does not have prerequisite admission requirements. It, therefore, admits all individuals who seek to utilize their learning potential. It does so by offering a challenging academic program, by developing advanced distance learning methods, and reaching out to potential students from the country's geographic and socio-economic periphery. The OUI offers a variety of learning methods, including face-to-face tutoring in small groups in over 60 study centers throughout the country as well as interactive online learning groups; this provides students with maximal flexibility in building their curriculum (https://www.openu.ac.il). Although there are no prerequisite admission requirements for undergraduate students, they need to demonstrate a high level of knowledge and skills to successfully pass university courses.

Overseeing the OUI's study centers and running an online learning environment open to students of all ages from Israel and abroad constitute an administrative challenge. It requires the cooperation and collaboration of many administrative departments working together with the academic departments in order to run, manage, and support the educational programs. Our research was conducted within the Department of Natural Sciences; however, collecting data for the study required the close cooperation of several administrative departments at the university. These included the Center for Technology in Distance Education, the Teaching Services System, and the Computer Center.

In this research, we study students who took chemistry courses in their online format. Three core courses were included in this study: 1) The World of Chemistry (WOC): an optional introductory course for students without any previous background in chemistry; 2) General Chemistry A (GCA): a mandatory course for both chemistry and life sciences students; 3) General Chemistry (GC): a mandatory course for life science students. All three courses were delivered through a Moodle environment and the Zoom platform. Each course included a textbook, one or two lab meetings (GCA and GC), course website, and 12 online tutoring sessions. Students could decide whether to participate in these tutoring sessions synchronously (live) or view them asynchronously (recorded) at their convenience.

2.2.2 The Weizmann Institute of Science

The Weizmann Institute of Science is a multidisciplinary basic research institution in natural and exact sciences. The Weizmann Institute conducts research and offers graduate education in various scientific disciplines, emphasizing cross-disciplinary investigation. We conducted our research at the Department of Science Teaching. This department's mission is to advance the field of science and mathematics education. In addition to Masters and Ph.D. programs, the Department of Science Teaching offers courses for teachers' PD. Some of the courses are delivered face to face in traditional classrooms, and some are provided online. The course we studied in this research was an online, one-semester-long course for teachers. Our research required direct cooperation with the department's technological staff, and on several occasions, we needed support from the department's administrative and institutional technological units.

Data were generated from several chemistry courses taught at the two institutions described above from 2016 to 2020. The general characteristics of the courses that were included in the study are described in Table 2.1 for the OUI and Table 2.2 for the Weizmann Institute of Science.

Table 2.1 Characteristics of undergraduate general chemistry courses at the OUI.

Course Name	Number of students [a]	Number of mandatory assignments [b]	Number of optional assignments	Course requirements
World of Chemistry (WOC) (3 credits) [c]	517	2	At least 1 out of 3	Final exam
General Chemistry A (GCA) (4 credits)	219	3	At least 2 out of 5	Two mandatory laboratories sessions (4 hours each) Final exam
General Chemistry (GC) (6 credits)	218	3	At least 2 out of 5	One mandatory laboratory session (4 hours) Final exam

[a] Total enrollment for online study groups in the years 2017-2020.[b] Including laboratory reports listed under "Course requirements." [c] During the research period, a bachelor's degree at the OUI required 108 credits.

Table 2.2 General characteristics of teachers' PD course at the Weizmann Institute of Science.

Course Name	Number of semesters per year	Participants per semester	Learning Materials	Requirements to complete the course
Introduction to materials and nanotechnology	1	30-40	• 13 pre-recorded video lessons, each comprising up to five, 25-minute-long videos • Course website • One face-to-face tutoring session	• Opening Forum – the participants introduce themselves and add a link to one nanotechnology application • Submit 13 short quizzes • Participate in 4 assignments on a Padlet board* • Submit a final course assignment

* Padlet Board is an online virtual board where students and teachers can collaborate, reflect, and share ideas in a secure environment (https://padlet.com/).

2.3 Research Population

2.3.1 Participants from the OUI

A total of 954 students were enrolled in at least one of the three chemistry courses described in Table 2.1 in an online format in 2017-2020 (7 semesters). Of these 954 students, 64 were counted twice since they enrolled in two of the courses during different semesters. Student descriptions appear in Tables 2.3 and 2.4.

Table 2.3 Student educational background

Educational Background	
Certification type	**First course at the OUI?**
Bachelor's degree - 13% (120)	No - 41% (387)
Matriculation Certificate or took academic courses in high school * - 68% (650)	Yes - 59% (567)
No Matriculation Certificate - 16% (158)	
Missing– 3% (26)	

*High School Academy is an academic program for high-school students.

Table 2.4 Students' demographic characteristics

Area of residence according to a socio-economic status *	**Gender**	**Average Age (Years)**
4 - 40% (383)	Female - 58% (555)	23
5 - 17% (160)	Male - 42% (399)	
7 - 27% (262)		
8 - 15% (144)		
Missing - 1% (5)		

* Based on a division into socio-economic clusters of local authorities according to Israel's Central Bureau of Statistics (1 – the lowest value, 10 - the highest value).

2.3.2 Participants from the Weizmann Institute of Science

Our sample included three cohorts with 95 Israeli chemistry teachers who took the course from 2016 to 2019. The teacher's description appears in Table 2.5

Table 2.5 Teachers' demographic characteristics

Area of residence, according to a socio-economic status *	Gender
1-3 26% (25)	Female - 83% (79)
4-6 29% (28)	Male - 17% (16)
7-10 42% (39)	
Missing - 3% (3)	

* Based on a division into socio-economic clusters of local authorities according to Israel's Central Bureau of Statistics (1- the lowest value, 10 - the highest value).

2.4 Research Tools

The research combines both qualitative and quantitative tools (mixed methods) that are known to increase the precision and trustworthiness of the results (Leech & Onwuegbuzie, 2007). This section briefly describes the research tools and methods used to study the courses from each of the two institutions. We will then elaborate on a pre-processing phase we developed while using the EDM techniques and the quantitative analysis process. These phases are relevant to all the dissertation chapters. The qualitative methods used will be discussed in greater detail within the relevant chapters.

The OUI: We first used semi-structured interviews to identify the characteristics of students' SRL. Based on these characteristics, we identified several parameters that could be analyzed using the EDM techniques. Finally, we created prediction models using a logistic regression model. Our study of courses at the OUI provides answers to research questions 1 and 4. Due to data limitations, we could not address questions 2 and 3. However, we tackle these questions in our study at the Weizmann Institute.

The Weizmann Institute: We assessed the chemical content and progress in acquiring knowledge throughout the course. Specifically, we evaluated the courses regarding three dimensions: 1) knowledge, 2) the complexity of learners' understanding, and 3) identification of learners' difficulties. We evaluated the first aspect using a pre-post questionnaire, the second using the Structure of the Observed Learning Outcome (SOLO) taxonomy (Biggs & Collis,1982), and the third by analyzing online activity reports and semi-structured interviews. In addition, we used case studies to identify the characteristics of students' SRL. Based on these characteristics, we identified several parameters that could be analyzed using EDM techniques. Finally, we created a prediction model using logistic regression. Our study of the PD course at the Weizmann Institute provides answers to research questions 1-4.

2.5 Quantitative Data Analysis

2.5.1 Stages of Pre-Processing Data

We defined the stages of pre-processing online educational data, starting from the data collection stage, data preparation for data-mining analysis, and data interpretation. The existing research focuses on either the technical features (Romero et al., 2014) or data interpretation (Pelánek et al., 2016). Our approach addresses all these aspects and stresses the need to collaborate with different people to better understand how data are managed locally. With this procedure, we enhanced the reliability of the data, clarified the procedures required for working with raw or partially processed data, and avoided the pitfalls of working with inadequately processed data.

We divided the workflow of pre-processing online educational data into four stages (presented in Figure 2.1): (1) data gathering, (2) data interpretation, (3) database creation, and (4) data organization. Each stage consists of several sub-stages. The data-gathering stage involves listing the sources from which the data files will be collected and planning the timetable for their collection. The data interpretation stage deals with mapping data from the log files and validating their quality. The database creation phase includes protecting the participants' privacy and uploading the files to a relational database management system (by SQL). Finally, data from various sources are filtered and integrated (by SQL queries). In the next section, we elaborate on and exemplify each stage.

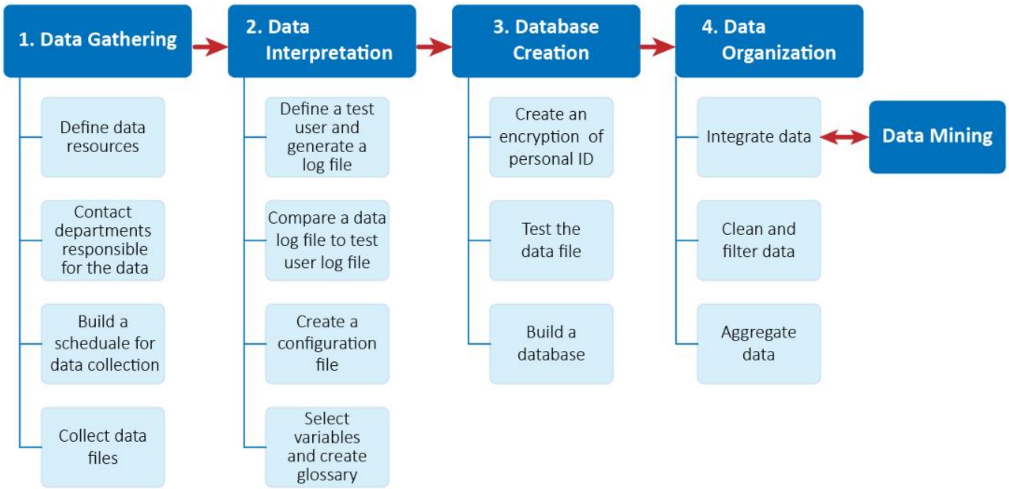


Figure 2.1 A schematic description of the suggested stages for pre-processing data.

2.5.1.1 Data collection and gathering

During the process of collecting data for our study, we faced two main challenges related to data availability and the retrieval process. As described above and stressed by Siemens (2013), carrying out research at big institutions requires administrative cooperation from several university departments and employees. Generally, the researchers have limited control over which data are available to them and in which format. The type of data available depends on the specific software used by the institution, its different features, and the way it is stored and archived. Both institutions in our study used Moodle as their LMS. The Moodle environment allows users to add customized features to the software in plugins, either written by the users or downloaded from public repositories. However, since the implementation of the Moodle system influences tens and even hundreds of course websites, system administrators often follow a slow and cautious policy concerning making changes in a working system. As a result, they tend to avoid installing system updates that are unnecessary or not required for many courses run by the institution. They also conduct extensive testing before a complete installation. Owing to information security considerations, we, as researchers and teaching staff, were restricted from adding such plugins to their course website.

In attempting to circumvent the above obstacles, first we tried to use PIWIK (https://piwik.pro/) and Google Analytics (https://analytics.google.com) to obtain statistical information about students' activities in the courses' LMS. However, the resulting data included the students' activities and the academic, administrative, and technical staff data and were unreliable. We, therefore, decided to work with the raw Moodle log files, as provided by the institutions.

Another challenge in collecting data is related to the study's timeframe, as opposed to the time required for data retrieval. Usually, LMS log files are saved either on the institution's servers or on an external repository (e.g., clouds or third-party servers). This can create several obstacles, particularly for ongoing research in which data are collected more than once a year. First, regular software updates may change how information is indexed or categorized, thus making interpreting back files confusing. Second, updates of the university's computers or servers can change the way data are stored in the archives, making its retrieval more difficult. Third, if a third party manages the institutional archive, recovering the information requires another interaction and cooperation that could affect the study period. However, since the amount of data has increased considerably over the years, the institutional policy may change concerning the period required to move log files to an external archive.

To summarize, in contrast to conducting an experiment in which the researchers have substantial control and access to the collected data, research based on data mining involves various challenges that can limit the availability of the data and its format in the timeframe of the study.

Data sources

As mentioned above, in both institutions, we based our analysis on Moodle log files. In addition, at the OUI, we obtained data from other information systems: the registration system, the online assignments system, and the grading system. The data file details are described below.

1. Folders of Moodle log files: Course activity reports show the number of views for each resource on the course's website. Each file in these folders contains information about a course in a particular semester and includes a free text column that describes an action performed by a Moodle user, which is identified by a Moodle ID string.

2. Grades and demographic profiles: Each file holds a complete set of the student's characteristics (from a particular semester), such as the student's profile, achievements, and information about submitting assignments that the student should have completed. In these files, the student's ID is used for identification purposes. To

set up a single student identifier and protect students' privacy, the national ID number and the Moodle ID string were combined into a new encoded and unique student identifier.

2.5.1.2 Data Interpretation

Before creating a database, it was essential to understand the meaning of each variable. Given a raw Moodle log file, determining what each line represents was often unclear and challenging. This is because titles in the log file did not necessarily match the titles seen by the user in the course's website (e.g., specific video identification in Moodle log files is often an internal string, provided by the system and not by the movie's title). We, therefore, created an activity configuration file in which each activity on the website has a clear connection to its representation in the log file.

To create such a file, we logged into each course as a guest user and carried out different activities in the system. We then immediately checked the way these activities were recorded in the Moodle log file. For example, we double-clicked to open a file, clicked to open a video, downloaded files, and answered a quiz. The configuration file we created was based on this accurate interpretation of all activities and served as an organizing scheme for the data in our database. The configuration file also helped us identify the differences between various courses and between different semesters of the same course due to software updates or changes in the website content. For example, Angeli and coworkers (Angeli et al., 2017) described the contrast between quantitative questionnaire-based research, in which the researcher knows the set of possible answers in advance, and the complex analysis of online behavior, where the meaning of the data attributes is not always clear. An online behavior analysis requires the researchers to carefully examine the data attributes that appear in the log files to prevent misinterpretation. For this purpose, an information reliability glossary is needed, as described below.

Information reliability glossary

The LMS records the buttons clicked by the users and can document the sites that users have visited. This often leaves the wrong impression that computers can track all user actions. However, major differences often exist between the actual actions and the way they are recorded by the LMS. Whereas writing is recorded as is, and the text entered by the user is archived, actions such as watching or reading become “opening” or “downloading” in the log files, with no clear ability to interpret what the actual action was. Another problem concerns data that are accurately collected by the LMS (e.g., IP addresses and time of action); however, one cannot base reliable conclusions on it, as explained below. To create a unified language of concepts that will be reliable and prevent ambiguous interpretation, we created a shortlist for the type of attribute we intended to analyze. This list, detailed below, spans only part of the documented attributes; it should be viewed as a flexible tool that can be expanded according to the LMS type, the collected data, and the research goals. Creating such a list should be regarded as an essential part of any research in this field.

1. User Type

The user type category can often indicate whether the user is a student or an instructor. However, different LMSs do not always separate students from the academic, technical, and administrative staff; therefore, usage statistics may be inaccurate. For example, Figure 2.2 describes an error resulting from counting the activities of all users (including academic and technical staff, instructors, and possibly administrative staff) as opposed to only students from the courses in our study. As can be seen, in these examples, the relative error can be significant and reach as high as 32%. The descriptive statistics presented in Table 2.6 indicate that an average of 22% of the records per course does not represent the students. Analysis of the total number of users’ activities, as opposed to only those of students, can create a secondary bias if, for example, technicians entered a specific course module several times due to a technical problem, making it look as if this was the most popular activity on the course website. It should be stressed that separating the teaching and administrative staff is not always easy, especially for extensive courses with several tutors and a large technical team. In our study, this separation was achieved by integrating the Moodle activity data with students’ grades, which by definition, did not include other types of users.

2. Timestamp

The timestamp indicates the exact time and date of each user activity. It can be used to explore dates with increased activity throughout the course period (e.g., towards the final exam). Ideally, one could deduce from it the time each user devoted to each activity. However, this would often be unreliable because, for example, students could simultaneously work in the LMS and surf the Internet (Cerezo et al., 2016). It is possible to measure the user's overall time in the system only if the "logout" button was clicked. However, if the user left the course website without "logging out," the system would not calculate it accurately. Another indication of time is the time gap between different activities of the same user on the same date. However, this could be unreliable if the user took an undocumented break (without logging out before the break) or had technical problems that required reloading the page several times. This creates duplicate records in the data (see Figure 2.3). If a user entered a specific activity twice in the same time frame, two records with similar time signatures would appear in the log file. The researcher should define a time difference threshold (e.g., one minute) below which the time differences can be neglected, and two consecutive records of the same user and the same activity can be considered identical. An example of the differences due to duplicated rows for a time difference threshold of one minute is shown in Figure 2.2. Descriptive statistics of these data for all courses are presented in Table 2.6. Owing to the challenges of using timestamps, we did not evaluate the total time that learners were engaged in an activity. Instead, we used "weeks" as our time unit. We divided the courses into weeks starting from the first day of the course.

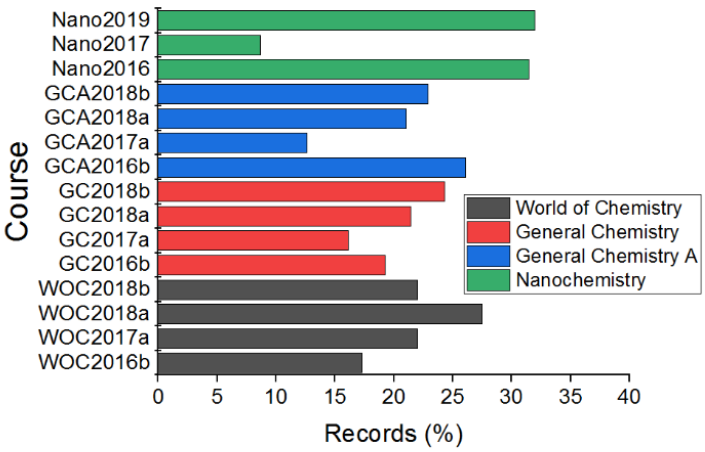


Figure 2.2 Differences (in percentages) between the number of records of all users and students only, per course, per semester. The Y-axis presents the course name and semester from which the data were taken. Black: WOC; Red: GC; Blue: GCA; Green: Introduction to Materials and Nanotechnology.

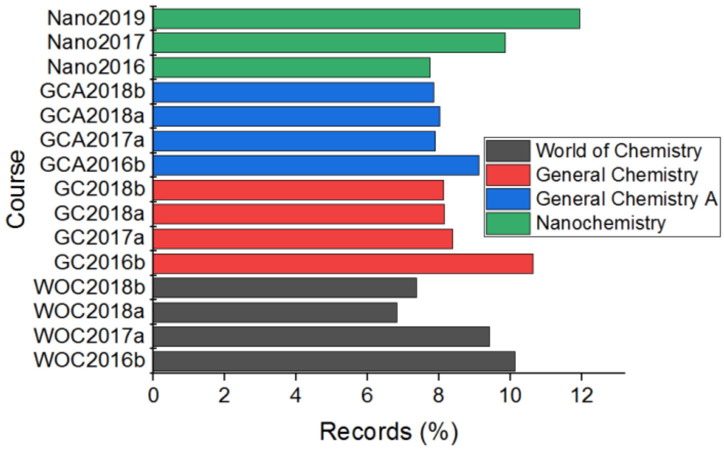


Figure 2.3 Percentages of records with the same timestamp per course, per semester. The Y-axis presents the course name and semester from which the data were taken: Black: WOC; Red: GC; Blue: GCA; Green: Introduction to Materials and Nanotechnology.

Table 2.6 Descriptive statistics of records removed from the log files of all courses (N = 15)

Category	**Average**	**Median**	**Standard deviation**	**Maximum**	**Minimum**
Non-student records [a]	22%	22%	6%	32%	9%
Identical Rows [b]	9%	8%	1%	12%	7%

[a] 149,094 entries out of 690,846 entries from all courses, [b] 45,505 entries out of 541,752 entries from all courses.

3. IP address

The IP address can tell the researcher the approximate location from which the user connected to the LMS. Ideally, knowing the IP address can help determine whether students are learning from home or during other activities (e.g., while on the train, in a coffee shop, or in the library) and statistically test their environment preferences. However, suppose the user used a proxy server or a LAN router (e.g., when surfing in a private network or in a workplace or organization). In that case, this information becomes unreliable and thus, prevents interpreting the data correctly. Therefore, in this study, we did not use the IP address in our analysis.

4. File opening

Different file types embedded in the LMS can lead to various reports on similar user actions. The log file generally indicates whether the user entered an online activity such as a presentation, video movie, or quiz. However, it does not indicate what the user did during the activity. Different types of files or links to these files (e.g., Scorm, Power-Point, PDF, Word, Excel, Jpeg, and links to Vimeo/YouTube) are uploaded to the LMS by the teaching staff. Observing the Moodle log can indicate whether a user opened a file or not. However, it is impossible to know how the specific file was used. For example, if the user opens a video on the website in a YouTube or Vimeo format, the action is reported as "played." A scorm video file is reported as both "opened" and "uploaded". An H5P format, on the other hand, provides more information about the speed at which the video was viewed and when the play/pause option was selected. In this study, we used video plays as a file opening category.

5. Activity counts

Activity counts represent the number of times all users entered a given activity. Upon careful interpretation, they can help measure the dropout rate (or degree of persistence) of the course at hand. Note that a simple count does not imply how many times each user repeated the activity or whether users completed the activity at all. However, one can differentiate between unique user activities and total activities. When researchers are interested in examining a specific video pattern, they should decide if they want to count the number of users who played the video (unique plays per user) or the number of times the video was played (total plays). Figure 2.4 compares two counts for the course "WOC" at the OUI during the winter semester of 2018 (with 89 students). As is evident, with both counting methods, the number of videos played declined between the first and last lesson. However, both the overall level of decrease and the decreasing trend is different: In examining the total number of videos played (Figure. 2.4a), one can note a relatively stable trend for the first three lessons (with a small increase for lesson 4), a slight decrease for the following lessons, and a sharp decrease between lesson 8 and 12. Overall, there were 360 plays of the first lesson and only 141 plays of the last lesson, suggesting a dropout rate of 61% (a 39% persistence rate). For the unique number of plays (Figure. 2.4b), the trend is different. There is a sharp decrease between the first two lessons, a moderate decrease for the rest of the semester, and a second sharp decrease between the last two lessons. Overall, the number of videos played dropped from 89 in the first lesson to 46 in the last, pointing to a dropout rate of 48% (a 52% persistence rate). Nevertheless, note that a decrease in the total number of videos played does not necessarily imply the dropout rate since students may have chosen to view the same video fewer times towards the end of the course or used other resources. Moreover, more video lessons become available as the semester progresses; thus, students have more opportunities to view the first video than the last one. Therefore, discussing dropout rates may be more reliable when they are based on a unique number of plays rather than the total number of plays. On the other hand, a video that the same students played several times may inform us about students' interest in that particular topic or difficulties they may have encountered with its content. Upon careful analysis, the number of repeated plays of a specific video lesson may also indicate students with distinct learning behaviors (Hassner et al., 2014). Finally, predicting the dropout rate cannot be based solely on the number of videos played, and other parameters, such as grades should be considered. In our analysis, we mainly used unique videos that were played.

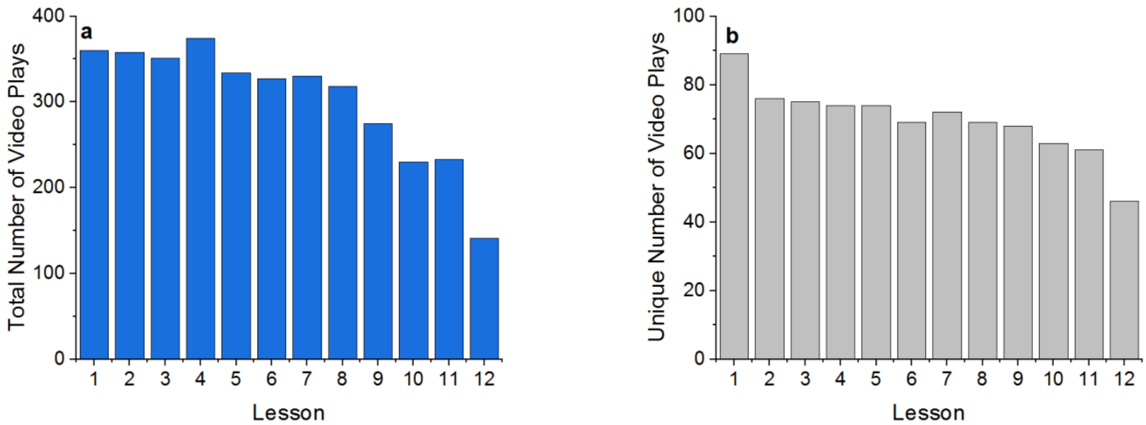


Figure 2.4 A number of videos played throughout the semester for the WOC course at the OUI during the 2018a semester. a. Total videos played (blue); b. Unique videos played (gray).

2.5.1.3 Creating a database

In order to maintain the principles of research ethics and students' privacy following the EU's General Data Protection Regulation (GDPR) and Israel's Protection of Privacy Law, identifying fields such as name and surname were removed. In addition, national ID numbers and moodle_id identifiers were encrypted. In addition, we checked each file to ensure that it represents the correct data in terms of the course's name, semesters, and all the activities included on the course website. These data were then used to create a relational database based on the configuration files previously discussed.

2.5.1.4 Data organization

At the data organization stage, data from various sources are filtered and integrated. We applied SQL queries to filter the data according to the variables' glossary discussed above. The results of each query are aggregated for further analysis. An example of aggregation is the creation of a new variable that presents the total number of participants that entered the course LMS or the number of times that each video was opened in a different month of the semester. These parameters are the starting point for developing students' online behavior models.

The suggested stages for pre-processing data were applied in this research. Data analysis from two different institutions enabled us to reach generalizable and reliable conclusions that do not necessarily depend on the institute, the course, or the learner characteristics.

2.5.2 Collective Variables Used for the Quantitative Video Analysis

Following the data organization stage, we used the filtered data to develop and define new collective variables to describe students' learning patterns. Below, we list and briefly describe the central variables used for the study's quantitative analysis. The rationale for developing these variables is presented and discussed in greater detail in chapters 3 and 5.

a) Skipping Index: Skipping lectures in online learning is a common phenomenon (Warner et al., 2015). To measure skipping, we created a skipping index. This index counts the total number of video sessions that the learner did not play during the course. For example, suppose the learner did not open 4 out of 12 video sessions. In that case, the “Skipping Index” will be 4. Note that if a learner played the video lectures in non-sequential order but eventually played them all, we did not consider that as “skipping.”

b) Student Cumulative Opening Pattern (SCOP): we developed this variable to measure learners’ progress in the course from week to week. The SCOP represents cumulative video sessions that each student opened by that specific week. This variable does not count the re-watched video session. For example, if student A played one video at week one, another one at week two, and zero videos at week three, at week three, his SCOP will be two.

c) Linearity Index: Building on Warner et al. (2015), we created an index to measure if a learner played the video lectures in non-sequential order. Using our “Linearity Index,” we calculate the number of times a learner did not play the video according to the course sequence. For example, if the learner played the video in the following order: 2,3,4,5,1,6,7,8 – the “Linearity Index” will be 1. Since we count unique video plays, this index does not count cases where students replayed a specific video.

d) Number of Active Weeks: We defined learners as active if they played at least one new video during that specific week. We count the number of weeks each student was active in the course, according to this definition.

e) Binge: recently, studies have begun to examine binge-watching in the context of online educational settings (Yoo et al., 2017). However, there is not yet an agreed-upon definition. In this study, we define a learning activity a binging when a learner played 75% or more of the videos during the last third of the course. We elaborate on the binge pattern in chapter 5.

2.5.3 Logistic Regression

In order to address research question number 4 and predict whether a student is going to succeed in a course, there is a need for a statistical methodology that could explain a dichotomous outcome (successful/unsuccessful) based on a collection of dichotomous, discrete, and continuous independent variables. The logistic regression approach (Osborne, 2015) provides such an analysis by moving from predicting an event occurrence to predicting its probability to occur. This type of analysis was successfully used in a number of educational studies (e.g., Artino & Stephens, 2009; Yair, Rotem, Shustak, 2020). This method involves defining a new dependent variable, Logit, defined by:

(1) $\text{Logit}=\text{Ln}\ (p/(1 - p))$

where p is the probability of an event occurring, in our case – success in a course (Osborne, 2015). The logit function is then estimated by a regression model that is a linear function of a set of independent variables $\{X_k\}$ with coefficients $\{b_k\}$:

(2) $Logit(Y)=b_0 + b_1X_1 +... + b_kX_k$

Here Y is the dependent variable, b_0 is the intercept, and $\{b_k\}$ measures the slopes or the effects with respect to $\{X_k\}$.

Since the data are not normally distributed, we used chi-square and Mann-Whitney U tests for preliminary logistic regression (MacFarland et al., 2016; Onchiri, 2013). Additionally, we used the Hosmer-Lemeshow test to examine model fit (Paul et al., 2013; Fagerland et al., 2012). We further evaluated the logistic regression models by plotting the area under the curve (AUC) to estimate their accuracy based on the Receiver Operating Characteristic (ROC) curve. The ROC curve is plotted with sensitivity in the Y-axis and specificity values in the X-axis. The sensitivity measures the probability that a given statistic correctly predicts the actual condition with respect to a pre-defined threshold. For example, a model predicts that the student will successfully complete the course and that the learner has actually completed it. Specificity measures the probability that a given statistic correctly predicts a non-existing condition with respect to the threshold. For example, a model predicts that the student will not complete the course and the student has actually not completed it. The AUC provides a biased presentation, and its values range between 0 and 1. Higher values represent better classification or discrimination (Raju, & Schumacker, 2015).

2.6 Ethical and Privacy Considerations

This research received the Review Board (IRB) approval of both institutions: the Weizmann Institute of Science and the OUI (#9390). As already, mentioned in order to follow the principles of research ethics and students' privacy and in accordance with the EU's GDPR and Israel's Protection of Privacy Law, identifying fields such as name and surname were removed. In addition, national ID numbers and Moodle identifiers were encrypted.

2.7 Overview of Chapters 3-5

In chapter 3, we study online chemistry courses in the OUI and focus on research questions 1 and 4 of the dissertation. We examine learning processes in undergraduate online general chemistry courses to identify indicators that predict students' success in the course. We also focus on student engagement and SRL, which are central factors that determine success in online courses.

In chapters 4 and 5, we study the "Introduction to Materials and Nanotechnology" online PD course at the Weizmann Institute of Science. Chapter 4 develops a framework that integrates traditional evaluation tools and EDM techniques for evaluating an online teachers' PD course. This framework enabled us to assess learning outcomes and difficulties in the course and to address research questions 2 and 3. Chapter 5 characterizes teachers' learning patterns using five case studies that exemplify different learner types. The learning patterns that emerged in the case studies provided guidelines for a quantitative analysis carried out with EDM techniques. Using this analysis, we addressed questions 1 and 4; we distinguished between teachers who completed the course and those who did not and identified indicators that predict teachers' success.

3. Predicting success in online general chemistry courses

3.1 Highlights

- In this chapter, we identify indicators of students' success in online chemistry courses.
- We show that self-regulated learning is strongly associated with the completion rate in online courses.
- We found that the completion rates are strongly associated with online learning patterns.
- Logistic regression models predict the success rates with a high probability.
- The findings emphasize how students' choices affect their potential for success.

3.2 Introduction

As discussed in the literature review that appears in Chapter 1, completing online courses is known to be more difficult than traditional face-to-face courses. A primary goal of this research was to use indicators of learners' engagement and SRL to produce a generalizable model for identifying students who have a high probability of completing the course as opposed to those who do not. This model allows recommendations on specific interventions that could potentially help increase the completion rate of online courses. Because online course data generally present information about learning behavior, this chapter includes measures related to the frequency of playing online lessons and of assignment submissions that indicate SRL.

3.3 Research Questions

In this chapter we address research questions 1 and 4 (section 2.1):

Q1) What characterizes learners who are likely to complete online chemistry courses and those that are less likely to do so?

Q4) What is the earliest stage in the online course in which one can predict course success, and which course indicators are required to make these predictions?

3.4 Research Set-up and Participants

This chapter presents a study of undergraduate online general chemistry courses offered at the OUI. Here we studied students who took the chemistry courses online (see Tables 2.1, 2.3, and 2.4).

3.5 Methodology

The research design included both qualitative and quantitative tools (a mixed method). The integration of quantitative and qualitative research methods is known to increase the precision and trustworthiness of the results (Leech & Onwuegbuzie, 2007). First, we used semi-structured interviews to identify the characteristics of students' SRL. Based on these characteristics, we identified several parameters that could be analyzed using EDM techniques. Finally, we used these data to create the prediction model using a logistic regression approach.

3.5.1 Analysis of Semi-structured Interviews

To better understand students' learning habits, we conducted 13 semi-structured interviews. These interviews included participants enrolled in one of the three courses studied. Twelve interviews were conducted by phone and one in a face-to-face meeting. Each semester, we posted an advertisement on one of the three courses' websites (alternating between the three courses throughout the year) and invited volunteers for interviews following the final exam. We used an interview protocol of twenty questions organized around subthemes (see Appendix 1). Each interview lasted 20–60 minutes, was audio-recorded and transcribed.

The interviews with students who successfully completed the course (10 interviews) were analyzed according to Shakedi (2003). We began with a preliminary analysis in which we identified 60 categories relevant to the students' learning organization. Next, we narrowed this list down to 34 by mapping the categories into overlapping groups. For example, "pausing a video" or "watching a video by breaking into different parts" was grouped into "strategic viewing." Category names in these two phases were constructed inductively in a "bottom-up" manner and derived from the interview material. Finally, we performed a "top-down" analysis by grouping the categories based on SRL dimensions defined by Barnard and her colleagues (2009). These include goal setting, environment structuring, task strategies, time management, help-seeking, and self-evaluation. First, the doctoral student conducted the initial analysis outlined above.

The research team then validated it in two stages. In the first validation phase, the doctoral student met with one of her advisors to discuss 20% of the interviews and went through the three analysis stages. These discussions led to changes in the original categories and continued until a consensus was reached. Following this discussion, the doctoral student re-analyzed the remaining 80% of the interviews according to the validation process.

The second advisor then performed a Cohen's Kappa validation, accounting for a chance agreement among coders (McHugh, 2012). To this end, the second advisor received a spreadsheet containing statements from 30% of the interviews grouped according to the categories that emerged in the “bottom-up” stage. Separately, she also received the six dimensions defined by Barnard et al. (2009). The second advisor then classified the categories to the dimensions of Barnard et al. (2009). We then compared these results to the doctoral student analysis results. The Kappa value of this comparison was 0.76, which is considered moderate (McHugh, 2012).

3.5.2 EDM Analysis

The EDM analysis was based on data from Moodle log files, course grades, and students’ demographic profile data. The Moodle log files contained the course activity reports showing the number of views for each course website resource. The grades and demographic data included a complete set of student characteristics (from a particular semester) such as the district of residence according to socio-economic status (SES), gender, educational background, achievements, and assignment submission status. Each Moodle log file contained data about a course in a particular semester and included a free text column describing a Moodle user's action, identified by a Moodle ID string.

The resulting database naturally contained numerous fictitious user activities. These can create a bias in the activity trends, consequently leading to inaccurate conclusions. To enhance data reliability, we performed a pre-processing phase that included four consecutive pre-processing stages: data gathering, data interpretation, database creation, and data organization (see Chapter 2).

First we conducted semi-structured interviews. After this phase, we identified several parameters that can be analyzed using data mining techniques (defined in the Methods chapter in section 2.5.2). During our analysis, we followed Soffer & Cohen (2019) and divided the students into two groups: 1) students who successfully completed the course and 2) those who did not complete the course. In our analysis, we focused on the course assignment submissions and the parameters of the video plays. Finally, we used these data to create the prediction model using logistic regression. Since the pedagogy of the three courses is similar, and there is an overlap in the course staff and content, we consolidated the data. We also added the course variable as a control to the models.

3.6 Results

3.6.1 Summary of Interviews

Let us start with a description of the interviews, which represents the first stage of the analysis. According to the SRL framework, the interviews were used to characterize students' learning behavior in the online chemistry courses. They also helped us identify the main variables for the regression model. Out of the thirteen students, ten successfully completed one of the three courses analyzed in this study, whereas three students did not complete the course. A summary of the interview analysis is presented in Figure 3.1. The headers of each list, in bold, are the existing SRL dimensions, drawing on Brenard et al. (2009): (1) goal setting, (2) environment structuring, (3) task strategies, (4) time management, (5) help-seeking, and (6) self-evaluation. Although we used their suggested SRL dimensions, our categories differ from their questionnaire items. The only similar things are "find a comfortable place to study" (which we called "appropriate location"), "preparing a weekly schedule," and "asking friends for help." In Figure 3.1, the categories under each headline represent students' SRL characteristics that emerged from our analysis of the interviews.

The list of categories is organized according to a heatmap scale, demonstrating the frequency they appeared in our interviews. Note that we count each category only once for each interview, even in cases where it emerged multiple times. We do this to prevent a situation where a specific category is prominent, even though it came up only in one or two interviews. This analysis allowed us to examine SRL in the context of online education in chemistry. Although students did not mention the content explicitly, they did describe their difficulty with chemistry content. In addition, as we see in Figure 3.1, the course assignment was found to play a significant role in the learning organization process. As can be seen in Figure 3.1, course assignments appeared in numerous dimensions.

Table 3.1 presents the number of times that a specific dimension appeared in the interviews together with an exemplary quote. The “goal setting” dimension represents students who mentioned their long-term goals and described why they registered for the course. The “environment structuring” category refers to the physical environment (place of study) and the online learning environment and study materials. We included several different strategies under the “task strategies” dimension. These included learning according to the course assignments, note-taking, and preparing for the online session by reading the chapter in advance. We included learning patterns such as preparing a weekly or daily learning schedule and setting aside a few hours to study each day under the “time management” category. Under the category of “help-seeking,” we included students who described how they discuss problem-solving strategies with their classmates through a WhatsApp group or turn to the course's staff through email, the Moodle forum, or by phone. Finally, we placed students who used the course assignments to self-evaluate their understanding of the course materials under the category of “self-evaluation.

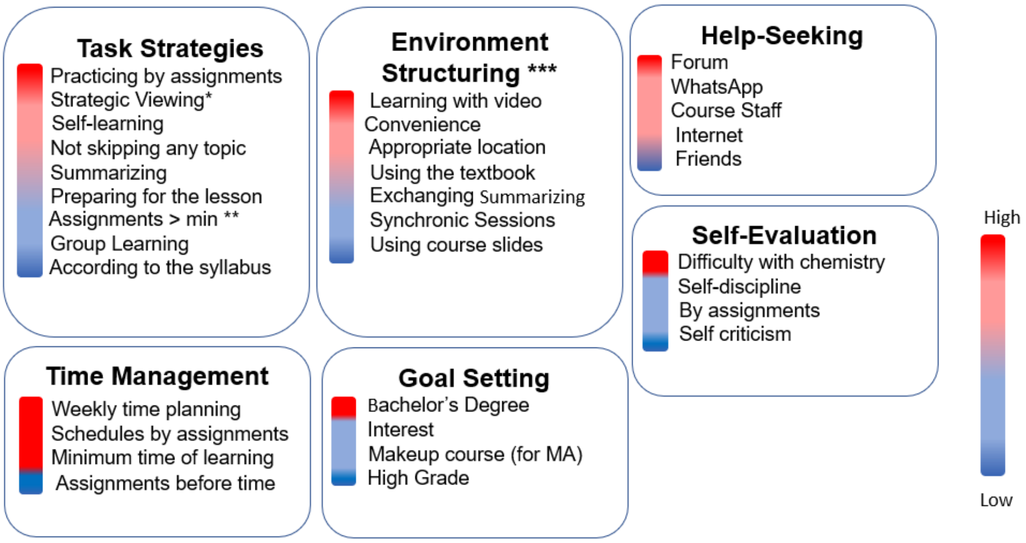


Figure 3.1 A summary of the interview analysis

* Strategic viewing: refers to students who described a proactive intervention in the context of watching a video lesson. This includes, for example, dividing a lesson into several parts independently, pausing a video, rewinding, or fast-forwarding.

** Assignments > min: refers to students who submitted more than the minimum required number of assignments.

*** Learning environment refers to the student’s place of study and the online learning environment and study materials.

Recruiting students who did not complete the course for the interviews was a real challenge in this study. Owing to the open admission, many students in their first semester register for a single course. If they drop it, they usually drop out of the university or move to another institution altogether and are less likely to cooperate. Nevertheless, three such students agreed to participate in the interviews. Although this is not enough to draw strong conclusions, some learning patterns did emerge. For example, all three students reported not setting aside a specific time for learning in their schedule. Although they practiced some kind of learning strategy, it did not involve learning activities such as notetaking or submitting optional assignments. These students also practiced self-evaluation to some degree (similar to students who successfully completed the course). However, they did so without any objective reflection, such as communicating with classmates or working on an assignment. Their self-evaluation was mainly subjective, for example, reporting that the course was difficult for them or that online learning was unsuitable for them.

The interviews helped us identify different learning behavior patterns. As is evident from the above analysis, students use the course assignments to manage their time, decide about their learning strategy, and evaluate themselves. In addition, students who succeed in the courses describe planning a specific time to participate in a live session or to watch a recorded one. Owing to ethical considerations, we did not crosslink the interviewees' answers with their Moodle log files or grades. Nevertheless, based on the interview responses, we focused the analytical part of the study on two major parameters: submission of optional assignments and video session opening patterns. Submission of optional assignments reflects the student’s choice, according to the SRL principles, and it is related to the task strategy category (see Table 3.1). Video session opening patterns are related to two SRL categories: task strategy and time management. These kinds of patterns are reflected in both the Moodle log files and the grade data, as described in detail below.

Table 3.1 An analysis of interview with students who successfully completed the course (n=10)

Dimension **(Number of Interviews where the category appeared)**	**Sample quotes from the interviews** **(Names changed to pseudonyms)**
Goal Setting (10)	"I am studying for a Bachelor's degree in Chemistry." (Bob, male, GCA) "I am focused on achieving my goal. I am trying to get accepted to study engineering". (Avi, male, GC)
Environment Structuring (9)	"I listen to the video recordings at work and on the road." (Roni, female, WOC)
Task Strategies (10)	"I need the structure of the live session. I don't think I missed a lesson, but if I did, I completed the material later (from the recording). I returned to the recordings when I missed something and did not write it down … Every time we finished a topic, I tried to answer a question in the assignment. I would even open the assignment during a session ... I submit the assignment even if I am unsure and get a low grade since that way I get feedback from the lecturer and that is a real blessing for me". (Danielle, female, GCA) "I would watch the video recording once, and then re-watch the parts I did not understand" (Irit, female, GC) "I first watch all of the videos and complete all the assignments. Before the exam, I re-watch them. Solve one problem and stop. There's an option to mark specific parts (of the video), which is very useful; that way you can return to where you stopped watching… I devote more time to watching the recorded sessions and to the assignments. Then I only skim through the textbook, and the solutions to the assignments that are on the website" (Rachel, female, GCA) "I watch the video session at 175% speed, or if it's a specific explanation, at 150% speed. If it is material that I already saw, even 200%...I submitted (the assignments), not the minimum but also not all of them" (Melany, female, GCA) "I was able to submit all the assignments, but maybe I missed one" (Bob, male, GCA)
Time Management (6)	I have specific days and hours when I plan to learn according to my work schedule. I plan in advance the days dedicated to my studies ... I try to watch a lecture in its entirety, try to devote 3-4 hours each time". (Erica, female, GC) "Studies are always in my head, but I study whenever I can ... at least two hours a day". (Roni, female, WOC)
Help-Seeking (9)	"Yes, through WhatsApp. It is a pretty significant tool, both as a social tool for people experiencing the same difficulty and for practical things, for example, comparing a question in an assignment". (Dan, male, GC) "I sent an email with questions to the lecturer, 3-4 times during the course". (Avi, male, GC)

Self-Evaluation (8)	“The assignments are part of the learning process. They give me an indication of what I know. I solve them throughout the lessons and organize them at the end in a Word document.” (Roni, female, WOC)

3.6.2 EDM Analysis

Our results are based on 954 students who participated in the online courses between 2017-2020. Out of this number, 487 students completed the course, and 467 students were active on the course website but did not complete the course.

3.6.2.1 Submission of course assignments

In the chemistry courses at the OUI, students are required to submit 2-3 mandatory assignments during the semester. In addition, they need to submit at least one or two additional assignments out of a list of optional assignments. As the number of submitted assignments increases, their weight in the course’s final grade increases at the expense of the final exam. Figures 3.2 provide information about the assignment submission patterns throughout an entire semester. A close look at Figure 3.2 shows that the rate of students who submitted the minimum number of assignments is similar between the two groups (who successfully completed the course and those who did not complete it). Looking at the columns that presents students who submitted above the minimum assignment or the maximum number of optional assignments, we can distinguish between those students who successfully completed the course and those who did not. Our original goal was to detect the course completion status during the semester. Therefore, we needed a more accurate parameter that would allow us to distinguish between the two groups at an earlier stage.

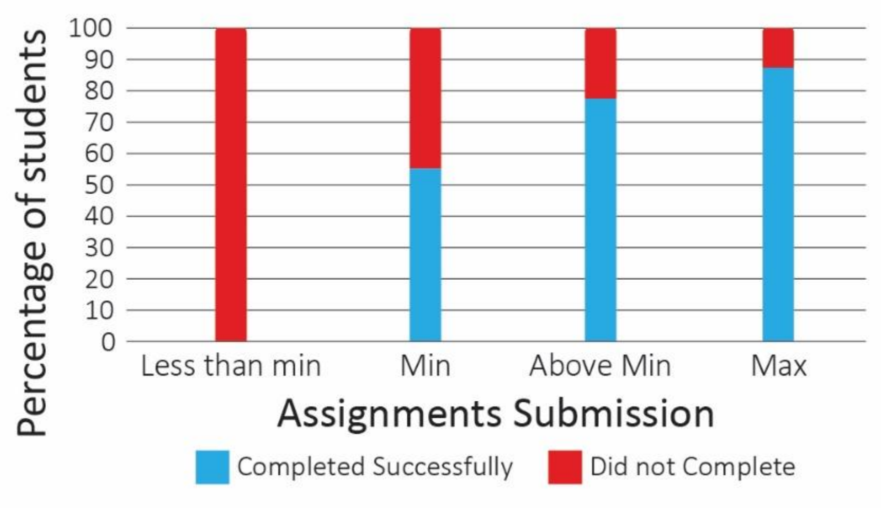


Figure 3.2 Submission of minimum (mandatory) and maximum assignment. Light Blue: Students who successfully completed the course. Red: Students who did not complete the course.

The assignments and their submission schedule are provided in advance, before the course begins. The teaching staff of all three courses made the first assignment mandatory, to take advantage of the students' motivation at the beginning of the course and to create a commitment to learning. Hence, it is not surprising that most of the students submitted the first mandatory assignment (see Table 3.2). Therefore, this assignment cannot be used as a predictor of course success. The student who chose to submit the first optional assignment had to do so by week 5 in the three courses (two weeks before the course’s mid-point). Not all of the students submitted this assignment, making it an informative variable for the logistic regression. Table 3.2 shows the percentage of students that submitted the first two assignments.

Table 3.2 Assignments' submission rate.

Course Name	Number of students	Submission rate of the first mandatory assignment (%)	Submission rate of the first optional assignment (%)
WOC	517	92%	53%
GCA	219	92%	64%
GC	218	93%	50%

3.6.2.2 Course video sessions

The courses in this study consist of 12 online sessions, which students can view either live (synchronous) or recorded (asynchronous). Since many students did not participate in the live sessions and opened the recorded sessions asynchronously, we did not distinguish between synchronous and asynchronous video opening. Note that, as with most online generated data, we know whether a student clicked and opened a video, but we have no way of knowing whether or not the student actually viewed the entire session (see section 2.5.1). Therefore, we referred to this as an opening pattern and not as a viewing pattern. Figure 3.3 shows unified data from all the courses; it counts the number of students who opened each session throughout the semester. The colors indicate two groups of students: (1) those who succeeded in the courses and (2) those who did not complete them. Each student was counted once per session for this analysis. As can be seen, the first group shows a steady pattern of sessions that opened – the number of students is constant throughout the semester, and almost all of them opened each video session at least once. On the other hand, the second group of students did not follow a steady pattern, and the number of students who opened each session significantly decreased throughout the semester.

Figure 3.4 presents a different view of these data; the percentages of students from each group that opened the sessions' first, second, third, and fourth quartiles are counted. As is evident, almost all the students in the first group, who successfully completed the courses, played the entire set of tutoring sessions. Most of the students who did not complete the course opened only some of the sessions. From Figure 3.4 we can see that some students skipped sessions. Therefore, we calculated the number of videos each student skipped (Skipping Index - see section 2.5.2 for definition). Students who completed the course skipped two videos, on average, whereas those who did not complete the course skipped six. A Mann-Whitney test indicated that this difference was statistically significant ($U=30977.50$, $Z=-14.954$, $p<0.001$).

By combining both figures and the Mann-Whitney test on the "Skipping Index," we observed not only that the successful students opened more sessions – but they were also consistent in doing so throughout the entire semester. On the other hand, many students who did not complete the course stopped opening the sessions long before the semester was over. It is important to stress that the online sessions were not mandatory. Based on these results alone, we could not determine whether students who did not complete the course decided to use other course learning materials. Nevertheless, all learning materials were available to all the students, to begin with.

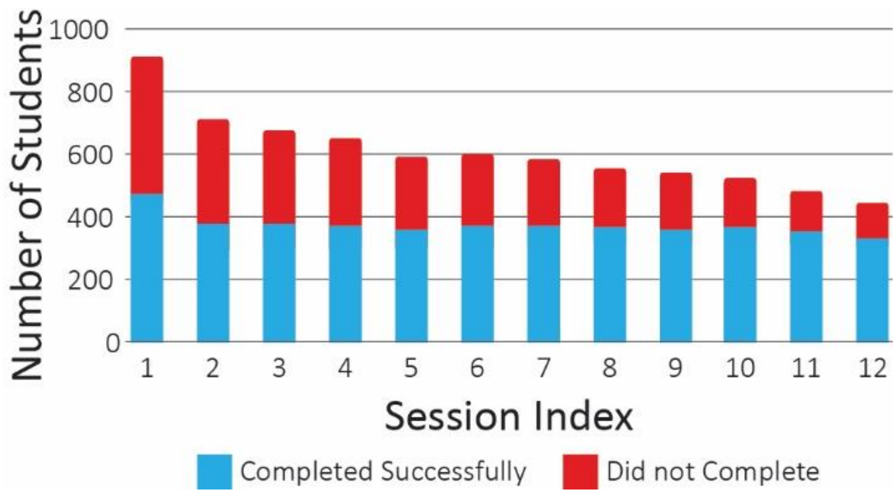


Figure 3.3 Sessions' opening pattern throughout the semester in all the courses. Light Blue: Students who successfully completed the course. Red: Students who did not complete the course.

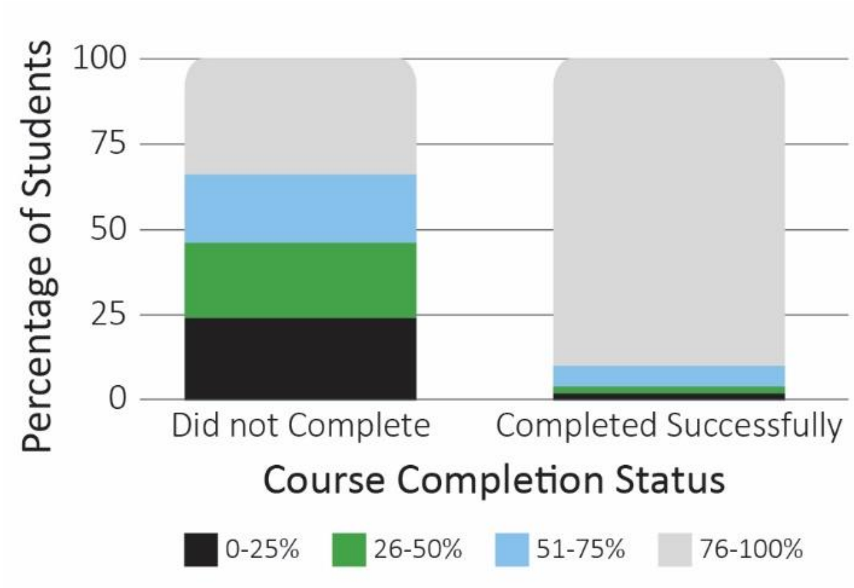


Figure 3.4 The opening rate of the course videos session: Colors represent the video opening percentages. Black: 0-25%. Green: 26-50%. Light Blue: 51-75%. Gray: 76-100%.

Figures 3.3 and 3.4 provide information about the opening patterns that accumulated throughout an entire semester. Both figures help distinguish between students who successfully completed the course and those that did not. A close look at Figure 3.3

shows that the number of students who opened each session during the first sessions is similar between the two groups. In addition, for each student, we calculated the number of weeks s/he was active in the course (number of active weeks - see section 2.5.2). Our analysis focused on the video lectures; therefore, we defined a student as active in a specific week if they played at least one new video during that time. We found that students who completed the course were, on average, active for ten out of the twenty-week course, which includes the final exam period. Students who did not complete the course were, on average, active for only six weeks. A Mann-Whitney test indicates that this difference was statistically significant ($U=33436.500$, $Z=-14.201$, $p<0.001$).

Our original goal was to detect the course completion status during the semester. Therefore, we needed a more accurate parameter that could distinguish between the two groups at the early stages. From the interviews, we learned that there were students who successfully completed the course and watched the online sessions from week to week (see Table 3.1). We, therefore, defined a new variable: the Student Cumulative Opening Pattern (SCOP) that counts the total number of different tutoring sessions that each student opened in a specific week (see section 2.5.2 for definition). This variable does not count the re-watched video session. Figure 3.5 presents the weekly average SCOP for each group (those who successfully completed and those who did not complete the course). As is evident, this parameter is quite informative in distinguishing between the two groups, even at earlier stages of the course. Note that the course itself lasted for 14 weeks. Data for weeks 15-20 represent the exam period. It is included here to show that students continued to open the video sessions at higher rates towards the exam date. The group of successful students used the video resources much more than the other group.

Nevertheless, we aimed to predict students' success in the course at the early stages of the semester. Next, we will focus on the first weeks of the semester.

We designed the SCOP variable not to count numerous plays of the same video. This is because calculating the total video plays would have made it difficult to learn about learners who continued in the course. Currently, if a learners' SCOP is nine, we know that they played nine of the course videos. If we had counted multiple video plays, we would not know whether the number nine represents a continuation of the course or a combination of views and replays of specific lessons.

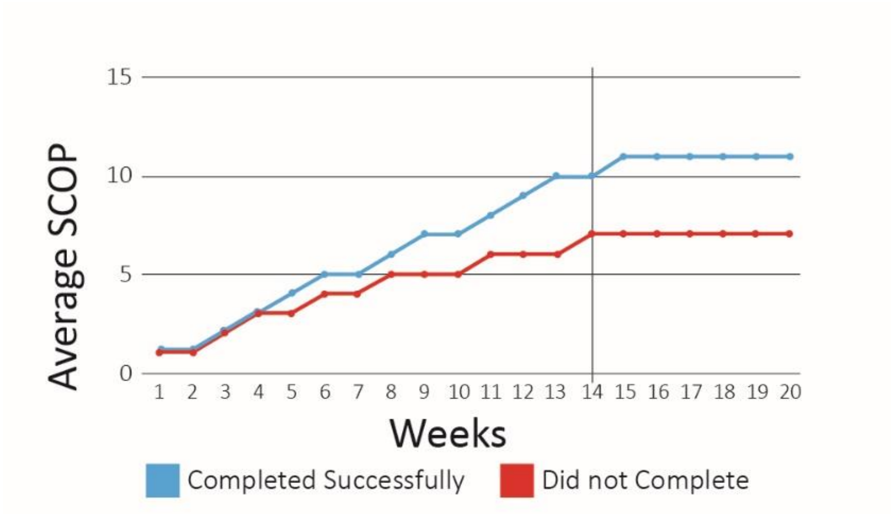


Figure 3.5. Student Cumulative Opening Pattern (SCOP) – weeks 1- 20. Light blue: Students who successfully completed the course. Red: Students who did not complete the course. Lines were added to lead the eye.

The SCOP variable we define does not evaluate whether the students watched the online sessions from week to week in a linear order (lesson 1, lesson 2, lesson 3, and so forth). Therefore, we calculated the linearity index (see section 2.5.2 for definition). For this linearity index, the average linearity for the students was zero – meaning that the average students linearly played the video. In addition, a Mann-Whitney test indicated that no statistical difference exists between the linearity index of students who successfully completed the course and those who did not.

3.6.2.3 Building Logistic Regression Models

Based on the results presented above, we defined two main independent variables for the logistic regression model: the first optional assignment submission and the SCOP. Both of these variables are related to student engagement in the course.

Before the logistic regression analysis, we conducted a correlation analysis for a set of independent variables to test for multicollinearity. Two categories of independent variables were used as control variables in the analysis. The first category consists of demographic variables that include gender and the place of residence. The age variable was not used since multicollinearity exists between this variable and an advanced diploma. We removed the age because the advanced diploma gave us an index of the student's educational background. The second category consists of educational background variables that included prior and current studies: the existence of a previously advanced diploma, an indication of whether the current course is the first course at the OUI, the semester index, and the course name. These factors are suitable as a control since although they may influence success in the course, they do not change during the semester and do not depend on students' learning choices during the course.

No significant associations were found between any of the control variables and the variables of the assignment submission status and the SCOP. However, we did find multicollinearity between the assignment submission status and the SCOP (namely, these variables are correlated). Therefore, we ran two different models of logistic regressions for each of them. Next, we carried out a logistic regression analysis using SPSS (IBM Corp. 2016). For each model, we chose which variables will be included in the model, and all of them were entered simultaneously. According to this approach, the analysis proceeds based on prior theory or research, and as such, it is considered more defensible (Osborne, 2015). The data were collected chronologically, and each model was based on data on students enrolled in the years 2016-2019 (n=797). Data from the first semester of 2020 (n=157) were used to validate these models.

Logistic regression models were built based on the data collected from all three courses. The course name was used as a control parameter. Model A used the submission status of the first optional assignment, which was the second assignment in all courses. Among the 797 students from the three courses, 478 (60%) submitted the first optional assignment. A Chi-Square test found a statistically significant association between the first optional assignment submission status and the overall course success ($\chi(1) = 129.49$, $p = .000$)**.** The effective size of this finding, Cramer's V, was moderate (Cohen, 1988) and significant ($\phi = 0.403$, $p = .000$.). As shown in Table 3.2, most of the students who successfully completed the course submitted the first optional assignment, whereas most of the students who did not complete the course did not submit it. These results justified building the model based on the first optional assignment submission rate.

Table 3.3 Submission status of the first optional assignment (n=797).

	Assignment's Submission status		
Courses' Completion Status	**Did not Submit**	**Submitted**	**Total**
Successfully completed	20% (81)	80% (318)	100% (399)
Did not complete	60% (238)	40% (160)	100% (398)

The results of model A are presented in Table 3.4. The Wald statistic, defined as the square of a regression coefficient divided by the standard error of that coefficient (Osborne, 2015), was applied to determine the statistical significance of each independent variable. The logistic regression model for the entire sample (797) was found to be statistically significant $\chi^2(6) = 129.079$, $p < .001$. Following our expectations described above, the submission rate of the first optional assignment ($p< 0.01$) was found to be a significant parameter for predicting the final course success status, along with the advanced diploma ($p< 0.05$). The course name was found to be insignificant, justifying the analysis of the three courses as a single database. The model explains 22.5% (Nagelkerke R^2) of the variance in the courses' success and correctly classifies 70% of the cases. The model is well fit to data according to the Hosmer-Lemeshow test. These results suggest that starting at the 5th week, when students submit their first optional assignment, we can determine the probability that a specific student will complete the course.

After establishing the optional assignment submission status as a predicting variable, we analyzed the data according to the SCOP variable. We conducted the Mann-Whitney U test to assess the statistical association between SCOP and the course's success since most of the variables did not present a normal distribution. We found that starting at the second week, students that succeeded in the course received a statistically significantly higher score in the Mann-Whitney U test than students who did not succeed in the course ($p< 0.05$). We chose to use the SCOP of week 7 in the logistic regression because, before week seven, the rate of the explained variance and the percentage of correct classification were lower. Model B's results, which are based on the SCOP variable as a predictor, are presented in the two rightmost columns of Table 3.4. It was found to be statistically significant, $\chi^2(6) = 63.54$, $p < .001$, suggesting that one can identify the probability to succeed in the courses based on the following parameters that were found to be significant: The SCOP at the 7th week ($p< 0.01$) and having an advanced diploma ($p< 0.05$). Again, the course name was found to be insignificant in this model. The model explains 13.6% (Nagelkerke R^2) of the variance in the courses' success and correctly classifies 63% of the cases. The model is well fit to data according to the Hosmer-Lemeshow test. Both models indicate that early prediction models based on student data collected before the course's mid-point enable identifying students who will probably succeed as well as those who probably will not succeed and might need extra attention.

Table 3.4 Models of logistic regressions of courses' success. (N = 797)

	Model A (optional assignment submission)		Model B (SCOP)	
Variable	**Wald**	**Sig.**	**Wald**	**Sig.**
Advanced diploma	7.1	* 0.008	4.699	*0.030
Course	0.604	0.739	1.709	0.425
(SES)	1.686	0.194	1.288	0.256
First course at the OUI	1.044	0.307	2.432	0.119
Gender	3.348	0.67	3.405	0.065
SCOP at week 7	----	-----	45.699	**0.00
Submission of the first optional assignment	113.784	** 0.000	-----	-----

*p<0.05, ** p<0.01

Evaluation of Models

Table 3.5 compares the predicted classifications of students' final status in both models against the actual classifications. In each model, we defined a student with a probability of 0.5 or higher to succeed as a student who will probably successfully complete the course and a student below 0.5 as a student who probably will not complete the course. The correct overall predictions of model A are higher than those of model B. This means that model A is more accurate than model B and that the first optional assignments' submission is a stronger predictor than the SCOP since it can provide valuable predictions at an earlier stage and because the prediction is more accurate than the actual results. Both models predict success better than they indicate course incompletion. Nevertheless, the models predict incompletion correctly in more than 60% of the cases.

We further evaluated the models by plotting the AUC and ROC. See Appendix 2 for details. AUC values for models A and B were 0.731 and 0.683, respectively. These values approach 0.7, which is considered acceptable in scientific research (Gašević et al., 2016; Mandrekar, 2010). To check the robustness of the logistic regression result, we ran a hierarchical logistic regression in which the user specifies the order in which parameters are added to the model instead of entering them all at once (Osborne, 2015). The results remained the same, which strengthened our original models. In addition, we used a 10-fold cross validation to evaluate the average accuracy of the models. The average value of the AUC was 0.70 for Model A and 0.65 for Model B. This further strengthened our original models.

To further validate our strategy, we applied the analysis on new data from the same courses in semester 2020a. Using this new sample of 157 students, we obtained results that are similar to our original findings (see Appendix 2). These findings support our reliance on the first optional submission status at week five and the SCOP at week seven as strong predictors of students' success.

Table 3.5 Actual and predicted classifications of course completion. (N = 797).

	Model A Predictions			**Model B Predictions**		
Actual Status	**Improbable to Complete**	**Probable to Successfully Complete**	**Correct Predictions (%)**	**Improbable to Complete**	**Probable to Successfully Complete**	**Correct Predictions (%)**
Did not Complete	243	155	61	247	151	62
Successfully Completed	84	315	79	124	275	69
Overall Percentage			70			66

3.7 Limitations

At the OUI, we faced three limitations. The first was that we only had access to students' assignment submission status (submitted/not submitted) but not their assignment answers. Thus, we could not address research questions 2 and 3, which focus on learning outcomes. Secondly, because the available log file data details were limited, we could not evaluate goal setting, environment structuring, help-seeking, or self-evaluation. However, by assessing students’ choices regarding whether or not to submit optional assignments and their video playing patterns, we could indirectly learn about the dimensions of time management and task strategies. Finally, although we knew from the interviews that many students re-watched the course videos, we could not accurately assess these re-watching patterns. This was because our data on multiple viewing included both students who re-watched lessons and those that simply re-opened them due to technical issues. Therefore, we decided not to analyze this pattern using the EDM analysis.

3.8 Discussion

This chapter focused on undergraduate online general chemistry courses and aimed to identify indicators that predict students’ success based on their engagement and SRL theory. Two main research methods were used: students’ interviews and EDM techniques, including logistic regression analysis.

To address the first research question (Q1) – *What characterizes learners who are likely to complete online chemistry courses and those that are less likely to do so?* ,we used semi-structured interviews and EDM techniques. To measure students’ level of SRL, most studies utilized a self-report questionnaire (Hadwin et al., 2007; Barnard et al., 2009; Pintrich & Schrauben., 1992; Magno, 2011). Except for the OSLQ (Barnard et al., 2009), most of the questionnaires developed to assess SRL were not designed specifically to study SRL in the context of online learning. In addition, the field of online education has grown significantly since the OSLQ was first developed. Therefore, using interviews in this study was a helpful tool to learn how students regulate their learning in online chemistry courses. The new categories we found (see Figure 3.1) and the examples presented in Table 3.1 can enrich the existing engagement and SRL theory. Interviews with students who successfully completed the course revealed various learning patterns and time management strategies. Several students followed the course materials according to the weekly session plans, whereas others followed the course's assignment schedule.

Regarding the video recordings, we found that only a few interviewees actually attended the live sessions; the rest viewed the recordings at their convenience. In addition, students reported that they communicated with each other through a social media platform (WhatsApp group) that is outside the course. This platform was used for consulting with each other and for answering questions. This finding supports previous studies (Rap & Blonder, 2016; Rap & Blonder, 2017) that found that students use social media platforms to interact with each other and discuss the course materials. Finally, we found that the assignment submissions, both mandatory and optional, were also used as a learning strategy and for self-evaluation. These learning choices guided us in choosing the learning variables that could be used to construct a model to predict students' success in the courses, namely, opening video sessions and submitting optional assignments. This helped us develop prediction models and address research question (Q4) "*What is the earliest stage in the online course in which one can predict course success, and which course indicators are required to make these predictions?*

Model A indicates that we can already identify students with a high probability to successfully complete the course at week 5 with the submission of the first optional assignment. This finding indicates that the optional assignment, which we view as a proxy of student choice, is an essential predictor of course completion. This expands on previous studies that found that the more assignments students completed on time and the earlier that they did so, the better they performed on quizzes and final exams (Li & Baker, 2018). Model B showed that students who eventually successfully completed the course had different video opening patterns than those who did not succeed in the course by week seven. The SCOP variable, which is the primary predictor in this model, is an indicator of students' engagement. Indirectly, it also indicates time management since it reflects their advancement in the course from week to week.

Model A is a stronger predictor, of course, success than model B. This can be understood considering that submitting an assignment better represents active learning than does opening a video (Gabbay et al., 2020; Glick et al., 2020). By active learning, we mean that participants are dynamically or experientially involved in the learning process, which is known to be a more important feature of successful online learning (Davis et al., 2018). Future research should examine whether embedding active learning features within video sessions is a stronger predictor of student success.

Finally, this chapter also contributes to the existing literature on SRL in an online learning environment (Aleven et al., 2010; Rakes & Dunn, 2010; You & Kang, 2014) and also contributes to specific research that examines the potential to predict success early via analysis of log file data (You, 2016). As shown in previous studies, one of the difficulties in developing incomplete persistence predictors from online courses has been the inability to detect dropout early enough to prevent it (Costa et al., 2017). The predictive models developed in the current study detect students with a high probability of not completing the course before the middle of the course. These results allow designing more effective interventions and scaffolds for students' learning.

3.9 Summary and Implications

Based on the interviews and the indicators for predicting early success in an online course, we wish to highlight several implications for lecturers, institutions, and students. It is essential to focus on students' learning strategies and their development. SRL is developed over years of learning experience in elementary, high-school, and post-secondary education. However, students can still develop these skills later on in life. Developing SRL skills can help them in future academic settings, especially with the growing importance of online learning and life-long learning (Pintrich, 2000; Zimmerman, 2008; Taranto & Buchanan, 2020).

Our results also have implementations for teachers and faculty. Previous research has shown that instructions alone cannot efficiently implement SRL strategies (Nawrot & Doucet, 2014). Therefore, lecturers who want to help students develop their SRL skills should make an effort to integrate the development of SRL skills into the context of the course they teach. For example, a study of chemistry courses showed that a workshop for learning strategies could help students improve their performance in the course (Cook et al., 2013). The chapter findings can be used to develop such workshops to guide students regarding their own responsibility for self-learning, for help-seeking at the early stages of the course, and the importance of proper time management and their choices during the course, particularly regarding answering and submitting the course assignments. Instructors should consider implementing pedagogy that enables students' choices such as optional assignments to identify students' current status during the course.

As for institutions, although the research findings highlight the potential of early prediction of the probability to succeed in an online course, we emphasize that this should be done carefully and accurately. Many institutions aim to automate this process by developing and implementing informative dashboards for instructors to help them monitor students' progress and acquire insights from this information (Ahn et al., 2019; You, 2016; Michaeli et al., 2020). This includes, for example, timestamps of each activity, counts of entrances to the website or specific files, content added by users (e.g., in a forum), and more (see chapter 3). However, drawing conclusions about the students' learning status in the course or the probability of successfully completing the course directly from this kind of dashboard should be done cautiously. Using this kind of data for statistical analysis requires a pre-processing phase that creates a reliable database and prevents bias in the activity trends (Romero et al., 2014). The models presented in this study were developed only after we implemented such a methodologic pre-processing phase.

Moreover, the prediction was based on a combination of three different sources: Moodle log files, course grades, and demographic data. This combination of data is generally not presented in the dashboards mentioned above. Finally, one should bear in mind that the raw data from the log files of online activities, such as the number of total page views and the frequency of students' login, provide little insight into why students complete an online course or withdraw from it (Li et al., 2020).

That being said, we recommend that creating new dashboards for a specific course be done, along with an evaluation of the relevant courses; this will involve both researchers and the course staff. This evaluation would guide them in choosing the most pertinent SRL indicators. Academic institutions could also consider embedding an automatic weekly statistical analysis in a dashboard that will present lectures, along with the probability of students’ success in the course.

4. A Multi-Dimensional Course Evaluation Framework for Online Professional Development of Chemistry Teachers

4.1 Highlights

- We developed a three-dimensional evaluation framework for teachers' online PD courses.
- We used this framework to evaluate a nanotechnology PD course.
- We assessed learning outcomes and analyzed Moodle activity reports.
- We showed how teachers expand their knowledge and skills on topics not part of the high-school science curriculum.
- We identified teachers' difficulties during the online course.

4.2 Introduction

The previous chapter focused on undergraduate chemistry students at the OUI. This chapter deals with teachers who participated in a PD course on nanotechnology at the Weizman Institute of Science. In contrast to the OUI, at the Weizman Institute, we had access to teachers' course assignments. These data allowed us to analyze teachers' assignment answers and thus address research questions 2 and 3, focusing on learning outcomes.

Importantly, we developed a multi-dimensional evaluation framework for online PD courses. This framework combines EDM techniques with more traditional evaluation tools and allows one to evaluate learners' knowledge, their complexity of understanding, and identify their difficulties during the course. We applied this framework to assess the online nanotechnology course for teachers' PD at the Weizmann Institute of Science. Combining the traditional evaluation approach with EDM techniques provides a more comprehensive assessment of learning outcomes and difficulties. First, we will discuss relevant literature on teachers' PD, nanotechnology education, and the SOLO taxonomy.

4.2.1 Science Teachers' Professional Development and Online Courses

Science teachers enroll in professional training courses for various reasons (Mamlok-Naaman et al., 2018). These include learning content updates (Blonder, 2011), meeting government requirements, and advancing their careers (Hofstein et al., 2003; Taitelbaum et al., 2008). Designed for adult learners, teachers' PD courses are typically based on the andragogy theory. According to this theory, adults learn better when they understand why they are required to learn certain topics (Morland & Bivens, 2004). Therefore, courses designed specifically for teachers' PD should aim to advance their knowledge and skills in their relevant field of expertise, which they can later implement in their teaching (Mamlok-Naaman et al., 2018; Shulman, 1987; Jones et al., 2020).

The online environment offers a convenient format for nearly all adult learners. This convenience is due to its time flexibility and accessibility that meet adults' preference for open learning with no time and distance hindrances between them and the learning sources (Milligan & Littljohn, 2014). In the last few years, the number of online PD courses has been growing rapidly (Milligan & Griffin, 2016; Salmon et al., 2015). During the Covid-19 pandemic, this format became dominant and replaced all other forms of teachers' PD (Hartshorne et al., 2020). Although researchers have highlighted the importance of assessing how online PD courses affect learning outcomes and their relevance to the learners' professional work experience (Egloffstein, 2018: Milligan & Littlejohn, 2014), this need has not been fully addressed.

4.2.2 Teachers' Professional Development in Nanotechnology

The emerging field of Nanotechnology (Jackman et al., 2016) has been integrated into the high- school chemistry curriculum via various elective units and learning activities (Delgado et al., 2015). This integration is particularly challenging since, at the nanoscale, matter can have different properties at both the molecular level and the macroscopic scale, giving rise to the unique functionality of nano-materials (Jones et al., 2013). Nanoscale science and technology (NST) deals with the ability to create materials, devices, and systems with fundamentally new properties and functions by exploring their structure at the atomic, molecular, and macromolecular levels (Roco, 2001). NST is an interdisciplinary field that combines content knowledge from chemistry, biology, physics, materials science, medicine, and engineering (Yonai &

Blonder., 2020). To address the uniqueness of this field, numerous nanotechnology PD courses that aim to introduce teachers to the nanoworld have been developed around the world (Jones et al., 2013; Dori et al., 2014, Lin et al., 2015; Sgouros & Stavrou; 2019). In this study, we examined one PD course that was designed according to the list of eight essential NST concepts defined by Sakhnini & Blonder (2016). The list was compiled after implementing a three-round Delphi-study methodology to reach a consensus between experts in nanotechnology regarding the essential NST concepts that should be taught in high school. These essential NST concepts are as follows:

1. *Size-dependent properties* refer to properties that change as a function of the material's size according to the high ratio of surface area to volume (SA/V) in the nanoscale and are based on fundamental quantum mechanical principles.

2. *Size and scale* are used to characterize the extent or amount of an object (size), and to compare it to other objects (scale).

3. *Characterization methods* are used to study the properties of nano-materials and nanosystems. This concept includes tools for observing, imaging, learning, and manipulating the nano-material size, for example, a) Scanning Probe Microscopy (SPM); b) Electron Microscopy (EM), including Transmission electron microscopy (TEM) and Scanning electron microscopy (SEM).

4. *Functionality* transforms nanoscience into nanotechnology. A certain property endows the material with a specific activity.

5. *Classification of nano-materials* includes the chemical composition, electrical conductivity, source, and dimensionality. Here dimensionality is used to classify nano-materials according to the number of dimensions (0-3) in which a nanostructure expands beyond 100 nm.

6. *The fabrication approach of nano-materials* can be divided into top-down and bottom-up approaches. Top-down approaches locate each component of the material from the top, such that the arrangement of the material is determined by external intervention (e.g., lithography) at the scale of the resulting nano-material. In contrast, in the bottom-up approaches, the molecules or atoms in the gaseous phase or in solution are arranged in a pre-defined set of structures and directionality, sometimes on a specific platform. A leading example is a self-assembly, which describes the ability of molecules to arrange themselves into ordered structures 'on their own' while satisfying the laws of thermodynamics.

7. *Innovations and applications of nanotechnology* include a) Current and future applications; b) Techniques for mimicking nature; c) Risks and benefits of nanotechnology; and d) Tailoring nano-materials to a specific application.

8. *The making of nano-materials* is related to the nature of scientific research, i.e., how nanoscience research is performed and how innovations are transformed into applications.

Most in-service science teachers are unfamiliar with nanotechnology since it was not widely taught in undergraduate programs when they were students, and it is not part of the school curriculum (Jones et al., 2020). Jones et al. (2020) claim that this necessitates offering a nanotechnology PD course for teachers, updating their knowledge, and preparing them so that they can integrate contemporary scientific research into the existing science curriculum. Blonder & Mamlok-Naaman (2016) found that following a PD course in nanotechnology, some science teachers could implement the teaching methodologies to which they were exposed in the PD course when they later taught chemistry to their students. However, one of the significant challenges of NST education lies in transferring this contemporary field to teaching practice and integrating it into the school curriculum (Sgouros & Stavrou, 2019). For this purpose, the development of appropriate materials for teaching is required. Studies have shown that it is important to engage teachers and support them in the process of interpreting and integrating nanoliteracy in a meaningful way into their teaching (Mamlok-Naaman et al., 2010; Jones et al., 2020). This raises the need to evaluate the effectiveness of these processes at the content level. For this purpose, we applied here the SOLO taxonomy.

4.2.3 The SOLO taxonomy

The 'SOLO' taxonomy, initially developed by Biggs and Collis (1982, 1989), classifies learning outcomes in terms of their complexity. It provides a systematic way to describe the range of performances produced by learners in a specific academic activity, such as writing an essay or answering an open-ended question (Minogue & Jones, 2009). The SOLO taxonomy describes five levels of complexity: 'Pre-structural,' 'Uni-structural,' 'Multi-structural,' 'Relational,' and 'Extended Abstract'. These levels of complexity are organized by various characteristics, including movement from the concrete to the abstract, the use of an increasing number of organizing aspects, increasing consistency, and relating to and extending key principles (Biggs & Collis, 1982; Biggs & Collis, 1989) according to the following assessment system: 'Pre-structural' responses indicate no understanding. At the 'Uni-structural' level, learners can choose only one aspect of a task. Dealing with several unrelated aspects is termed 'Multi-structural.' When learners integrate a few aspects into a whole, the level is termed 'Relational.' Finally, if learners can generalize and transfer aspects of a task to different contexts, their level of learning complexity is termed 'Extended Abstract'. It is assumed that assignments can be answered in a way that reveals the complexity of the learners' understanding (Alexandron et al., 2016; Lister et al., 2006; Tsaparlis et al., 2018). In this study, we used the SOLO taxonomy to evaluate teachers' ability to apply the content they learned in the course to the chemistry curriculum and an unfamiliar nanotechnology application. The SOLO levels are shown in Table 4.1 in the Methods section.

4.3 Research Questions

Online video-based learning could be a passive learning experience (Brame, 2016). However, video lessons are a central resource in online courses (Johnson et al., 2014; De Waard et al., 2012). To bridge this gap, online course designers need to think about how they can keep learners engaged in the course. Indeed, many studies focus on evaluating the level of learner engagement (Baldwin et al., 2018). Although this is no doubt important, this approach focuses less on evaluating learning outcomes, knowledge, and understanding in the context of online learning. To address this shortcoming, we focus on knowledge evaluation and the learners' complexity of understanding, and the difficulties they encountered.

In this chapter we focus on answering research questions 2 and 3:

Q2. How can we evaluate learning outcomes in the context of online learning?

Q3. How can we identify learners' difficulties in the online course?

To address these two questions, we developed a framework to evaluate learning outcomes and identify learners' difficulties in the online course. For this purpose, we defined two additional sub-questions regarding teachers' improvement of knowledge and their level of complexity of understanding:

Q2a) Did teachers improve their knowledge of nanotechnology?

Q2b) To what level of complexity did teachers develop their understanding of the NST concepts?

4.4 Research Set-up and Participants

4.4.1 Course Design

In 2008, a face-to-face course on nanotechnology was developed at the Weizmann Institute of Science as part of a master's degree program for science teachers (Blonder, 2011). To reach a wider population of chemistry teachers, the course was redesigned and converted to a Small Private Online Course (SPOC) called 'Introduction to Materials and Nanotechnology.' The course was designed and given by one of the advisors (Prof. Ron Blonder) and coordinated by a doctoral student. This course exposed chemistry teachers to six out of the eight NST essential concepts described above: *Size-dependent properties, Size and scale, Characterization methods, Classification of nano-materials, the fabrication approach to nano-materials*, *Innovations, and applications of nanotechnology*. During the course, teachers were asked to find appropriate connections between the six main NST concepts and the high-school chemistry curriculum. The two other concepts, "*Functionality and the making of nano-materials,"* were part of the course but were not taught explicitly since they require exposure to specific laboratory techniques (Blonder & Sakhnini, 2015; Akerson et al., 2000).

The course was presented in a Moodle environment; it included 13 pre-recorded video lessons, each comprising up to five 25-minute-long videos. A short quiz with automatic feedback followed each video. Other tools included a Padlet board (an online collaborative bulletin board) and discussion forums to create an active learning environment. It also integrated one face-to-face (F2F) session that included a laboratory experiment and a visit to an SEM facility at the Weizmann Institute to characterize the experimental products. A new lesson was opened each week, and teachers could proceed according to their schedule. However, the final assignment had a deadline at the end of the semester.

The course was organized according to the NST content model (Sakhnini & Blonder, 2015). Accordingly, most of the lectures were devoted to the scientific aspects of each concept and its technological applications. The NST concepts outlined in the lectures were not presented with a direct connection to the high-school chemistry curriculum. However, the knowledge included in the high-school chemistry curriculum was used to explain each concept. For example, when the concept of *fabrication of nano-materials* was presented, the lecturer presented, as an example, the chemical reaction of oxidation-reduction in the synthesis of gold nanoparticles according to a bottom-up approach. During the course, teachers were asked to suggest where and how NST concepts can be linked to the chemistry curriculum and were requested to upload their answers to a Padlet board. In the final course assignment, teachers were asked to choose a nanotechnology application that was not mentioned in the course and explain it in terms of three NST concepts they had learned in the course. In addition, each teacher was required to read three assignments by their peers and provide feedback.

4.4.2 Learning Outcomes

Three main learning outcomes were proposed for the course: 1) The teachers will become familiar with the six nanotechnology concepts according to the NST essential concepts model, 2) Teachers will know how to describe nanotechnology applications according to the NST essential concepts, and 3) Teachers will be able to connect NST concepts to the chemistry curriculum.

4.4.3 Participants

Ninety–five teachers participated in the online Introduction to materials and nanotechnology course (see Table 2.5 for the teachers' characteristics).

4.5 Methodology

In order to address the research questions, we applied qualitative and quantitative tools (a mixed method) as detailed below. We selected tools that are appropriate for the course design, such that each tool enabled us to evaluate a different dimension. We evaluated teachers' improvement in knowledge using a pre-post questionnaire and assessed the teachers' complexity of understanding level by means of the 'SOLO' taxonomy. The LMS log files were analyzed to identify patterns of video learning. The interviews helped us see that these patterns can explain learners' difficulties in specific topics.

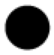
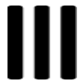
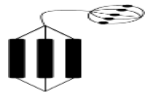
4.5.1 The Knowledge Pre-post Online Questionnaire

A knowledge pre-post online questionnaire was embedded in the course's LMS to determine whether the teachers had advanced their knowledge of nanotechnology (see Appendix 3). The questionnaire was given twice: at the beginning of the course and again at the end. The participants were asked to explain a list of concepts related to the NST concepts. The participants' answers were graded as follows: 0, for a wrong answer or no answer; 1, for a partial answer; and 2, for a full answer. We compared the pre-post responses in the online course.

4.5.2 Content Analysis Using the SOLO Taxonomy

The SOLO taxonomy was applied for two purposes: (1) To evaluate the teachers' ability to connect the field of nanotechnology to the chemistry curriculum, and (2) to evaluate teachers' final course assignments in terms of complexity. The SOLO classification analysis in this study was conducted as follows: The definition of the SOLO levels in the current study (Table 4.1) was first suggested by the doctoral student according to the course's structure and content. Next, these definitions were validated by two of the advisors until a consensus was reached regarding the analysis protocol and the definition of each SOLO level.

Table 4.1 Defining the SOLO taxonomy levels in terms of the structure and content of the nanotechnology PD course.

	Pre-structural	**Uni-Structural**	**Multi-Structural**	**Relational**	**Extended abstract**
Level Symbol					
Level of connecting the NST concepts to the school chemistry curriculum *	Did not specify a connection to the chemistry curriculum	Mentioned one connection to the chemistry curriculum and did not explain its relationship to the NST concepts	Mentioned connections to several topics from the curriculum but did not explain their relationship to the NST concepts	Mentioned one or more topics from the curriculum and explained the relationship to the NST concepts	Generalized and transferred aspects of an NST concept to different curricular topics and contexts beyond the course contents

* A similar analysis was done for the final course assignment

4.5.3 EDM techniques

EDM methods: Moodle course activity reports were analyzed to learn how the video lessons were used. A pre-processing phase was performed on the log files from the LMS to enhance the data reliability, to clarify the procedures required for working with raw or partially processed data, and to avoid the pitfalls of working with inadequately processed data (Romero et al., 2014; see chapter 2). Data were analyzed according to unique user activities, meaning that when a specific video lesson was considered, we counted the number of users that played that video at least once.

4.5.4 Semi-Structured Interviews

Semi-structured interviews: To understand teachers' learning habits during the course, we conducted nineteen semi-structured interviews with participants in 2017 and 2019 (Two of them were pilot interviews). Following the completion of each course, we sent an email to the participants to ask for volunteers for the interviews. Each interview lasted 20–60 min and was audio-recorded and transcribed by a doctoral student. The questions of the semi-structured interviews were validated by the two advisors. In addition, we conducted three additional interviews with participants from the 2016 course to learn about the long-term use of the course website after the course had ended. We used the interview data to better understand the learning pattern found from analyzing the Moodle activity reports. This chapter does not present the full results of the qualitative data analysis. Instead, data collected during the interviews are shown as selected utterances used to enrich and explain the quantitative data with illustrative verbal descriptions (Dorfman& Fortus., 2019). A complete analysis of the interview is presented in chapter 5 using case studies.

4.6 Results

Next, we will present the results according to the dimensions of the evaluation framework.

4.6.1 Advancing Nanotechnology Knowledge

Chemistry teachers' knowledge improvement was evaluated using an online pre-post-questionnaire. Analysis of the questionnaires revealed a significant improvement in the teachers' conceptual understanding of nanotechnology. Figure 4.1 presents the average scores for each item in the pre- and post-questionnaires. The NST concepts in Figure 4.1 are listed according to the order that they appeared in the course. Since the variables were not distributed normally, we decided to analyze the data by using non-parametric tests. We applied the Wilcoxon signed-rank test to the overall difference between the average pre- and post-questionnaire answers. A significant improvement in the teachers' understanding was found in five out of the six NST concepts ($p < 0.05$) that were discussed in the course (Figure 4.1).

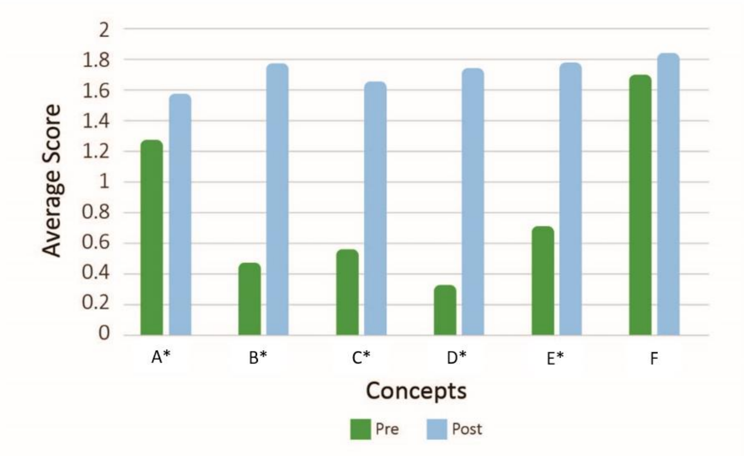


Figure 4.1 Pre-post knowledge questionnaire (n=41). Green – pre, Light Blue - post. List of concepts: A. *Size and scale*, B. *Size-dependent properties*, C. *Characterization methods*, D. *Fabrication of nano-materials, E.* Innovations and applications of nanotechnology, and F. Classification of nano-materials. *p<0.05

4.6.2 Complexity of Understanding

Teachers' level of complexity of understanding was evaluated using the 'SOLO' taxonomy regarding two aspects presented next.

1. Connecting NST concepts to the school chemistry curriculum

After learning each of the following concepts: *size and scale*, *size-dependent properties*, *characterization methods,* and *fabrication of nano-materials*, teachers were requested to suggest connections for the NST concepts in the high-school chemistry curriculum and were requested to post their answers on a Padlet platform. In these online Padlet assignments, the teachers were required to think about applications of the course content that are relevant to their day-to-day work (Salmon et al., 2015). Between 33 and 51 teachers participated in each Padlet session. To evaluate these assignments, first, we identified the number of times a specific concept was connected to a particular topic from the chemistry curriculum. Importantly, we found that 90% of the insertion points confirmed the results of a previous study by Blonder & Sakhnini (2016). To understand how teachers connect specific NST concepts to each topic in the high-school curriculum, the teachers' responses were evaluated using the SOLO taxonomy (Biggs & Collis, 1982; Biggs & Collis, 1989). Figure 4.2 shows the distribution of each of the SOLO levels according to each NST concept. As shown in

Figure 4.2, most of the teachers reached the multi-level and the 'Relational' level (Figure 4.2).

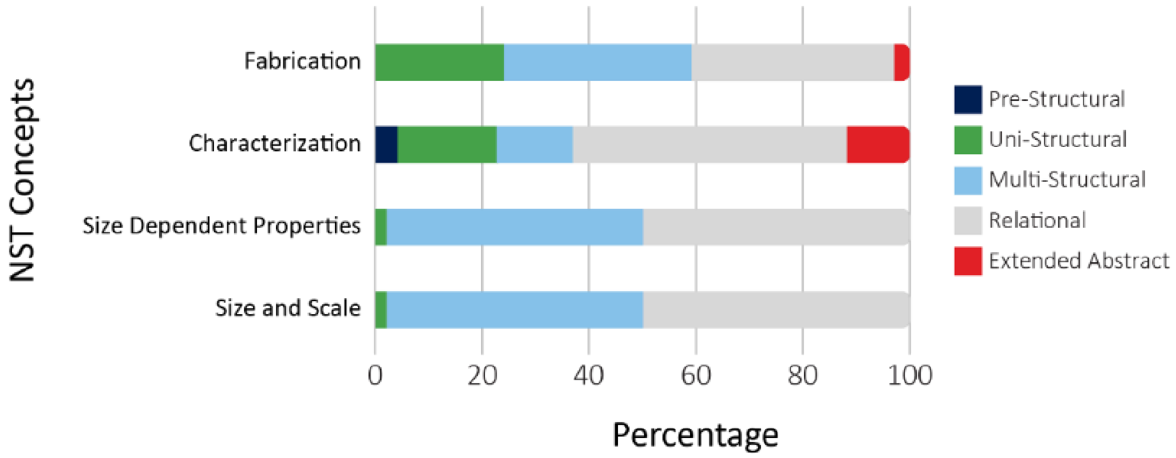


Figure 4.2. Distribution of teachers' level of complexity according to the SOLO taxonomy.

The following examples demonstrate the SOLO classification of several connections that were suggested by the teachers regarding the third NST concept of *characterization methods* (SPM, EM and Resolution):

- *Pre-structural*: Dana answered, *"I do not think I would incorporate the use of SEM during my teaching. Perhaps only as an enrichment lecture on its structure and possible usage."* We classified the teachers' answers as 'Pre-structural' when they did not specify a connection to the chemistry curriculum and did not provide an explanation as to why.
- *Uni-structural*: Michelle answered, *"I would incorporate learning about the SEM when teaching metallic bonding."* This answer was classified as 'Uni-structural' because the teacher mentioned one connection of the concept *characterization methods* to the chemistry curriculum (metallic bond) but did not explain it.
- *Multi-Structural*: Adi answered," I *suggest explaining about the SEM in a few topics: atomic structure, molecular geometry, and physical chemistry."* We classified teachers' answers as 'Multi-structural' when they mentioned several topics from the curriculum but did not explain the connection to the NST concepts.
- *Relational*: Miri answered," I *would refer to SEM only at the end of the atomic structure unit. At this stage, students should be able to internalize the differences between atoms, molecules, etc. Then, we can use SEM images for*

demonstrations". Teachers' answers were classified as 'Relational' when they mentioned several topics from the curriculum and explained the connection between them and the NST concept.

Extended abstract: Rachel answered,

> *"Nanotechnology is not mentioned in the 10th- and 11th-grade curriculum, but as teachers, we have to talk about it. Because of its importance as a scientific development in the world, …I first teach the basic nanotechnology concepts in the 9th grade, primarily for enrichment. Then, when I teach Van der Waals forces, I discuss how we can see nano-materials. I relate the explanation of the Van der Waals forces to the principles of operating the AFM microscope. Sometimes after I teach Van der Waals forces, I build together with my students the microscope model from clay and bole caps to demonstrate the topic."*

This answer was classified as 'an Extended Abstract' since the teacher generalized and transferred aspects of a task to different curricular contexts and provided a pedagogical explanation for her decision.

2. Evaluation of the final course assignments:

Note that the course's final assignment was constructed in such a way that comprehension complexity is presented at the 'Relational' level. The analysis of 143 concepts from 50 final assignments is presented in Figure 4.3. In the final assignment, teachers were asked to use three NST concepts and connect them to the nanotechnology application they chose to present. As shown in Figure 4.3, teachers chose from the five central concepts that were taught in the course (they were asked not to use the "size and scale" concept). The complexity of their usage ranged between 'Uni-structural' and 'Relational.' Only one teacher exhibited an 'Extended Abstract' level. This means that almost none of the teachers reached the highest level of complexity. On the other hand, no one exhibited the lowest 'Pre-structural' level either.

The following are examples of our analysis of the concept of "*surface area to volume ratio*" (a sub-concept of the "*size-dependent properties*" NST concept) that appeared in the final assignment according to the SOLO taxonomy. Owing to ethical considerations, a detailed description of the applications the teachers chose to present was excluded.

- *Uni-Structural*: One of the teachers, Tal, explained that nano-materials have a large surface area relative to ordinary materials. This most likely means that she understood the concept learned in the course but did not associate it with other concepts nor to the application she decided to explain in the final assignment.
- *Multi-Structural*: Another teacher, Idit, explained the concept of surface area and its dependence on the nanometric scale. She also explained the application she chose but without associating it with the definition of the concept or the example from the application. That is, the teacher understood the concept and the nanotechnological application but did not exhibit the ability to associate them with each other.
- *Relational*: At the relational level, the concept of surface area was defined, and the definition was related to an application. As Diana describes:
  "In nanotechnology, there is special emphasis on the ratio between surface area and volume because this ratio impacts the nanoparticle elements, for example, the melting point. Unlike non-nano materials, in which the melting point is constant, in nano-materials, different melting points depend on the size or, more accurately, on the ratio between the surface area and the volume. In the application I studied, the surface area of the nanocapsule influences the effectiveness of the medication."
- *Extended Abstract*: Only in one assignment did the complexity of understanding reach this level. The explanation of the teacher, Rachel, was beyond what was required in the course assignment, and the level of details indicated that she had achieved an extensive complexity of learning. This teachers' assignment was about using nanotechnology for the domesticated transfer of a drug for the treatment of cancer. The teacher, Rachel, described the context of a surface area: *"The expansion of the drug's surface-to-volume ratio was manifested in a number of stages during the transition process: 1) Decreasing the storage space of the drug molecules. 2) Increasing the surface area that "sticks" to the cancer cell by the nanocapsule coating, and 3) Increasing the surface area of the nanoparticles by molecules that are partly hydrophilic and partly hydrophobic, which leads to an increase in the surface area of the drug for antibody-specific binding to the cancer cells".*

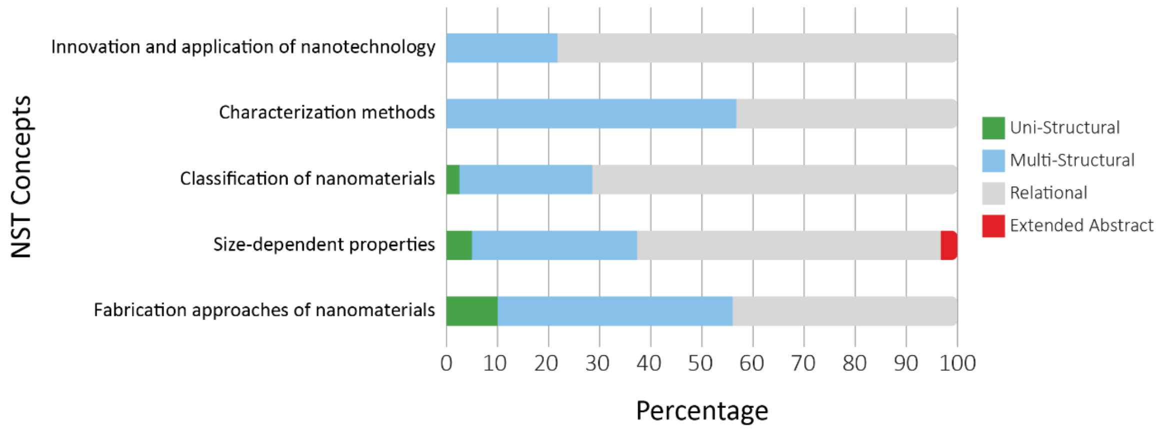


Figure 4.3 Distribution of the appearance of NST concepts in the final assignment. Colors represent the complexity level according to the SOLO taxonomy.

4.6.3 Evaluation of Learners' Difficulties

In addition to the learning outcome evaluation, another dimension of the evaluation framework is learners' difficulties. Here we present the results based on combining the semi-structured interviews and the EDM techniques. In the next chapter, we expand the use of these tools by examining learners' learning patterns in the online course.

The pre-recorded video lessons were the main learning resource for the online course. Since the course and the target population did not change during the three cycles in which the course was given, we combined data regarding teachers' video playing from all three courses' cohorts. Figure 4.4 presents the average percentage of unique playing of each video. Each bar represents the percentage of users who played the first video of each lesson at least once. We noted that in lesson 4.4, the teachers were asked to read an article and had no video lesson; therefore, it does not appear in Figure 4.4. Figure 4.4 shows a decrease in the video playing up to lesson 8, followed by a slight increase in the video playing in lessons 9 and 10. Finally, the video playing level stabilized during the last four lessons.

A possible explanation for the low participation in lessons 7 and 8 was revealed from the interviews. Most teachers reported that the topic of quantum mechanics studied during those lessons was very difficult for them, causing some of them to skip those lessons. Tanya explained: *"I ended up giving up the photoelectric effect lesson. It was difficult for me. I didn't have enough time…"*. Another teacher, Alma, said*: "There were parts in the quantum mechanics lessons where I played the video a little faster …"*. On the other hand, other participants chose to watch these lectures several times to reach a better understanding. Note that since we considered unique video playing as a factor, this did not affect Figure 4.4.

Analysis of the Moodle activity reports revealed that 43% (n=41) of the teachers returned to the course website for a few months and even three years after completing the course*. For example, Talia explained: "I use other teachers' final assignments posted on the course website to guide my students*." Another teacher, Romi, explained*, "I use the course website to find examples of the development of science to show my students recent innovative scientific discoveries."* The interviews showed that teachers returned to the course website from time to time in order to refresh their knowledge, explain an NST concept to a fellow teacher, or prepare classroom materials. Some teachers also applied the technological tools (e.g., video and the Padlet board) that they had worked with during the PD course in their chemistry teaching.

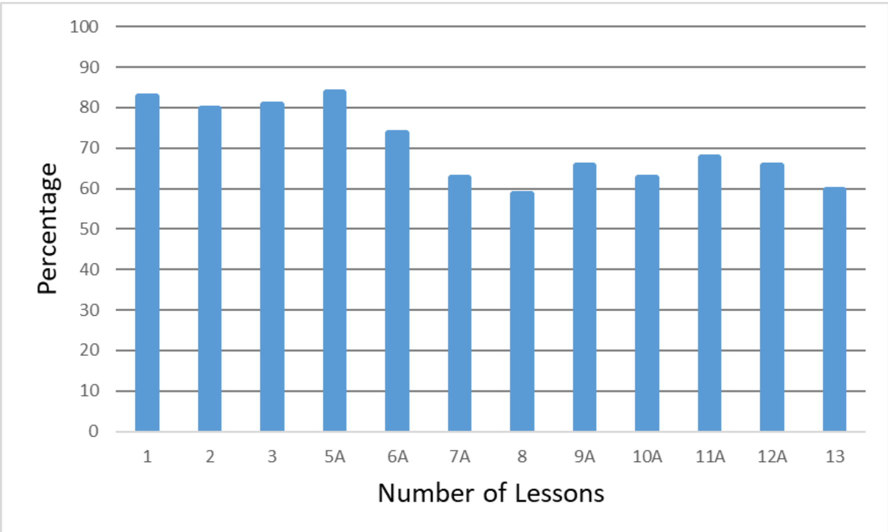


Figure 4.4. Percentages of unique video playing for each lesson

4.7 Limitations

The number of respondents was smaller than the number of total graduates in the three cohorts. This is likely because the post-questionnaire was not mandatory. Furthermore, teachers who did not complete the course did not answer the post-questionnaire. Sixty-one teachers (64%) completed the course and submitted the final assignment; 50 assignments were analyzed according to the SOLO taxonomy. An additional 11 assignments were excluded from the analysis because the NST concepts were not mentioned in their text. We must, however, take into account one limitation of using the SOLO taxonomy: it does not guarantee an accurate and complete account of what was learned (Minogue & Jones, 2009). We tried to overcome this limitation by having one of the advisors validate the content of the teachers' answers. We further addressed this weakness with the use of our multi-dimensional evaluation framework.

4.8 Discussion

The main goal of this chapter was to evaluate learning outcomes and identify learners' difficulties. The following section presents our answers to the chapter's sub-research questions by discussing the results drawn from the different research tools we applied.

Q2a) Did teachers improve their knowledge of nanotechnology?

The results of the pre-post questionnaires indicate that teachers' understanding, and conceptual knowledge of nanotechnology had improved with regard to all NST concepts, except for the *classification of nanotechnology materials*. One way to account for this finding is that teachers were already familiar with this concept because it is similar to their understanding of the Periodic Table. This assumption is well-founded since, as shown in Table 4.1, most teachers received a high score for this *concept in the pre-*questionnaire**.**

In the current study, the improvement of knowledge was only one of the criteria used to evaluate the effectiveness of the online course. Focusing on the advancement of knowledge is not enough. Teachers' PD programs should also guide and support teachers in implementing what they learned in their school curriculum (Blonder & Mamlok-Naaman, 2016; Loucks-Horsley & Matsumoto, 1999). The evaluation criterion addressed was the complexity of understanding the NST concepts and the ways teachers connect them with the chemistry curriculum.

Q2b) To what level of complexity did teachers develop their understanding of the NST concepts?

The Padlet boards were used to create an active learning environment and to enable engagement. As demonstrated in the analysis, the course design allows teachers to apply new content obtained in the PD course in their classes during and immediately after the course. We found that they could tailor the integration of NST concepts to their teaching. We thus concluded that to support teachers, course designers should consider providing the participating teachers with opportunities to connect advanced science topics to the school curriculum during the PD courses. It is important that the teachers themselves will make this connection since they are most familiar with their school context and the specific characteristics of their students (Gess-Newsome, 2015; Dorfman et al., 2019).

Using the SOLO taxonomy, we found that the online nanotechnology course brings most teachers from the ‘Multi-structural’ to the ‘Relational’ Level. Recall, however, that the course’s learning outcomes did not aim to advance all teachers to the ‘Extended Abstract’ level. Therefore, teachers who achieved the ‘Relational’ level in the Padlet assignments and in the final assignment met the course’s learning outcomes. The few teachers who reached the ‘Extended Abstract’ level surpassed the intended course target.

Using the SOLO taxonomy as a design tool and not only as an evaluation tool might help teachers reach a higher level of complexity. This idea was suggested by Biggs & Tang (2015), who claimed that designing the curriculum according to the level of the intended learning outcomes using the SOLO taxonomy would help implement a constructive alignment. Namely, educators should consider the outcomes they intend their learners to reach and align teaching and assessment according to these outcomes.

Another possibility we suggest examining in future research is to teach those enrolled in the PD course how to use the SOLO taxonomy to monitor their own progress during the course. According to Tan et al. (2008), the SOLO taxonomy can serve as a cognitive organizer (a type of learning strategy) that can enhance the level of SRL. Although there is always some degree of subjectivity in assessing open-ended tasks (Minogue & Jones, 2009), the SOLO taxonomy represents only one dimension of the evaluation framework. The accuracy of understanding the NST concepts was also evaluated in the post-questionnaire.

Q2c) How can we identify teachers’ difficulties in the online course?

Teachers indicated that the materials in the less frequently opened lessons were more difficult for them. This caused some of them to skip these lessons or drop out of the course entirely. Despite that this pattern was discovered after the courses had ended, in future courses, similar activity reports can be analyzed during the course and can be used to ascertain teachers’ difficulties on the fly; thus, effective interventions can be provided.

An additional key finding in this study was that the course website serves as a valuable source of information for teachers, allowing them to continue learning after they have completed the course. This finding could also address a need highlighted by Jones et al. (2020), who contend that the rapid development of nanotechnology requires teachers to continue learning about new developments in the field following the PD course. Our findings suggest that by keeping the course website open after the course has been completed, teachers' continuing development can be supported and sustained. The next chapter further examines this pattern by analyzing how teachers plan to implement the scientific knowledge and pedagogy they acquired in the online PD course.

4.9 Summary and Implications

The increase in online courses in the last two decades, and particularly during the COVID-19 pandemic, poses a challenge to evaluate these courses effectively. With the emergence of big data in the educational field, many researchers have evaluated online students' activities using EDM techniques. Along with its many advantages, such as recognizing content that students tend to skip over, it combines several evaluation tools that can provide new and important insights into the learning process and can facilitate interpreting online activity reports. In this research, we developed and demonstrated the use of a multi-dimensional framework to evaluate online PD courses. This framework is based on evaluating knowledge, the complexity of understanding, and identifying learners' difficulties.

Using this framework, we showed how chemistry teachers expanded their knowledge and skills on topics not part of the high-school science curriculum. Using the SOLO taxonomy, we could evaluate the complexity level of the connection that teachers found between the NST concepts and the chemistry curriculum. Using the EDM techniques, we analyzed the frequency and timing of learning resource use. Along with the analysis of the interviews, we could identify teachers' difficulties. In this respect, we showed that the whole is greater than the sum of its parts. Using this evaluation framework, we learned about teachers' learning outcomes and how we can design better online PD courses in the future.

The evaluation framework was tested on a nanotechnology PD course for chemistry teachers; however, it is by no means limited to these types of courses, and it could be applied to a variety of online PD courses. When this framework was applied, an analysis of activity reports could be conducted during the course to detect learners facing difficulties during the course; this would allow lecturers to conduct relevant interventions. We wish to emphasize that the developers of online PD courses for teachers should strive to maintain a strong link between content and the existing curriculum and remain available to advise teachers long after the course has been completed.

We evaluated the online course “Introduction to Materials and Nanotechnology” in terms of learning outcomes and difficulties. In assessing teachers’ difficulties in the course, we demonstrated the benefits of integrating EDM techniques and interviews. This chapter focused on course evaluation, but it did not distinguish between learners who successfully completed the course and those who did not. In the following chapter, we differentiate between these two groups in a manner similar to our analysis in chapter 3. We also continue to discuss the benefits of combining interviews and EDM techniques to evaluate and characterize teachers’ learning patterns and predict success at an early stage of the online course.

5. Learning Patterns of Chemistry Teachers in a Professional Development Online Course

5.1 Highlights

- Using a case study methodology, we identified teachers' learning patterns in an online professional development course.
- We learned that some teachers studied continuously from week to week, whereas others 'binge' or practice 'interval learning'.
- We found that online learning patterns are strongly associated with course completion rates.
- Our logistic regression model predicts students' completion rates starting at week 5.

5.2 Introduction

As discussed in the Introduction and chapter 3, the completion rate of online courses is lower than in traditional face-to-face courses. Following chapter 3, we used SRL indicators to produce generalizable models that can identify students who have a high probability of completing the course and those who have a low probability of doing so.

This chapter focuses on teachers who participated in the PD online course: "Introduction to Materials and Nanotechnology." The study population was presented extensively in chapters 2 and 4. This chapter completes and expands the data analysis presented in chapter 4, emphasizing the central theoretical frameworks of the study – students' engagement and SRL.

5.3 Research Questions

In this chapter we address research questions 1 and 4:

Q1) What characterizes learners who are likely to complete online chemistry courses and those that are less likely to do so?

Q4) What is the earliest stage in the online course in which one can predict course completion, and which indicators are required to make these predictions?

5.4 Methodology

Following previous chapters in this dissertation, our research design combined qualitative and quantitative tools (mixed methods). In the qualitative section, we describe and analyze five case studies. The case studies rely on interviews, submission of assignments, analysis of the final assignment, and a personal summary of the learning process that teachers were requested to include with the final assignment. The doctoral student initially prepared each of the case studies, which were then validated separately by two advisers, who read and provided comments. The Ph.D. student then edited the case studies according to their comments. Finally, the case studies were revalidated by the two advisors. In the quantitative section, we used EDM techniques to analyze several parameters identified in the qualitative section, which were also considered appropriate for quantitative analysis.

5.4.1 The Case Study Method

Rather than just using a qualitative analysis of the interviews as we did in chapter 3, in this chapter, we apply case study analysis because, at the Weizmann Institute, we had access to more diverse qualitative materials, as detailed below.

A case study is a strategy used to characterize and analyze a particular situation. It zooms in on the circumstances of a specific situation and illustrates a more general principle. It provides a unique example of real people in a natural context and enables researchers to better understand how ideas and abstract principles fit together (Rap & Blonder, 2017). Case studies are used to observe effects in authentic contexts and to investigate and report dynamic events or human relationships. The use of case studies to portray the experiences that participants underwent can enable the researcher to examine the participants' thoughts more closely (Cunningham., 1997; Cohen et al., 2007).

Semi-structured interviews were the primary research tools that helped us to gather data for the case study analysis. The interviews took place between August 2017 and August 2019. Fifteen interviews were conducted by phone, and four were conducted face-to-face. We conducted nineteen interviews in total, two of which were pilot interviews that were not used in the final analysis. Most of the interviewees were teachers who had completed the online course; however, we also interviewed two teachers who did not submit the final assignment and consequently did not complete the course. The interviews were transcribed and read several times before the analysis phase. We conducted a follow-up interview with one of the teachers about a month after she was first interviewed.

In addition to the interviews, we analyzed personal summaries of the learning process that teachers were asked to include in their final assignment. Teachers were asked to reflect on what they most liked about the course and what they did not like and to provide suggestions for future improvement. These feedback summaries served as an additional tool to expand and enrich the data for the case studies. One of the five teachers mentioned in the case study analysis did not complete the course and therefore did not submit a personal summary.

Finally, we gathered details about the assignment submission rate and the SOLO taxonomy analysis for each case study. The use of SOLO taxonomy in the case studies helped us better understand the relationship between the teachers' learning patterns and the complexity level of understanding they had achieved.

5.4.1.1 Selection criteria

Next, we will describe five case studies that represent teachers' online learning patterns and the time management that arose in the interviews as well as personal summaries of the learning process. The five case studies describe four teachers who completed the course and one who did not. The case studies include teachers' patterns of a learning organization; some were effective, and some were less effective. The aim was to provide examples and to present a variety of learning patterns and time management strategies that emerged from the teachers' learning.

Each case study will include the following:

1. *The teacher's profile:* This is a brief description of the teacher regarding her academic background, her reasons for enrolling in the course, and her approach to online learning. The aim is to provide a broad picture of the teacher described and to shed light on the diverse population of teachers enrolled in the course. The teachers' names in the case studies were replaced with pseudonyms, and any other identifying characteristics were removed.

2. *Learning strategy and time management:* This is a description of the learning process and organization during the course in terms of timing and the use of course materials. We focused on these issues because it is well established in the literature that there are different strategies for online learning. Time management has been found to be an essential skill that affects perseverance and success in the context of online courses (Nawrot & Doucet; 2014; Kizilcec, et al., 2017.
3. *Reference to scientific content:* We highlighted cases where teachers particularly addressed specific scientific content that was learned in the course.
4. *Assignment submission:* We included a description of the submission of assignments during the course.
5. *Implementation of acquired scientific knowledge and pedagogy:* We reported teachers' decisions to implement in their classroom the new knowledge and skills they had learned in the nanotechnology course.

5.4.2 EDM analysis

The EDM analysis included Moodle log files, course completion status, and teachers' demographic profile data. The demographic data included the teachers' place of residence according to their SES and gender.

To follow the principles of research ethics and learners' privacy as well as to follow the GDPR and Israel's protection of privacy law, identifying fields such as name and surname were removed. The research received IRB approval from the Weizmann Institute of Science Ethics Committee.

The resulting database naturally contained numerous fictitious user activities that can bias the activity trends, leading to inaccurate conclusions unless careful strategies for data cleaning, filtering, and indexing are applied. To enhance the data reliability, we performed a pre-processing phase that included four consecutive pre-processing stages: data gathering, data interpretation, database creation, and data organization (see section 2.51), as we described in chapters 2 and 3. We conducted a Mann-Whitney U test to assess the statistical association between ordinal variables and course completion status since most of the variables did not present a normal distribution. Finally, we used the data described here to construct a prediction model with the SCOP variable that we described in section 3.5.2.

5.5 Results

5.5.1 Case Studies

The course “Introduction to Materials and Nanotechnology” was attended by teachers with different years of experience from all over Israel. Some were at the beginning of their careers as teachers, and some transitioned to teaching. The main reasons teachers chose to attend the course were their interest in the topic of the course and the fact it was an online course. Most of the course participants were women (see sections: 2.3.2).

Case Study 1: Rona – “Continuous learner through multiple views of each lecture.”

Teacher profile: Rona is an experienced chemistry teacher with a master’s degree in chemistry. She loved to study and enrolled in the course out of interest and for PD credit. She was not familiar with the field of nanotechnology before the course began, but she thought it would be relevant to her students. Rona lives in northern Israel. If the course was given close to where she lives, she would have preferred to enroll in an in-person method of instruction; however, Rona enrolled in the online PD course since she did not have that option. Nevertheless, she still believes that an online course has many advantages, especially the possibility of going back and rewatching the lectures.

Learning strategy and time management: During the course, Rona spent many hours studying, devoting time to complete the course assignments, and viewing each lecture two or three times. She explained that a combination of her personality, interests, and motivation to earn PD credits helped her persist and successfully complete the online course. She also noted that as the course progressed, it was easier for her to persist. It was difficult for her to estimate how much time she should devote to each task during the course. Consequently, she contacted the course’s teaching staff several times via email to inquire about this issue. She would have preferred to have received more explicit guidelines regarding how much time she should spend studying weekly. Rona described how her way of learning changed during the course:

“*At the beginning of the course, I would first listen to the lectures and only then complete the quizzes. But over time, as the course progressed, I would solve the* *quizzes and watch the lectures simultaneously. In this way, when I reached a specific topic, I could view it in the video lesson, which was more convenient*."

Regarding the scientific content: Rona told us that one phase in the middle of the course was not easy for her, especially the lectures on quantum mechanics. She explained:

"*There were parts in the middle (of the course) that were difficult for me, mostly the quantum mechanics lectures. When I studied this topic in the university 35 years ago, it was at the same time as all the math courses. But now, I did not remember any of it. I have a masters' degree in chemistry, but after so many years, it was difficult. I cannot say I fully understood quantum mechanics, but it was interesting in terms of the implementation.*"

Assignment submission: Rona submitted four (out of six) Padlet assignments during the course, and all her responses were at the 'Relational' SOLO level. Additionally, she submitted 90% of the course quizzes. In Rona's final assignment, she reached the 'Relational' level in two concepts and the 'Multi-structural' level in one concept. As presented in Figure 4.2 in chapter 4, most teachers' explanations were at one of these levels.

Implementation of the acquired scientific knowledge and pedagogy: Rona indicated that the course's topic inspired her to add a school activity about nanotechnology. In addition, she plans to use the "Dilution" experiment presented in Lesson 3 ("Size-dependent properties"). She is considering incorporating additional topics from the PD course, following the advanced training but has not yet decided which ones.

<u>Case Study 2 -Danny – "Single video viewing without repeating them".</u>

Teacher profile: Danny has been teaching chemistry for three years. He said that he is well acquainted with physical chemistry, enrolled in the PD course out of interest, and wanted to know how to teach this subject to his students. Danny lived in the south periphery of Israel, far from the training centers, and chose the online PD because it allows him to control his own schedule. Danny reported that he is highly self-disciplined and enrolled in the course out of interest and not just for PD credits.

Learning strategy and time management: Danny reported watching lectures and then answering the course quizzes. He said he watches every lecture once but knows when he loses his concentration. When this happens, Danny replays the recording. He said he made sure that he watched the video lectures from week to week and made up for the ones he missed during the week-long holiday. He told us he would have preferred to receive a schedule detailing how much time should be spent on each task.

Danny's view of the collaborative Padlet board was very positive. He added several posts and would have been interested in even more collaborations between the teacher-learners. He thought the advantages of the Padlet were most evident in the final assignment, where all the assignments were shared on the same Padlet. He explained: "In this approach, the course participants become content creators, and it is a great idea...this allows for active participation and knowledge sharing between teachers with different backgrounds and interests."

Regarding the scientific content: In the personal summary of the learning process Danny specifically addressed a section on quantum theory, and whereas most of the other interviewees stressed how difficult this topic was for them, Danny noted: "A part that I liked and, in my opinion, is crucial, is the lack of compromise regarding the theoretical knowledge in quantum chemistry. I am particularly interested in this subject, and I also liked it during my undergraduate studies.

Assignment submission: Danny answered all six Padlet tasks. In one of the six, he reached the extended abstract level, which was rare among most of the teachers. In the other five, he only reached the multi-structural level. Danny submitted all of the quizzes, and in the personal summary, he noted that he did so because he thought that the quizzes were mandatory. In the final assignment, he reached the relational level of complexity.

Implementation of acquired scientific knowledge and pedagogy: Regarding using the technological tools demonstrated in the course (for example, video combined with questions and Padlet board), Danny said that there is a problem with the school infrastructure in the school where he teaches because there are not enough workstations. In terms of combining topics studied in the course, he would like to incorporate experiments with nanotechnology and the subject of "*a particle in a box*" into his chemistry lessons.

Case Study 3 - Michelle– “Skipping pattern”.

Teacher Profile: Michelle is a new teacher with a bachelor’s degree in chemistry; she made a career change and began teaching chemistry. She registered for the PD course to gain more knowledge in chemistry teaching and to receive credit for PD. Michelle used the course website and attended the in-person meeting where an experiment was conducted at the Weizmann Institute of Science. Nevertheless, Michelle did not submit the final assignment due to lack of time and therefore did not officially complete the course. She explained that as a new teacher, she was very busy at school because it took her longer to prepare for every lesson she taught. She stated that she was satisfied with the course, but in the future, she would be happy to retake it to receive credit for the advanced training.

Learning strategy and time management: Michelle’s approach was first to view the lecture and then solve the relevant quiz. She watched each video lecture only once. Michelle did not plan her learning time in advance and tried to study whenever she had free time. She did not summarize the lectures but downloaded the presentations to a folder on her computer. Michelle explained that online learning was a challenge for her due to a lack of self-discipline and loss of concentration during classes. However, solving the quizzes directly after watching the video lecture helped her remain engaged. Although she did not complete the course, she thinks the course was excellent because of the knowledge she gained.

Regarding scientific content: At the beginning of the course, Michelle’s strategy for addressing the challenging scientific content was to refer to her undergraduate chemistry course summaries. She described it as follows: “*I was already familiar with most of the course material but did not directly link it to nanotechnology. When something was unclear, I would read lecture summaries from my undergraduate studies.*”

Michelle emphasized that her difficulties with the scientific content continued. When asked whether she had skipped certain parts of the course, she explained that she found the photoelectric effect topic the most difficult. “*In this part, I gave up and said to myself: ‘I have a hard time, and I have no time to delve into the material.*” She later noted that she had already learned about the Schrödinger equation in her undergraduate studies, and even back then, she found the subject too complex.

Assignment Submission: Michelle participated in only two of the first Padlet assignments and reached the relational level of understanding in the first assignment and the multi-level understanding in the second. She explained that she was not comfortable with this kind of collaborative activity. When asked why Michelle explained that she prefers face-to-face interactions with the lecturer and the other participants. She answered 50% of the course quizzes. As mentioned before, Michelle did not submit the course's final assignment.

Implementation of acquired scientific knowledge and pedagogy: Michelle explained that the course provided her with important knowledge. Nevertheless, she needed time to process it to understand how to impart it to her students. Michelle felt most comfortable teaching the topic of size and scale. She explained that the course content was relevant to the high-school chemistry curriculum. However, regarding the technological tools, she does not plan to implement them due to her personal preference for the in-person method of teaching.

A few months after the course ended, we interviewed Michelle again to ask her how the course had contributed to her professional experience and whether she had returned to the course materials on the course website. Michelle said she mainly uses the lessons from the beginning of the course: "size and scale" and "size-dependent properties." She used these materials when she was training middle-school teachers in order to expose them to chemistry. She showed them the videos from the course website. She believes that if middle-school science teachers are familiar with nanotechnology and appreciate its beauty and applicative nature, they will be more successful in encouraging students to take chemistry in high school.

<u>Case Study 4 –Karen- 'Binging' the course with a friend.</u>

Teacher Profile: Karen has a bachelor's and a master's degree in chemistry, and she chose the course out of interest. She was unfamiliar with an elective unit that deals with physical chemistry and thought the course could help her become more familiar with relevant topics. She chose the online course because it was convenient and provided flexibility.

Learning strategy and time management: Karen emphasized her lack of self-discipline. She did not set a specific learning time and did not study consistently. Eventually, however, Karen studied with another teacher; this helped her complete the course and submitted the final assignment. She explained:

“*At first, I tried watching the video lessons, but I couldn’t keep up. I didn’t study for a few weeks, so I ended up doing a ‘marathon’ with a friend. We watched all the videos in one week and also completed the quizzes. Sometimes we downloaded the lecture slides, but mostly, we learned from the video lectures.*”

However, Karen and her friend skipped most of the quantum theory lectures. She explained that since she had completed the course towards the end of the semester, she had already heard from the other teachers that these lectures were more difficult and decided to skip them. She suggested both in the interview and in the personal summary that a complex topic such as quantum theory should be taught in person and not in the online format. Karen felt that the Padlet assignments were less helpful and productive for her. Because she did not watch the lectures weekly, she did not post on the Padlet simultaneously with all the other teachers. Therefore, it was difficult for her to add a new or an original idea that other teachers had not already posted on the Padlet board.

Assignment Submission: Karen participated in all six Padlet assignments, and her explanation was at the multi-structural level. She submitted less than 50% of the course quizzes. The SOLO analysis of her final assignments showed that she had reached the relational level of complexity of understanding for one concept and the multi-level for two other concepts.

Implementation of acquired scientific knowledge and pedagogy: Following the course, Karen was interested in integrating the online quizzes and videos into her classroom.

<u>Case Study 5: Delilah- Intervals Learning.</u>

Teacher Profile: Delilah is a relatively new chemistry teacher who is in her twenties. She attended the course to receive PD credit and because she loved nanotechnology when she studied it in her undergraduate studies. Her goal was to learn how to introduce students to a complex subject such as nanotechnology. In addition, she enrolled in an online course because of convenience because she lives far from any university campus. It is difficult for her to combine attending in-person lessons while working full time and being a mother. In addition, Delilah appreciated the flexibility of being able to re-watch lectures.

Learning strategy and time management: Delilah said that she changed her learning strategy during the course: “*At first, I watched the video lectures and then answered the quizzes. Later on, as the course progressed, I began answering the quizzes at the* *same time that I viewed the lesson*." She also reported what we call "interval learning," which means combining several lessons and watching them together. She explained: "*I did not learn every week. Rather, I accumulated four lessons at a time and watched them in succession.*" Delilah thought that the Padlet assignments were related to the chemistry curriculum and that the NST concepts were especially significant. However, these kinds of assignments required using the Padlet board; she explained that since she did not necessarily study according to the weekly course schedule, she was frustrated. She explained: "*Sometimes I did not post on the Padlet on time… only a few weeks afterward. This was a bit frustrating for me because most of the other teachers had already posted.*" In addition, she mentioned that the in-person lab meeting towards the end of the course served as a trigger to watch the video lectures she had missed in order to be ready for this meeting. In the personal summary she submitted with her final assignment, Delilah noted that because the course was online, she did not have to waste time commuting to campus or finding childcare arrangements.

Regarding the scientific content, similar to other teachers, Delilah also stressed that she had issues with quantum mechanics and difficulties with the mathematical aspects. In addition, Delilah stated that she used past materials from her undergraduate studies. In her personal summary, she explained that the chapters dealing with the mathematical development of quantum mechanics were stimulating and reminded her of her undergraduate days; however, she felt that they were too complicated for an online PD course.

Assignment Submission: Delilah submitted all the course quizzes and three of the six Pedlet assignments, where she demonstrated a relational level of understanding according to the SOLO taxonomy. In her final assignment, she reached the relational level in all three concepts.

Implementation of acquired scientific knowledge and pedagogy: Delilah explained that she would not be able to integrate technological tools from the course into her school teaching: "*Unfortunately, I cannot combine the technological tools…I teach in a school with a complex population. I have tried to integrate technology before, and it wasn't easy. I do, however, incorporate a lot of videos in class.*" In terms of the course content, she found that it possible to implement the scientific materials learned in her school teaching and to use some of the course videos with her students.

Comparison of case studies

To compare the case studies, we analyzed them according to the SRL dimensions used in chapter 3. These dimensions are based on the OSLQ (Barnard et al., 2009). The analysis presented in this chapter is top-down since we examined how each of these pre-defined dimensions is reflected in the case studies (Shkedi, 2003). Table 5.1 presents the case studies according to the OSLQ dimensions together with an example from each case study. Because teachers did not provide details about their physical and online learning environments, and the course video lessons served as the primary learning resource, the dimension of “environment structuring” was omitted from the table. The table also includes an additional dimension that addresses the implementation of acquired scientific knowledge and pedagogy.

Table 5.1 SRL Characteristics that emerged from the case studies

SRL Characteristics	Case	Description
Goal Setting	Rona	Personal interest, relevance to her students, convenience, PD credit.
	Danny	Personal interest, positive familiarity with the subject, learning how to introduce the topic to HS students.
	Michelle	Gain more knowledge in nanotechnology.
	Karen	Exposure to the field of nanotechnology because of the HS elective unit.
	Delilah	PD credit, to learn about how to introduce the topic to HS students.
Task Strategies	Rona	Multiple views of each lecture; an adaptable learning strategy.
	Danny	Single viewing of videos, self-identification of the loss of concentration, no skipping.
	Michelle	Solving the relevant quiz following a single viewing of the video, skipping difficult parts of the course.
	Karen	Studying with a friend. Skipping difficult parts of the course.
	Delilah	Multiple views of each lecture; an adaptable learning strategy.
	Rona	Many hours spent studying, devoting time to the course assignment.
	Danny	Watched the video lectures from week to week and made up for ones that were missed.
	Michelle	Did not set a specific learning time but studied consistently.

Time Management	Karen	Did not set a specific learning time; eventually, she studied with a friend, and completed the course in one week.
	Delilah	Did not learn every week; accumulated four lessons and watched them in succession.
Help - Seeking	Rona	Contacted the course's teaching staff several times via email.
	Danny	There was no reference to this issue.
	Michelle	Read lecture summaries from her undergraduate studies.
	Karen	Studied with a friend.
	Delilah	Read lecture summaries from her undergraduate studies.
Self-evaluation	Rona	Did not fully understand quantum mechanics. She was motivated to complete the course.
	Danny	A high level of self-discipline.
	Michelle	Difficulty in understanding the Schrödinger equation. Lack of self-discipline.
	Karen	Lack of self-discipline.
	Delilah	Had issues with quantum mechanics and difficulties with the mathematical parts.
Implementation of acquired scientific knowledge and pedagogy	Rona	Planned to implement course content in her teaching.
	Danny	Planned to implement course content in his teaching.
	Michelle	Implemented course content in her teacher training.
	Karen	Planned to implement the technological tools used in the course.
	Delilah	Plans to implement the course content in her teaching.

Some patterns described by the teachers can be reflected by their learning patterns on the course website. We can therefore analyze them using data mining techniques. The log file did not enable us to explore all the characteristics that emerged from the case studies.

The results presented below are for the 95 teachers who participated in the course - 61 teachers who completed the course and 34 teachers who were active on the course website but did not complete the course.

5.5.2 EDM Analysis and Examples from the Case Studies

5.5.2.1 Padlet Assignments Submission

Although the course's six Padlet assignments were not mandatory, teachers were encouraged to participate in all of them. Figure 5.1 provides information about the Padlet submission patterns throughout an entire course. A close look at Figure 5.1 shows differences in the pattern of Padlet submission between teachers who successfully completed the course and those who did not complete it. Teachers who completed the course participated with an average of four Padlet assignments, whereas those who did not complete the course participated in only one Padlet assignment, on average. A Mann-Whitney test indicated that this difference was statistically significant ($U=682$, $Z=-2.892$, $p<0.001$). In chapter 3, we used the optional assignment as a predictor of course completion. However, here we cannot use the Padlet assignments as a prediction variable since teachers could have completed them at the end of the course. Teachers were not required to submit the Padlet assignments by a specific date. For example, in case studies 4 and 5, Karen and Delilah explained that they completed the Padlet assignments towards the end of the course.

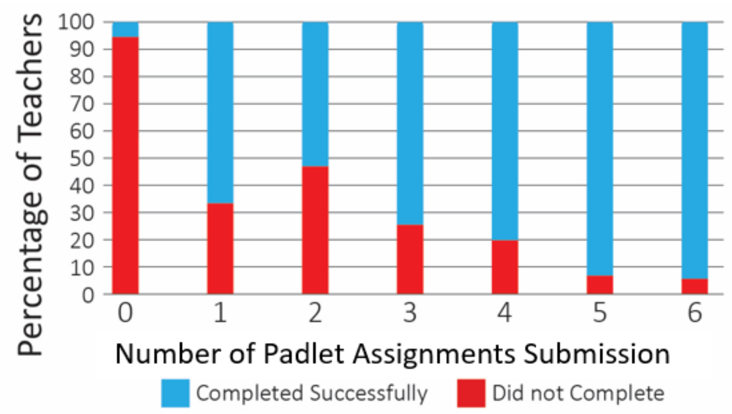


Figure 5.1. The number of submissions of Padlets (optional assignments). Light Blue: Students who successfully completed the course. Red: Students who did not complete the course.

5.5.2.2 Course Video Lessons

Skipping lectures in online learning is a common phenomenon (Warner et al., 2015); it is reflected in the case studies presented above (Michelle and Karen – case studies 3 and 4). In case studies 3 and 4, we identified teachers who reported that they had skipped video lectures about quantum mechanics. We examined how the teachers viewed each video lecture to better understand this pattern of skipping. In the previous chapter, we counted the playing of only the first video in each lesson and concluded that some teachers skipped lessons due to content difficulties (see Figure 4.4). In this chapter, we focused on all video lectures that comprise the same lesson. Recall that the course included 13 pre-recorded video lessons, some of which comprised 1-5, 25-minute-long sub-videos. This division enabled us to analyze how teachers viewed these sub-videos (the index of each video lecture, including the topic and the specific lesson number, are detailed in Appendix 5). Figure 5.1 presents the mean percentages of unique video playing for each lesson for teachers who completed the course (Figure 5.2 A) and those who did not (Figure 5.2 B).

These three figures (4.4. 5.2a, 5.2b) also allow us to examine the percentage of teachers who skipped each video. These figures show a decrease in the number of teachers who played the remaining videos of the lesson after each initial video lesson. The one exception to this trend is video 7E, which dealt with a topic included in the high-school chemistry curriculum and is, therefore, more likely to interest the teachers. It can be seen that the number declined between the first and last lessons, but that this decrease is much sharper for those teachers who did not complete the course, as presented in Figure 5.2 a +b.

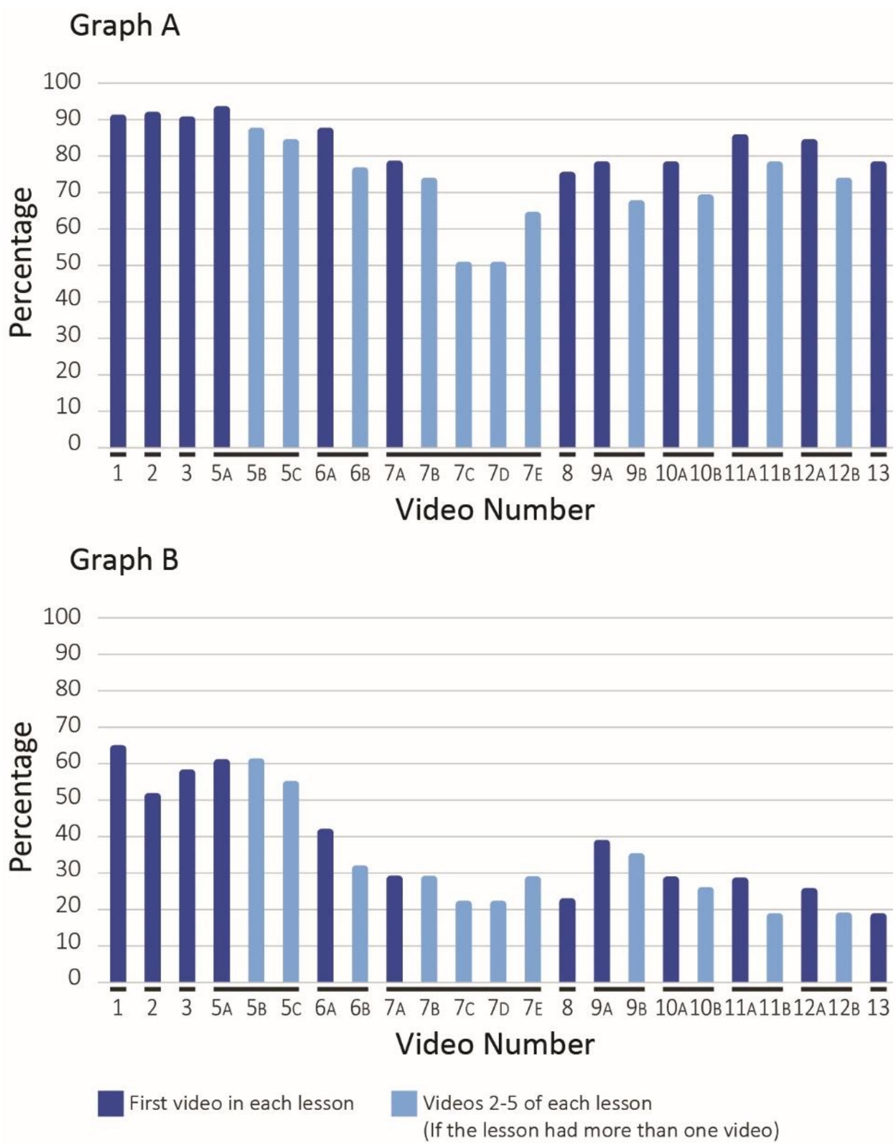


Figure 5.2: Average percentages of unique video playing for each video.
A. Teachers that completed the course successfully. B. Teachers that did not complete the course.

Figure 5.3 presents a different view of these data and counts the percentages of teachers from each group that opened the sessions' first, second, third, and fourth quartiles. As is evident, most of the students in the group that successfully completed the courses played the major set of video lessons. Most of the students who did not complete the course opened only some of the sessions. Figure 5.3 shows that some students skipped lectures.

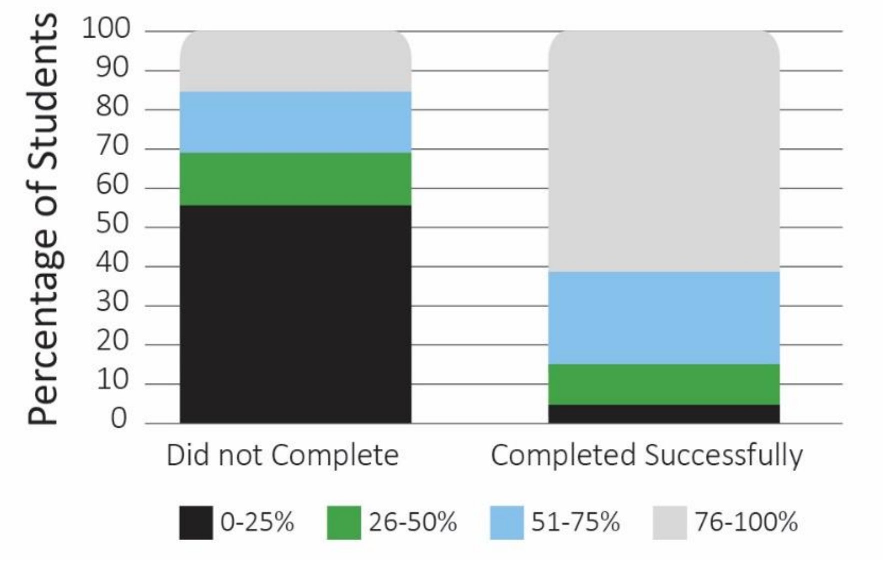


Figure 5.3 The opening rate of the course video sessions: Colors represent video opening percentages. Black: 0-25%. Green: 26-50%. Light Blue: 51-75%. Gray: 76-100%.

Following the above analysis, we calculated the number of videos each teacher skipped (Skipping Index - see section 2.5.2). Teachers who completed the course skipped four videos, on average, whereas those who did not complete the course skipped 12. A Mann-Whitney test indicated that this difference was statistically significant ($U=267$, $Z=-4.862$, $p<0.001$). As we highlighted in case studies 3 and 4, some teachers skipped some of the quantum mechanics videos. To determine whether this affected our results, we divided the videos into those videos relating to quantum mechanics and those that did not. For each video group, we calculated the differences between teachers who completed the course and those who did not. Using a Mann-Whitney test, we found a statistically significant difference between teachers who successfully completed the course and those who did not in both groups: (videos relating to quantum mechanics, $U=492$, $Z=-4.048$, $p<0.001$) (videos that did not relate to quantum mechanics $U=302$, $Z=-5.551$, $p<0.001$). From this, we can conclude that teachers who did not complete the course not only skipped the more difficult quantum mechanics videos but other videos as well.

In our study, teachers described their learning patterns using the word "week," detailing different learning timeframes. Some described studying from week to week (case study 1- Ronna), whereas others indicated that they had studied every few weeks (case study 5 - Delilah), or sometimes most of the course in one week (case study 4 - Keren).

Therefore, in the file analysis, the primary time unit is one week, which we measured starting from when a new lesson was opened on the course website.

In one of the case studies, we encountered a teacher who waited until the end of the course to watch most of the video lessons in one week (case study 4). We defined this learning pattern as 'binging the course.' Binging online content has emerged as a trending behavioral phenomenon among users of online streaming services such as Netflix or Amazon (Yoo et al., 2020). Only recently have studies begun to examine binge-watching in the context of online educational settings (Yoo et al., 2017). Nine of the teachers in our sample exhibited this pattern, seven of whom successfully completed the course. Five of the teachers who binged the course skipped between one and five video lectures. In chapter 3, we did not calculate the binge parameter. This is because the binge pattern did not emerge from the qualitative analysis. Instead, this parameter was added in Appendix 4.

In another case study (case study 5 - Delilah), we encountered a teacher who did not follow the video lessons every week but, instead, accumulated four videos and then watched them all in succession. Following Dermy et al. (2020), we defined this learning pattern as 'Interval Learning.' This pattern is far more difficult to quantify through the log file data because individual students could learn in different intervals regarding the length and the break between consecutive intervals. For example, a teacher can learn for three weeks and then "rest" for four weeks, whereas another student can learn for two weeks and then "rest" for three weeks. Because of this difficulty in assessing interval learning, we evaluated the number of weeks that each teacher actively played a new video lesson during the course. Our analysis focused on the video lectures; therefore, we defined a teacher as active in a specific week if the teacher played at least one new video during that time (number of active weeks – see section 2.5.2). We found that teachers who completed the course were, on average, active for seven out of the 20-week course period (this includes the weeks in which students worked on the final assignment). Teachers who did not complete the course were, on average, active for only five weeks. A Mann-Whitney test indicates that this difference was statistically significant ($U=406$, $Z=-3.536$, $p<0.001$).

The log file data were used to construct a prediction model with the SCOP variable that we described in chapter 3. Recall that this variable represents students' cumulative opening patterns (SCOP) of the video lessons. The SCOP counts learners' weekly advancement in the course lectures (but does not count multiple playing of the same video). We applied this variable to predict which teachers are likely to complete the course and which are not. Figure 5.4 presents the weekly average SCOP for each group (successfully completed and did not complete). As is evident, this parameter is quite informative for distinguishing between the two groups, even at the early stages of the course. Note that the course itself lasted 13 weeks; therefore, data for weeks 14-20 represent the period when students worked towards the final assignment. We included these weeks to show that teachers continued advancing in the video sessions while working on the last assignment. As can be seen, the group of successful students used the video resources much more than the students who did not complete the course. The trends presented in Figure 5.4 are similar to the ones we found in the OUI, presented in chapter 3 (see Figure 3.5). The SCOP variable we define does not evaluate whether the teacher played the online sessions from week to week in a linear order (video 1, video 2, video 3, and so forth). Therefore, we calculated the index for linear learning presented in chapter 2. The average linearity for those teachers who successfully completed the course was 0.54, and for teachers who did not complete the course, it was 1.15. A Mann-Whitney test indicated that this difference was statistically significant ($U=560$, $Z=-2.416$, $p<0.001$). This pattern is not shown in Figure 3.4.

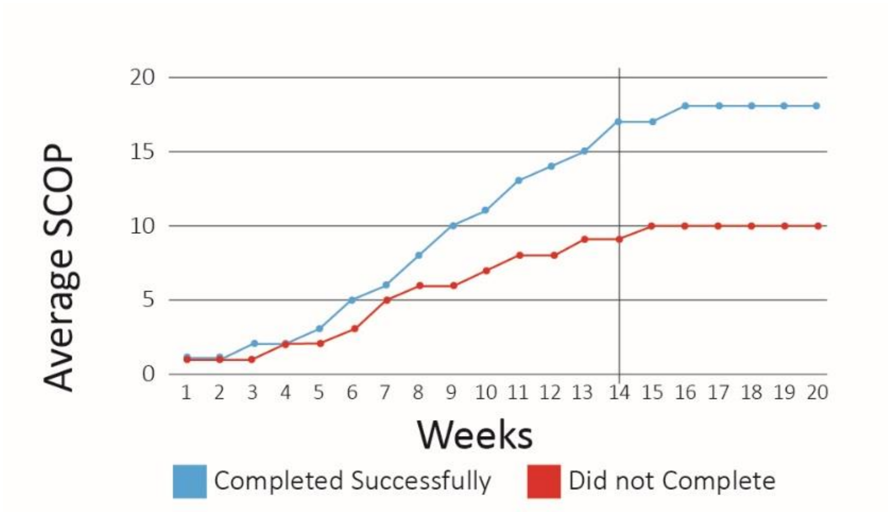


Figure 5.4. Teachers Cumulative Opening Pattern (SCOP). SCOP– weeks 1- 20. Light blue: Teachers who successfully completed the course. Red: Students who did not complete the course. Lines were used to lead the eye.

5.5.2.3 Building a Logistic Regression Model

To predict when a teacher is most likely to complete the course as early as possible, we built a logistic regression model based on the SCOP variable. Since the previous parameters, we calculated in this chapter were calculated at the end of the course, we did not include them in the model. The course year, gender, and SES were used as control variables (no multicollinearity was found between these variables). Because most variables did not present a normal distribution, the Mann-Whitney U test was used to assess the statistical association between SCOP and course success. It was found that starting at the fifth week, teachers who completed the course received a statistically significantly higher score in the Mann-Whitney U test than those who did not complete it ($p < 0.05$). We used the Hosmer-Lemeshow test goodness of fit test for logistic regression; it indicates how well the data fit the model (Paul et al., 2013). This test found that the data were fit starting at week 5.

After teachers with missing variables were removed, the sample we used for the logistic regression model included 88 teachers. The model's results, based on the SCOP variable as a predictor, are presented in Table 5.2. This variable was found to be statistically significant, $\chi^2(4) = 18.261$, $p < .001$, suggesting that one can identify the probability of succeeding in the courses based on the following, statistically significant parameters: The SCOP at the 5th week ($p < 0.01$), gender – male ($p = 0.05$). Although being male was a significant predictor of course success, it is important to stress that men constituted only 17% of the population. Therefore, we chose not to attribute too much importance to this finding. The model explains 23.9% (Nagelkerke R^2) of the variance in succeeding in the course, and it correctly classifies 73.9% of the cases. The results indicate that early prediction models based on teachers' data collected before the course's mid-point enable one to identify students who probably will succeed (and those that probably will not).

Table 5.2 Models of logistic regressions of succeeding in the course. (N = 88).

	SCOP	
Variable	**Wald**	**Sig.**
District of residence (SES)	0.000	0.650
Year	2.452	0.117
Gender	5.684	**0.017
SCOP at week 5	7.626	**0.006

**p<0.01

Model Evaluation

Table 5.3 compares the predicted classifications of teachers' final status (completed/did not complete) according to the model in comparison with their actual classifications. In this model, we defined a teacher with a probability of 0.5 or higher as an individual who probably will successfully complete the course and a teacher below 0.5 as an individual who probably will not complete the course. The model predicts the course's completion better than it predicts its incompletion.

We further evaluated the models by plotting the area under the curve (AUC) to estimate their accuracy based on the Receiver Operating Characteristic (ROC) curve (see Appendix 2 for details). The AUC value was 0.734. These values approach 0.7, considered acceptable in scientific research (Mandrekar, 2010). In addition, we used a 10-fold cross validation to evaluate the average accuracy of the model. The average value of the AUC was 0.73.

Table 5.3 Actual and predicted classifications of course completion. N = 88.

	Predictions		
Actual Status	**Probably will not Complete**	**Probably will Successfully Complete**	**Correct Predictions (%)**
Did not Complete	9	19	32.1
Successfully Completed	4	56	93.3
Overall Percentage			73.9

5.6. Research Limitations

The chapter's main limitation was that the quantitative analysis used a relatively small sample size. Because of this shortcoming, we decided to primarily use correlation analysis rather than more advanced statistical tests requiring a larger sample size. However, we took the liberty to use a more advanced statistical model for the SCOP variable because this variable, used in chapter 3, and its development, rely on a much larger sample. We hope that our results will serve as a basis for further research that applies the more advanced statistical analysis of the patterns that emerged from our case studies.

Another limitation concerns our regression model. As we previously showed, the model used in this analysis predicted at week five, with 95% accuracy, which of the teachers enrolled would eventually complete the course. However, the model's weakness is that it does not consider the six percent of teachers who 'binged the course.' At week 5, these teachers' learning patterns mirror those who did not complete the course. This, however, changed towards the end of the course when a few teachers binged the lectures they had missed. Unfortunately, our model cannot reveal this pattern. Finally, the parameters we defined to analyze teachers' learning patterns using EDM do not fully reveal the learning patterns that emerged from the case studies. Despite this limitation, we could still use EDM techniques to distinguish and characterize differences between teachers who successfully completed the course and those who did not.

5.7 Discussion

Like in chapter 4, this chapter also studies the online “Introduction to Materials and Nanotechnology” teachers’ PD course and examines learners’ difficulties. However, in this chapter, we focus on learning patterns. Our goals were the same as in chapter 3: characterize learning patterns and predict success in an online course based on their engagement and the SRL theory. However, whereas chapter 3 focused on undergraduate students, in this chapter, we study teachers in the context of an online PD course. We used a mixed-methods approach that combines case study analysis and EDM techniques. In contrast to chapter 3, where the qualitative research was based only on interviews, we relied on interviews, feedback summaries, and course assignments and case studies in this chapter.

The case studies were used to identify and better understand teachers’ learning patterns in the online PD. As professionals accustomed to reflection (Mamlok-Naaman & Eilks, 2012; Laudonia et al., 2018), teachers knew how to describe in detail what they had learned and how they did so. We characterized teachers’ course participation by examining five representative case studies. These case studies reveal various learning patterns, such as single access to each video or parallel access to the videos and the quizzes for technical convenience to save time as well as different time management patterns.

The five case studies represented patterns that emerged from the interviews and teachers’ feedback summaries of the learning process they included in their final assignment. Some of the patterns could also be shown using EDM techniques. This illustrates how specific learning patterns can be found by analyzing the Moodle log files. We also determined whether these patterns are effective and characterized learners who had completed the course. The different patterns that emerged from the case studies can reflect teachers’ SRL and their engagement in the online environment—for example, time planning influences the completion of an online course (Handoko et al., 2019). The third case study (Michelle) exemplifies a teacher who did not have a clear learning schedule and did not complete the course. This supports existing studies that stress how time management is essential in online course success (Inan et al., 2017). Time management is usually assessed by a self-reported SRL questionnaire (Barnard et al., 2009; Pintrich, 1992; Magno, 2011). For example, in the OSLQ questionnaire, students report a weekly or daily schedule

(Barnard et al., 2009). However, an online course that enables flexible learning time allows learners to manage their time according to their plan during the entire course period and not only with daily or hourly planning. In this chapter, this principle was demonstrated in the case study representing ‘interval learning.’ This time management method is not generally appropriate for every online course; it depends on the course content and design. In the course examined in this chapter, the Padlet assignments were not planned for ‘binge’ or ‘interval’ learning. Teachers who did not follow the course from week to week reported that the Padlet assignments were frustrating for them because they felt that the timing of the submission was essential in order to contribute to their classmates’ learning. Indeed, teachers who submitted the Padlet assignments towards the end of the course reported that they did not benefit from them as they could have.

Another aspect discussed in the previous chapter (chapter 4) and that was expanded here refers to teachers’ difficulties in the course and how they addressed them. On the one hand, when faced with a complex topic, several teachers chose to avoid it. For example, after hearing from others about several difficult lessons, Karen decided to avoid these challenging topics and skipped these lessons when she was “binging the course with a friend.” On the other hand, we encountered examples of teachers’ help-seeking patterns as ways to deal with the challenges. Two teachers who were relatively young and were mentioned in the case studies reported that when they had encountered difficulties in quantum mechanics, they turned to their bachelor’s degree course summaries, which were available to them (cases studies 3 and 5, Michelle and Delilah). In the case study of “investing a lot of time,” Ronna explained that because of the difficulties she had with quantum mechanics, she would be unable to implement this topic in her teaching despite her in-depth learning.

In the previous chapter, we found that teachers had specific difficulties with the lessons on quantum mechanics. In this chapter, we examined how teachers dealt with these difficulties. We learned that teachers who decided to skip the quantum mechanics lessons could still complete the course, based on the case studies and the EDM analysis. This probably results from the course’s design. Teachers could choose which NST concepts they want to elaborate on in the final assignment, allowing them to sidestep the topic of quantum mechanics altogether.

The third case study teacher, Michelle, did not complete the course. However, she reported that she implemented the course materials in her teaching. This shows that even teachers who did not complete the course decided to implement the knowledge they had acquired. This provides further support for the claim raised by Rabin and colleagues (2019). They suggested that in the context of a PD course, success should not be evaluated only according to the metrics of persistence and course completion but also by assessing learners' fulfillment of their expectations. Teachers who enrolled in the "Introduction to Materials and Nanotechnology" course were already in-service, with several years of teaching experience. Analyzing their learning patterns through SRL theory revealed how previous knowledge assisted them in acquiring new scientific knowledge.

5.8 Summary and Implications

SRL and learner engagement are essential factors in every type of learning. However, their importance increases in the context of online education, considering the flexibility that the online domain provides in choosing the place and time for learning (Li et al., 2020). The current chapter analyzed how teachers study in an online PD course. Five case studies of teachers were presented: four who completed the course and one that did not. The teachers who completed the course and agreed to be interviewed reported that they were comfortable with the online learning platform. We found that different learning patterns could result in teachers successfully completing the course. In addition to the learning patterns that emerged in chapter 3, in this chapter, we identified two additional learning patterns: interval learning and 'binging.' In addition, we built a logistic regression model based on the model from chapter 3 and showed that the main predictor in the model, the SCOP variable, reflects teachers' engagement in the course because it assesses their weekly advancement in the course lectures.

6. Discussion and Conclusions

This dissertation’s main goals are to identify students’ and teachers’ learning patterns in online chemistry courses as well as to develop assessment tools to predict learners’ success. Previous research showed that online learning is characterized by low completion rates relative to face-to-face courses (Narayanasamy & Elçi, 2020; Lakhal & Khechine, 2021). The two main theoretical frameworks often used to explain these phenomena are SRL and student engagement (Soffer & Cohen, 2019; Artino & Jones, 2012; You, 2016). This study builds on these existing theories and applies them to examine undergraduate students enrolled in online chemistry courses as well as chemistry teachers participating in an online PD course. We characterized learners based on the SRL approach and expanded the current understanding of online learning by identifying patterns that lead to successful distance learning. We also identified challenges that often result in the incompletion of online courses. Two predictive models that determine, at an early stage, students’ likeliness to complete the course serve as a central tool developed in the context of this study. Finally, we developed a framework for evaluating online courses for teachers’ PD. This framework integrates the evaluation of online activities with traditional evaluation tools.

6.1 Pre-Processing Stage

At the outset of this study, we encountered the challenge of analyzing data extracted from LMS. The immense amount of data archived by LMSs, pertaining to users’ activities has enabled researchers to accurately analyze students’ learning patterns in online learning environments (Aldowah et al., 2019). However, such data typically contain numerous fictitious user activities that can bias the activity trends. Unless careful data cleaning, filtering, and indexing strategies are applied, this could lead to inaccurate conclusions. As the number of publications in the field continues to grow (El Aouifi et al., 2021; Yoon et al., 2021), it is essential to point out the challenges in collecting this type of data reliably (Hershkovitz & Alexandron, 2020). This study offers a unique and detailed perspective on possible challenges in conducting research based on EDM techniques. We wish to emphasize the need to “separate the wheat from the chaff” by implementing a well-documented phase of early pre-processing and interpretation of data before attempting to evaluate online learning patterns based on log files taken from LMSs. To address this challenge, we defined different stages of pre-processing online educational data, which we considered to be critical for reliable data mining. We divided the pre-processing phase into four main stages: data gathering, data interpretation, database creation, and data organization. Our analysis of undergraduate chemistry courses and the chemistry PD course presented in this study assisted us in validating and exemplifying each stage. To avoid bias, we wish to emphasize that these pre-processing stages should be performed by researchers working in large institutions, where they are not necessarily the instructors of the courses under research and where they have little or no control over the format, quality, and extensiveness of the reports produced by the institutional LMS.

We attempted to generalize the technical and cooperative nature of this type of process, along with its specific terminology in the form of four consecutive work stages (Figure 2.1, chapter 2). Interestingly, we found that the challenge begins at the data-gathering stage. In this dissertation, we describe research that began with raw data. When researchers receive processed data, they should be aware of the pre-processing phases that preceded the data to evaluate its reliability. The data interpretation stage emphasized the need to carefully examine the data attributes in the log files to prevent misinterpretation. In the example of the pre-processing phase, we focused on specific features (i.e., user type, time step, IP address, file opening, and activity count). For other studies, the list of attributes can be expanded in line with the specific research goals and the data at hand. In the third phase of database creation, we emphasized the need to follow the GDPR and protect the participants' privacy. The last pre-processing stage was data organization, where various sources are filtered and integrated. These stages integrate the technical, cooperative, and interpretation aspects of this type of research.

This study also aimed to minimize the inclusion of irrelevant and erroneous data in the analysis and to increase researchers' awareness of hidden pitfalls of misinterpretations in the process. However, since different online learning environments provide additional data types, researchers can adjust the workflow suggested in this study according to their data. Overall, our findings led to three main recommendations regarding cooperation, automation, and interpretation.

1. Cooperation: Researchers in academic institutions often have limited control over the collected data, type, and format. Software updates and the institution's policy regarding these updates should also be considered in designing the research. Successfully engaging in data mining requires the cooperation of various staff

members from different academic institutions, who are usually not under the researchers' direct control. To obtain the necessary data in the study's timeframe, researchers should identify this staff early and strive to establish long-term working relationships with them. Researchers should also ensure that the team understands their role and their assistance to the study's eventual success (Knapp et al., 2015; Siemens, 2013).

2. Automatic processes: Automating the implementation of the technical aspects of the pre-processing data stages can increase both the quality and the amount of future data mining-based studies. This automation can help formulate an institutional policy for pedagogical design by building and adopting reliable, user-friendly reports. Such reports can significantly reduce the amount of effort and time that researchers must devote to the pre-processing stage as well as bridge the gap between the educational merit of this type of research and the technical expertise required to perform it (Luna et al., 2017). The application of EDM and LA in higher education may also help provide data and tools that institutions can use for real-time prediction (Aldowah et al., 2019).

3. Interpretation: This is a key issue in understanding the data at hand. As we have shown, several variables can be misleading. In chapter 2, we emphasized the challenge of relying on timestamps. Therefore, we did not evaluate the total time learners were engaged in an activity. Instead, we used "week" as our time unit for evaluation. Another reason for using the "week" time unit is that it provided a more reliable indicator of students' viewing patterns. Because we only knew when students played the video (but not whether they actually watched it), looking at their weekly usage patterns provided a more complete picture. This more careful analysis helped us use the patterns we identified in the data as predictor variables (the SCOP).

To sum up, the suggested stages for data pre-processing should be treated as a preliminary yet necessary phase in any study aiming to analyze educational datasets from an LMS environment. Although LA/EDM-focused researchers are aware of the need for data cleanup, institutional collaboration, and more accurate data interpretation, most published papers do not report all the work carried out in the pre-processing phase, which often remains "behind the scenes." More elaboration on these stages can help newcomers become familiar with these methods. Moreover, researchers interested in reproducing results using the same dataset can do so in the future. We aimed to highlight challenges and dilemmas that researchers most likely encounter during data preparation, particularly in meta-analysis studies. The cooperation of academic institutions' policymakers is required to provide researchers with a reliable and straightforward research environment that could significantly increase the quality and reliability of studies in this field for students, instructors, and institutions.

6.2 Combining EDM and Qualitative Research Methods

A wide array of general-purpose tools and frameworks for conducting EDM research have been developed in recent years (Slater et al., 2017). According to Romero & Ventura (2020), these tools are not easy for educators to use because they require selecting and applying specific methods/algorithms and providing appropriate parameters in advance to obtain good results/models. As a result, employing these methods requires education researchers to become familiar with data science methodologies and tools (Romero & Ventura, 2020).

To address this challenge, we used a mixed-methods approach. This notion is based on previous studies that showed how qualitative information gathered during research could assist researchers who usually apply a straightforward quantitative analysis (Alexandron et al., 2019; Hilliger et al., 2020). The current research combined qualitative analysis (interviews, case studies, and content analysis), which helped us choose the emerging parameters for the statistical analyses and the logistic regression models. The benefits of the mixed-method approach are well known. However, we claim that in the context of EDM, the contribution of this approach is significant. EDM is often used to analyze large-scale data, not all of which is necessary for addressing the relevant research question/s. Qualitative analysis can guide researchers in answering the research question/s by assisting them in filtering the data.

Another advantage of the mixed-method approach is that it can overcome the weaknesses of each method in the assessment of SRL. Traditionally, SRL is evaluated using a self-reported questionnaire. However, the main disadvantage of this approach is that many individuals suffer from self-report bias, and students' memories are often insufficient for them to accurately recall past behavior (Baker et al., 2020). Such closed-ended questions are limited in their ability to reveal new learning methods. This limitation is especially relevant in the context of online learning, which has opened up numerous new opportunities for non-traditional studying. Interviews helped us address the weaknesses of the SRL questionnaires, which are usually composed of Likert scale questions. Although interviews can also suffer from self-report bias, the interviewer can address this weakness, for example, by asking follow-up questions that challenge the interviewee. Interviews are helpful in this regard since they represent a far more open-ended evaluation method.

Another method developed in recent years and that is relevant for addressing the self-report bias of the questionnaire is the use of log file data (Aleven et al., 2016; You, 2016). Inferring SRL using log file data depends on students' interactions within each learning environment (Baker et al., 2020). According to Baker et al. (2020), using log file data to measure SRL has both advantages and disadvantages. The main advantage of this method is that, rather than assessing students' SRL at one or two data points, the log file traces students' SRL throughout the course. The main disadvantage is that it is impossible to receive a complete description of students' SRL based solely on log file data since the data only capture students' interactions on the course website. Next, we describe how we address each research question and the method that we implement to do so. In addition, we discuss the contribution and implications of the research findings.

6.3 Learners' Characteristics

Combining the mixed methods approach helped us address the first research question (Q1): *what characterizes learners that are most likely to complete online chemistry courses and those that are less likely to do so?* Our qualitative analysis identified several online learning patterns practiced by students and teacher-learners (see Figure 3.1 and Table 5.1). We then distinguished between these learning patterns by the six SRL dimensions outlined by Barnard et al. (2009): goal setting, environment structuring, task strategies, time management, help-seeking, and self-evaluation. Since the qualitative analysis was based on a small sample, we wanted to generalize our findings using the large-scale log file data. The details we had in the available log file data were limited; therefore, we were unable to evaluate goal setting, environment structuring, help-seeking, and self-evaluation. However, we could indirectly learn about the dimensions of time management and task strategies from assessing students' choices when they submit optional assignments and from their video playing patterns.

The qualitative analysis mostly helped us characterize the learning patterns of those students who successfully completed the course because most of the interviewees were the ones who had completed the online course. Various aspects related to how learners study and manage their time emerged from our interviews. This led us to realize that the current SRL categories are not detailed enough to evaluate online learning patterns. To address this shortcoming, we recommend expanding the SRL categories to better capture unique learning patterns exhibited in the context of online learning. These insights can be used to develop and update existing SRL questionnaires in the context of online learning. For instance, although the OSLQ questionnaire (Barnard et al., 2009) includes an item about the discussion in a forum under the category of “Task Strategies”, there are no items about video usage in the online courses. Because video lessons are currently the primary resource in online classes, the existing SRL questionnaires should be updated in a manner that addresses this shortcoming.

Although the number of interviews with participants who did not complete the course is small and not enough to draw strong conclusions, some learning patterns associated with this group did emerge. For example, in the two institutions studied, participants who did not complete the course reported not setting aside a specific time in their schedule for learning. Moreover, the quantitative analysis enabled us to characterize the online engagement patterns of both groups: those who completed the course and those who did not. In both institutions, we characterized learners according to a few engagement characteristics: the status of the submission of optional assignments, Skipping Index, SCOP, and number of active weeks. Statistically significant differences were found between participants who completed the course and those who did not regarding these parameters in both institutions. Since the SCOP did not enable us to assess the order of video plays, we also characterized learners by calculating the linearity index, which was statistically significant only in the online PD course. The case studies in chapter 5 reveal the binge pattern; therefore, we characterized the teachers according to this parameter, which was found only for nine teachers, most of whom had completed the course. Binging online content usually appears in research focused on TV viewing patterns (Deloitte, 2016; Jurgensen., 2013). Only recently have studies begun examining binge-watching in the context of online educational settings (Lu et al., 2017). Because this phenomenon remains to be defined, we created our own definition (see chapter 2). However, additional research is needed to better understand and define this new learning pattern.

6.4 Evaluation of Learning Outcomes and Difficulties

Recall that due to data limitations, we could only address our second research question (Q2): *How can we evaluate learning outcomes in the context of online learning?*) and our third research question (Q3): *How can we identify learners' difficulties in the online course?.* These questions focus on learning outcomes and difficulties for the "Introduction to materials and nanotechnology" online PD course. In chapter 4, we answered these two questions by evaluating knowledge, the level of understanding, and difficulties in the context of online learning. To this end, we developed a multi-dimensional evaluation framework. This framework combines EDM techniques with traditional evaluation tools such as the SOLO taxonomy and a pre-post knowledge questionnaire. Each of the dimensions in the framework provides unique insights into the learning process throughout the online course. Using the SOLO taxonomy enabled us to evaluate open-ended responses that are considered a challenge in online course evaluation (Admiraal et al., 2015). Since using the SOLO taxonomy requires manual evaluation, this tool is only suitable for a course with a limited number of participants and for assignments with open-ended questions that lend themselves to analysis with the SOLO taxonomy, such as the PD course we examined. Researchers who focus on the challenges of open-ended question evaluation in an online course should explore how to apply the SOLO taxonomy to a course with a larger group of students. Applying this approach in a course larger than the online PD course we studied could benefit from an automatic system such as natural language processing (NLP) for automated qualitative content analysis, which has been developed recently (Ariely et al., 2020; Çınar et al., 2020).

Another dimension of the evaluation framework focuses on learners' difficulties with the course content. Our study showed that analyzing log file data can help identify changes in learning patterns, which may reflect learners' challenges in dealing with the course content. As we showed in chapter 4, we could identify changes in learning patterns in our analysis of the log files. For example, the rate of playing the quantum mechanics video lessons was significantly lower than other lessons in the course. This result indicated that many teachers decided to skip these lessons. From the interviews, we knew that this was because teachers struggled with this specific topic. In other words, analysis of video playing data can indicate changes in learning patterns, which we can then investigate by using more qualitative tools. This insight can be applied in future online courses by developing visual dashboards that reflect participants' learning patterns for the course staff. Such dashboards can assist the staff in recognizing changes in the usage pattern of a specific course resource and initiate an online discussion to better understand what is causing it. Such a discussion can help the course staff evaluate whether these changing learning patterns result from difficulties with the course material. We will discuss additional dashboard implementations in the next section.

6.5 Predicting Learners' Online Course Completion

The main theoretical contributions of the study focused on the learner's engagement and SRL. This research highlights the importance of two key variables: submitting the first optional assignment and the video opening pattern according to the SCOP variable. Next, we will discuss our analysis findings and elaborate on their relevance to the existing theory.

From our interviews with teachers who took the PD course at the Weizman Institute and those enrolled in online chemistry courses at the OUI, we discovered how essential the assignments were for the learning process. We already know from several studies that students' interaction with course assignments and learning tasks is vital to their learning experiences (Kokoç et al., 2021; Zacharis, 2015). It also aligns with studies that specifically examined general chemistry courses (Cosio & Williamson, 2018; Richards-Babb et a., 2018). Previous models showed that the assignment deadline indicates the course's success (Kokoç et al., 2021). Others have termed delaying online assignment submission as procrastination behavior, which resulted in lower grades (Cerezo et al., 2016; You, 2016; Cormack et al., 2020). According to You (2016), late submissions directly reflect students' time management skills. Alexandron et al. (2020) referred to the submission of non-mandatory assessment items as a measure of engagement. We contribute to this discussion by emphasizing the importance of the optional assignments to predict success in the course. We suggest that optional assignments are related to SRL because the theory refers to students' choices in the context of learning (Roll & Winne, 2015). Recall that in the OUI's courses, submitting the first optional assignment was a significant predictor of course success at week 5. We could use it as an early predictor because students were given a deadline for submitting the assignment. In the PD course, we found a statistical association between the optional assignment submission status and course success. However, we could not use it as an early predictor of course success because teachers did not receive a deadline and could submit the optional Padlet assignments at any time, including at the end of the course.

In addition to optional assignments, we also focused on students' video opening patterns. We decided to focus on video opening for several reasons. First, learners from both institutions described videos as a central learning resource. Second, previous research has shown that students' success in video-based education is mainly dependent on their learning strategies for absorbing and internalizing content delivered by videos (Kennedy et al., 2008; Seidel & Shavelson, 2007).

Previous predicting models applied the number of clicks performed (Giannakos et al., 2015), the video sequence, and the number of weekly videos played by students (El Aouifi et al., 2021) as predictors of course success. The current study emphasizes the benefit of using a cumulative pattern of students' video playing from week 1 to the end of the course as a strong predictor of course completion. The SCOP variable, developed in the study, indicates learners' engagement in the course and indirectly opens a window to evaluate their time management, which is an essential feature of SRL theory (You, 2016).

To address research question number four, (Q4) *"What is the earliest stage in the online course in which one can predict course completion, and which indicators are required to make these predictions?"* we developed two logistic regression models. The first included the submission of the first optional assignment, and the second incorporated the SCOP variable. Next, we outline the different ways our models contribute to the research community and their potential implications.
According to Dalipi et al. (2018), most of the research on dropout prediction models is based on MOOC data. However, because MOOC courses can significantly differ from other online academic or PD courses, there is less knowledge about learning patterns in non-MOOC online learning. Our research addresses this need by examining course completion in these contexts.

Instead of using the commonly used phrase of "dropout" (Shea & Bidjerano, 2019; Wang et al., 2019), we follow Soffer & Cohen (2019), who used the more precise terms "complete/not complete successfully." We emphasize this because a typical academic course also includes books and other learning materials that students use to study. In the context of an online course, not viewing the lectures would make it seem as if students dropped out of the course, when in fact, they just did not use the LMS. Identifying actual dropouts is more challenging when students learn by "binging" the course or by interval learning. Therefore, the terms complete/not complete successfully are more accurate.

In our study of courses at the OUI, we found that the first optional assignment, which we view as a proxy of students' choice, is a predictor of course completion. Our predictive models can be beneficial for lecturers, helping them identify specific learners that are likely either to complete or not complete the course. These models can also help lecturers design interventions that assist learners who face difficulties in the course.

Our predictive models can serve as a basis for creating LMS dashboards. A significant number of studies (Fang & Zahiruddin, 2020; Kew & Tasir, 2021; Matcha et al., 2019; Michaeli et al., 2020), focus on the development of dashboard applications that can visualize data mining results and help intervene in the learning process whenever necessary. Some analytics dashboards are very system-specific, whereas others are developed to be used across different learning platforms (Matcha et al., 2019). We believe that our models can serve as a basis for such dashboard applications. They can also serve as a basis for designing learning activities in a personalized learning approach (Aviran et al., 2020; Fang & Zahiruddin, 2020). Although predictive models are primarily intended for educators, students can also benefit from them.

We also wish to emphasize the importance of students' responsibility for their time management and learning choices during the course, which might positively impact their potential for success (Inan et al., 2017). Traditional courses have fixed time schedules that involve students attending class regularly. However, online courses often do not require students to follow a specific schedule for accessing course material (You, 2016). Therefore, students enrolled in online courses need to make more of an effort to follow the course. These dashboards can also serve students by reflecting their learning choices and advancement in the course. Because students face difficulties in interpreting graphs produced by contemporary dashboards (Matcha et al., 2019), such applications should be made with caution and be guided by educators.

6.6 Validation of the Study and its Limitations

We believe the research has external validity that relates to online courses. The study focuses on chemistry, but apart from the specific scientific content, the parameters examined are not unique to this field. Therefore, we think that the research findings can be generalized to other online courses in science taught with a similar pedagogy.

This study has several limitations. Because it is impossible to assign the learners randomly, we used convenience sampling based on learners' registration for the courses. Some of the other limitations of the study stemmed from the use of log files. The log file that we received from each institution was a basic Moodle log file. This was a limitation because we did not have, for example, detailed video sequence data. We think that if we had a more detailed log file, we could have learned about more aspects of SRL. For instance, Roll et al. (2011) evaluated students' help-seeking patterns from the log file. They integrated an intelligent tutoring agent for help-seeking into a tutoring system for geometry. The tutor recorded detailed log files of students' interactions with the tutor, which enabled them to evaluate the help-seeking dimension.

In addition, we did not crosslink the interviewees' answers with their Moodle log files or grades due to ethical considerations. Nevertheless, based on the qualitative analysis, in the analytical part of the study, we focused on two major parameters described in section 6.5 that enabled us to predict success in the online course.

Our statistical models were logistic regressions with dichotomist results (Y/N). First, we tried to run models that would give us ordinal outcomes. These models could have allowed us to distinguish between students who took the test and passed, students who took the test and did not pass, and students who did not take the test at all. Although our results were significant, this model was not statistically strong enough. Therefore, we decided not to use it in this study and instead to rely only on our dichotomic models.

In addition, from one of the case studies (case study 4) that we presented in chapter five, we learned about two teachers that studied together. If they did so from only one account, the log file of the other teacher would represent a learner who did not play the video lesson. Since we found significant statistical differences in learning patterns, we think that this was a marginal phenomenon in the course we studied. It also did not come up in other interviews. However, this is a specific example of a much more general problem of analyzing and reaching conclusions based on EDM techniques. As described in the discussion, we tried to cope with this shortcoming by applying a mixed-methods approach using several research tools and cautiously interpreting our findings.

6.7 Future Research

This study opens a window for further research on a variety of topics. We already shared a few ideas for future research during the discussion, such as the NLP analysis for the SOLO taxonomy and developing dashboards to identify students who are most likely to not complete the course.

In a prospective study, it will be possible to carry out an experimental study to determine whether a pedagogical change or SRL workshops given at an early stage can increase the completion rates. Previous research found that providing students with general information concerning SRL did not promote persistence (Kizilcec et al., 2017). Kizilcec suggested that integrating SRL into the learning process and course design could more effectively promote student persistence (Kizilcec et al., 2017). Support for this suggestion was provided by studies that found that SRL training improves learners' performance (Hermanns & Schmidt, 2018). Future research should also examine why students did not submit the optional assignment. This might be related to students' motivation, SRL, or difficulty with the course content.

The SCOP variable we developed in this study represents learners' weekly advancement in playing the course video lectures. Course designers and researchers may offer alternative variables that are more suitable for their specific courses. We also hope that other researchers will integrate the "multiple video play patterns" into future research. This variable might be important because one of the advantages of online learning is the possibility of rewatching video lectures. Future research should also consider utilizing various technological features integrated with video lectures, enabling more active learning to predict student success. Future research should examine the use of prediction models and the ability to plan for personalized support. Although this study examined chemistry courses, SRL characteristics are also relevant to other fields.

Finally, the findings from this research are especially relevant in the context of the ongoing COVID-19 pandemic. With the growing exposure of more learners to online learning, it can be assumed that more learners will choose an online education in the future. We believe that online learning will continue to be important in the post-COVID-19 world. This dramatic change highlights the need for more research on developing learning theories that promote more effective online learning. The study presented in this thesis is a step forward in this direction. We expect future research to further address the critical open questions in the field regarding students' persistence, the learning process, course design, and student-instructor interactions.

7. References

שקדי., א'. (2003).**מילים מנסות לגעת מחקר איכותני-תיאוריה ויישום** (מהדורה חמישית). אוניברסיטת תל-אביב: רמות

Admiraal, W., Huisman, B., & Pilli, O. (2015). Assessment in massive open online courses. *Electronic Journal of E-learning, 13(*4*),* 207-216.

Ahn, J., Campos, F., Hays, M., & DiGiacomo, D. (2019). Designing in Context: Reaching beyond Usability in Learning Analytics Dashboard Design. *Journal of Learning Analytics, 6(*2*),* 70- 85.

Akerson, V. L., Abd-El-Khalick, F., & Lederman, N. G. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *37(*4*),* 295-317.

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational Data Mining and Learning Analytics for 21stcentury higher education: A Review and Synthesis. *Telematics and Informatics. 37*, 13-49.

Aleven, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist, 45(*4*),* 224-233.

Alexandron, G., Armoni, M., Gordon, M., & Harel, D. (2016). Teaching Nondeterminism Through Programming. *Informatics in Education, 15(*1*),* 1-23*.*

Alexandron, G., Wiltrout, M. E., Berg, A., & Ruipérez-Valiente, J. A. (2020). Assessment that matters: Balancing reliability and learner-centered pedagogy in MOOC assessment. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 512-517).

Alexandron, G., Yoo, L. Y., Ruipérez-Valiente, J. A., Lee, S., & Pritchard, D. E. (2019). Are MOOC Learning Analytics Results Trustworthy? With Fake Learners, They Might Not Be!. *International Journal of Artificial Intelligence in Education*, *29(*4*)*, 484-506.

Alturkistani, A., Lam, C., Foley, K., Stenfors, T., Blum, E. R., Van Velthoven, M. H., & Meinert, E. (2020). Massive open online course evaluation methods: systematic review. *Journal of medical Internet research, 22(*4*),* e13851.

Amaral, K. E., Shank, J. D., Shibley Jr, I. A., & Shibley, L. R. (2013). Web-enhanced general chemistry increases student completion rates, success, and satisfaction. *Journal of chemical education, 90*(3), 296-302.

Angeli, C., Howard, S. K., Ma, J., Yang, J., & Kirschner, P. A. (2017). Data mining in educational technology classroom research: Can it make a contribution? *Computers & Education*, *113*, 226-242.

Angrave, L., Zhang, Z., Henricks, G., & Mahipal, C. (2020, February). Who Benefits? Positive Learner Outcomes from Behavioral Analytics of Online Lecture Video Viewing Using ClassTranscribe. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education.* (pp. 1193-1199).

Appleton, J. J. (2012). Systems consultation: Developing the assessment-to-intervention link with the Student Engagement Instruments. Christenson, A.L. Reschy, C. Wylie (Eds.), Handbook of research on student engagement, Springer, New York, NY (2012), pp. 725-741

Ariely, M., Nazaretsky, T., & Alexandron, G. (In press). Machine learning and Hebrew NLP for automated assessment of open-ended questions in biology. *International Journal of Artificial Intelligence in Education.*

Arora, Y., Singhal, A., & Bansal, A. (2014). A study of applications of RBF network *International Journal of Computer Applications, 94*(2), 17-20.

Artino Jr, A. R., & Jones II, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education*, *15*(3), 170-175.

Artino Jr, A. R., & Stephens, J. M. (2009). Academic motivation and self-regulation: A Comparative analysis of undergraduate and graduate students learning online. The *Internet and Higher Education, 12*(3-4), 146-151.

Aviran, E., Easa, E., Livne, S., & Blonder, R. (2020). Implementation of a personalized online learning system towards creating hybrid learning and teaching in chemistry classes. In G. D., C. A., & C. C. (Eds.), *Early warning systems and targeted interventions for student success in online courses* (pp. 90-110). IGI Global.

Baldwin, S. J., & Ching, Y. H. (2021). Accessibility in Online Courses: a Review of National and Statewide Evaluation Instruments. *TechTrends*, 1-12.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. *In Learning analytics* (61-75). Springer New York.

Baker, R., Xu, D., Park, J., Yu, R., Li, Q., Cung, B., Fischer, C., Rodriguez, F., Warschauer, M., & Smyth, P. (2020). The benefits and caveats of using clickstream data to understandstudent self-regulatory behaviors: opening the black box of learning processes. *International Journal of Educational Technology in Higher Education*, *17*(1), 1-24.

Barak, M. (2007). Transition from traditional to ICT-enhanced learning environments in chemistry courses. *Computers & Education*, *48*(1), 30-43.

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S. L. (2009). Measuring self-regulation in online and blended learning environments. *The internet and higher education, 12*(1), 1-6.

Barnea, N., & Dori, Y. J. (1999). High-school chemistry students' performance and gender differences in a computerized molecular modeling learning environment. *Journal of Science Education and Technology*, *8*(4), 257-271.

Battle, G. M., Allen, F. H., & Ferrence, G. M. (2010). Teaching three-dimensional structural chemistry using crystal structure databases. 2. teaching units that utilize an interactive web-accessible subset of the cambridge structural database. *Journal of Chemical Education*, *87* (8), 813-818.

Biggs, J.B. & Collis, K.F. (1982). Evaluating the Quality of Learning: The SOLO Taxonomy. New York: Academic Press.

Biggs, J. B. & Collis, K. F.: 1989, ‘Towards a Model of School-Based Curriculum Development and Assessment: Using the SOLO Taxonomy’, *Australian Journal of Education 33*, 149–161

Biggs, J., & Tang, C. (2015). Constructive alignment: An outcomes-based approach to teaching anatomy. *In Teaching anatomy* (31-38). Springer, Cham.

Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher education, 33*(1), 71-84.

Blonder, R. (2011). The story of nanomaterials in modern technology: An advanced course for chemistry teachers. *Journal of Chemical Education, 88*(1), 49-52.

Blonder, R., Jonatan, M., Bar-Dov, Z., Benny, N., Rap, S., & Sakhnini, S. (2013). Can You Tube it? Providing chemistry teachers with technological tools and enhancing their self-efficacy beliefs. *Chemistry Education Research and Practice*, *14*(3), 269-285.

Blonder, R., & Mamlok-Naaman, R. (2016). Learning about teaching the extracurricular topic of nanotechnology as a vehicle for achieving a sustainable change in science education. *International Journal of Science and Mathematics Education, 14*, 345-372.

Blonder, R., & Rap, S. (2017). I like Facebook: Exploring Israeli high school chemistry teachers’ TPACK and self-efficacy beliefs. *Education and Information Technologies*, *22*(2), 697-724.

Bouhnik, D., & Marcus, T. (2006). Interaction in distance-learning courses. *Journal of the American Society for Information Science and Technology, 57*(3), 299-305.

Bozkurt, A., Akgün-Özbek, E., & Zawacki-Richter, O. (2017). Trends and patterns in massive open online courses: Review and content analysis of research on MOOCs (2008-2015). *International Review of Research in Open and Distributed Learning: IRRODL*, *18*(5), 118-147.

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education, 15*(4), es6.

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Computers & Education, 96*, 42-54.

Çınar, A., Ince, E., Gezer, M., & Yılmaz, Ö. (2020). Machine learning algorithm for grading open-ended physics questions in Turkish. *Education and Information Technologies*, *25*(5), 3821-3844.

Clark, T. M., & Chamberlain, J. M. (2014). Use of a PhET interactive simulation in general chemistry laboratory: Models of the hydrogen atom. *Journal of Chemical Education*, *91*(8), 1198-1202.
Cohen, A. (2017). Analysis of student activity in web-supported courses as a tool for predicting dropout. *Educational Technology Research and Development*, *65*(5), 1285-1304.
Cohen, L., Manion, L., & Morrison, K. (2007). Research methods in education (Sixthth ed., pp. 253–263). London: Routledge Falmer
Cohen, A., Shimony, U., Nachmias, R., & Soffer, T. (2019). Active learners' characterization in MOOC forums and their generated knowledge. *British Journal of Educational Technology, 50*(1), 177-198.
Cook, E., Kennedy, E., & McGuire, S. Y. (2013). Effect of teaching metacognitive learning strategies on performance in general chemistry courses. *Journal of Chemical Education, 90*(8), 961-967.
Cormack, S. H., Eagle, L. A., & Davies, M. S. (2020). A large-scale test of the relationship between procrastination and performance using learning analytics. *Assessment & Evaluation in Higher Education*, *45*(7), 1046-1059.
Cosio, M. N., & Williamson, V. M. (2019). Timing of homework completion vs. performance general chemistry. *Journal of Science Education and Technology*, *28*(5), 523-531.
Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior, 73*, 247-256.
Cunningham, J. B. (1997). Case study principles for different types of cases. *Quality and quantity, 31*(4), 401-423.
Dalipi, F., Imran, A. S., & Kastrati, Z. (2018, April). MOOC dropout prediction using machine learning techniques: Review and research challenges. In *2018 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1007-1014). IEEE.
Davis, D., Chen, G., Hauff, C., & Houben, G. J. (2018). Activating learning at scale: A review of innovations in online learning strategies. *Computers & Education, 125*, 327-344.
Delgado, C., Stevens, S. Y., Shin, N., & Krajcik, J. (2015). A middle school instructional unit for size and scale contextualized in nanotechnology. *Nanotechnology Reviews, 4*(1), 51-69.
Dermy, O., & Brun, A. (2020). Can We Take Advantage of Time-Interval Pattern Mining to Model Students Activity?. Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020) *(*pp-69-80)
DeWaard, J., Kim, K., & Raymer, J. (2012). Migration systems in Europe: Evidence from harmonized flow data. *Demography*, *49*(4), 1307-1333.
Dietz-Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of interactive online learning, 12*(1), 17-26
Dorfman, B. S., & Fortus, D. (2019). Students' self-efficacy for science in different school systems. *Journal of Research in Science Teaching, 56*(8), 1037-1059.
Dorfman, B. S., Terrill, B., Patterson, K., Yarden, A., & Blonder, R. (2019). Teachers personalize videos and animations of biochemical processes: results from a professional development workshop. *Chemistry Education Research and Practice*, *20*(4), 772-786.
Dori, Y. J., Dangur, V., Avargil, S., & Peskin, U. (2014). Assessing Advanced High School and Undergraduate Students' Thinking Skills: The Chemistry From the Nanoscale to Microelectronics Module. *Journal of Chemical Education, 91*(9), 1306-1317
Dutt, A., Ismail, M. A., & Herawan, T. (2017). A systematic review on educational data mining. *IEEE Access, 5*, 15991-16005.
Eidelman, R. R., Rosenberg, J. M., & Shwartz, Y. (2019). Assessing the Interaction Between Self-Regulated Learning (SRL) Profiles and Actual Learning in the Chemistry Online Blended Learning Environment (COBLE). *In Learning Technologies for Transforming Large-Scale Teaching, Learning, and Assessment* (pp. 231-255). Springer, Cham.
Eitemüller, C., & Habig, S. (2020). Enhancing the transition?–effects of a tertiary bridging course in chemistry. *Chemistry Education Research and Practice*, *21*(2), 561-569.
Egloffstein M. (2018). Massive Open Online Courses in Digital Workplace Learning. In D. Ifenthaler (ed.), *Digital Workplace Learning. Bridging Formal and Informal Learning with Digital Technologies* (pp. 149-166).Springer, Cham.

El Aouifi, H., El Hajji, M., Es-Saady, Y., & Douzi, H. (2021). Predicting learner's performance through video sequences viewing behavior analysis using educational data-mining. *Education and Information Technologies*, 1-16.

Fagerland, M. W., & Hosmer, D. W. (2012). A generalized Hosmer–Lemeshow goodness-of-fit test for multinomial logistic regression models. *The Stata Journal*, *12*(3), 447-453.

Fang, L., & Zahiruddin, A. K. M. (2020). Dashboard for the E-Assessment and E-Feedback System for Aerospace Engineering Examination Preparation in Singapore. In Early Warning Systems and Targeted Interventions for Student Success in Online Courses (pp. 64-89). IGI Global.

Feldman-Maggor, Y., Rom, A., & Tuvi-Arad, I. (2016). Integration of open educational resources in undergraduate chemistry teaching–a mapping tool and lecturers' considerations. *Chemistry Education Research and Practice*, *17*(2), 283-295.

Gabbay, H., Cohen, A., & Festinger, E. (2020). Early prediction of persistence and Performance in online language courses. In D. Glick, A. Cohen, & C. Chang (eds), *Early Warning Systems and Targeted Interventions for Student Success in Online Courses* (pp. 200-217). IGI Global.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71.

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education*, *28*, 68-84.

Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: results of the thinking from the PCK summit. In A. Berry, P. J. Friedrichsen, & J. J. Loughran (Eds.), *Re-examining pedagogical content knowledge in science education* (pp. 28–42). New York, NY: Routledge.

Giannakos, M. N., Chorianopoulos, K., & Chrisochoides, N. (2015). Making sense of video analytics: Lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *International Review of Research in Open and Distributed Learning*, *16*(1), 260-283.

Gibson, D. C., & Ifenthaler, D. (2017). Preparing the next generation of education researchers for big data in higher education. In *Big data and learning analytics in higher education* (pp. 29-42). Springer, Cham.

Glick, D., Cohen, A., & Gabbay, H. (2020). Do Student Written Responses to Reflection Questions Predict Persistence and Performance in Online Courses? A Text Analysis Approach. In D. Glick, A. Cohen, & C. Chang (eds), *Early Warning Systems and Targeted Interventions for Student Success in Online Courses* (pp. 1-21). IGI Global.

Gregori, P., Martínez, V., & Moyano-Fernández, J. J. (2018). Basic actions to reduce dropout rates in distance learning. *Evaluation and program planning*, *66*, 48-52.

Gupta, S., & Sabitha, A. S. (2019). Deciphering the attributes of student retention in massive open online courses using data mining techniques. *Education and Information Technologies*, *24*(3), 1973-1994.

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. Metacognition and Learning, 2(2-3), 107-124. IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Hampton, D., & Pearce, P. F. (2016). Student engagement in online nursing courses. *Nurse educator*, *41*(6), 294-298.

Handoko, E., Gronseth, S. L., McNeil, S. G., Bonk, C. J., & Robin, B. R. (2019). Goal Setting and MOOC Completion: A Study on the Role of Self-Regulated Learning in Student Performance in Massive Open Online Courses. *International Review of Research in Open and Distributed Learning, 20*(3),40-58.

Hartshorne, R., Baumgartner, E., Kaplan-Rakowski, R., Mouza, C., & Ferdig, R. E. (2020). Special issue editorial: Preservice and inservice professional development during the COVID-19 pandemic. *Journal of Technology and Teacher Education*, *28*(2), 137-147.

Hassner. T., Wolf. L., Lerner, A., & Leitner, Y. (2014). Viewing the viewers: how adults with attentional deficits watch educational videos. *Journal of Attention Disorders*, *18*(7) 585-593.

Hermanns, J., & Schmidt, B. (2018). Developing and applying stepped supporting tools in organic chemistry to promote students' self-regulated learning. *Journal of Chemical Education*, *96*(1), 47-52

Hershkovitz, A., & Alexandron, G. (2020). Understanding the potential and challenges of Big Data in schools and education. *Tendencias pedagógicas*, (35), 7-17.

Hershkovitz, A., & Nachmias, R. (2011). Online persistence in higher education web-supported courses. *The Internet and Higher Education*, *14*(2), 98-106.

Hilliger, I., Ortiz-Rojas, M., Pesántez-Cabrera, P., Scheihing, E., Tsai, Y. S., Muñoz-Merino, P. J., Broos, T; Whitelock-Wainwright, A & Pérez-Sanagustín, M. (2020). Identifying needs for learning analytics adoption in Latin American universities: A mixed-methods approach. *The Internet and Higher Education*, *45*, 100726.

Hofstein, A., Carmi, M., & Ben-Zvi, R. (2003). The development of leadership among chemistry teachers in Israel. *International Journal of Science and Mathematics Education*, *1*(1), 39-65.

Holme, T. A. (2019). Reproducibility, Replication, and Generalization in Research about Teaching Innovation. *Journal of Chemical Education*. *96*, 2359−2360.

Huang, L., Li, S., Poitras, E. G., & Lajoie, S. P. (2021). Latent profiles of self-regulated learning and their impacts on teachers' technology integration. *British Journal of Educational Technology, 52*(2), 695-713.

Inan, F., Yukselturk, E., Kurucay, M., & Flores, R. (2017). The impact of self-regulation strategies on student success and satisfaction in an online course. *International Journal on E-learning*, *16*(1), 23-32.

Islam, M. Z., Sajjad, G. S., Rahman, M. H., Dey, A. K., Biswas, M. A. M., & Hoque, A. K. M. J. (2012). Performance comparison of modified LMS and RLS algorithms in de-noising of ECG signals. *International Journal of Engineering & Technology*, *2*(3), 466-468.

Islam, A. N. (2014). Sources of satisfaction and dissatisfaction with a learning management system in post-adoption stage: A critical incident technique approach. *Computers in Human Behavior*, *30*, 249-261.

Jackman, J. A., Cho, D. J., Lee, J., Chen, J. M., Besenbacher, F., Bonnell, D. A., & Cho, N. J. (2016). Nanotechnology education for the global world: training the leaders of tomorrow. *ACS Nano, 10*, 5595−5599

Jones, M. G., Blonder, R., Gardner, G. E., Albe, V., Falvo, M., & Chevrier, J. (2013). Nanotechnology and nanoscale science: Educational challenges. *International Journal of Science Education*, *35*(9), 1490-1512.

Jones, M. G., Blonder R. & Kähkönen, A.-L. (2020). Challenges in Nanoscience. *In Education. 21st Century Nanoscience–A Handbook: Public Policy, Education, and Global Trends* (Volume Ten).

Johnson, L., Becker, S. A., Estrada, V., & Freeman, A. (2014). NMC horizon report: 2014 K (pp. 1-52). The New Media Consortium.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they Seem. *Journal of computer assisted learning, 7*(2), 75-83.

Kay, R., & Kletskin, I. (2012). Evaluating the use of problem-based video podcasts to teach mathematics in higher education. *Computers & Education, 59*(2), 619-627.

Kapusta, J., Munk, M., Halvoník, D., & Drlík, M. (2019). User Identification in the Process of Web Usage Data Preprocessing. *International Journal of Emerging Technologies in Learning (iJET)*, *14*(09), 21-33.

Kennedy, G.E., Judd, T.S., Churchward, A., Gray, K. & Krause, K-L (2008). First year students' experiences with technology: Are they really digital natives? Australasian Journal of Educational Technology, 24(1), 108-122.

Kew, S. N., & Tasir, Z. (2021). Learning Analytics in Online Learning Environment: A Systematic Review on the Focuses and the Types of Student-Related Analytics Data. *Technology, Knowledge and Learning*, 1-23.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & education, 104*, 18-33.

Knapp, B., Bardenet, R., Bernabeu, M. O., Bordas, R., Bruna, M., Calderhead, B., & Deane, C. M. (2015). Ten simple rules for a successful cross-disciplinary collaboration. *PLoS Comput Biol, 11*(4), e1004214.*Biol 11*(4): e1004214.

Kovacs, G. (2016, April). Effects of in-video quizzes on MOOC lecture viewing. In *Proceedings of the third (2016) ACM conference on Learning@ Scale* (pp. 31-40).
Lakhal, S., & Khechine, H. (2021). Technological factors of students' persistence in online courses in higher education: The moderating role of gender, age and prior online course experience. *Education and Information Technologies*, *26*(3), 3347-3373.
Laudonia, I., Mamlok-Naaman, R., Abels, S., & Eilks, I. (2018). Action research in science education–an analytical review of the literature. *Educational action research, 26*(3), 480-495.
Leech N. L. and Onwuegbuzie A. J., (2007), An array of qualitative data analysis tools: a call for data analysis triangulation, *School Psychol. Quart., 22*, 557–584.
Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers& education*, *48*(2), 185-204.
Lemay, D. J., & Doleck, T. (2020). Grade prediction of weekly assignments in MOOCS: mining video-viewing behavior. *Education and Information Technologies*, *25*(2), 1333-1342.
Lewis, S. E., & Lewis, J. E. (2007). Predicting at-risk students in general chemistry: comparing formal thought to a general achievement measure. *Chemistry Education Research and Practice, 8*(1), 32-51.
Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education, 45*, 100727.
Li, Q., & Baker, R. (2018). The different relationships between engagement and outcomes across participant subgroups in massive open online courses. *Computers & Education*, *127*, 41-65.
Lin, S. F., Chen, J. Y., Shih, K. Y., Wang, K. H., & Chang, H. P. (2015). Science teachers' perceptions of nanotechnology teaching and professional development: a survey study in Taiwan. *Nanotechnology Reviews*, *4*(1), 71-80.
Liñán, L. C., & Pérez, Á. A. J. (2015). Educational Data Mining and Learning Analytics: differences, similarities, and time evolution. *International Journal of Educational Technology in Higher Education*, *12*(3), 98-112.
Lister, R., Simon, B., Thompson, E., Whalley, J. L., & Prasad, C. (2006). Not seeing the forest for the trees: novice programmers and the SOLO taxonomy. *ACM SIGCSE Bulletin, 38*(3), 118-122.
Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). Designing professional development for teachers of science and mathematics. Corwin press.
Lu, T., Bradlow, E., & Hutchinson, W. (2017). Binge consumption of online content. *Preprint on webpage at https://faculty. wharton. upenn. edu/wp-content/uploads/2017/07/JoyLu_ JobMarketPaper_2017. pdf*.
Lu, O. H., Huang, A. Y., Huang, J. C., Lin, A. J., Ogata, H., & Yang, S. J. (2018). Applying learning analytics for the early prediction of Students' academic performance in blended learning. *Journal of Educational Technology & Society*, *21*(2), 220-232.
Luna, J. M., Castro, C., & Romero, C. (2017). MDM tool: A data mining framework integrated into Moodle. *Computer Applications in Engineering Education*, *25*(1), 90-102.
MacFarland, T. W., & Yates, J. M. (2016). Mann–whitney u test. In *Introduction to nonparametric statistics for the biological sciences using R* (pp. 103-132). Springer, Cham.
Magno, C. (2011). Validating the academic self-regulated learning scale with the motivated strategies for learning questionnaire (MSLQ) and learning and study strategies inventory (LASSI). *The International Journal of Educational and Psychological Assessment*, *7*(2), 56-73.
Mamlok-Naaman, R., Blonder, R., & Hofstein, A. (2010). Providing chemistry teachers with opportunities to enhance their knowledge in contemporary scientific areas: a three-stage model. *Chemistry Education Research and Practice*, *11*(4), 241-252.
Mamlok-Naaman, R., & Eilks, I. (2012). Different Types of action research to promote chemistry teachers' professional development –A joined theoretical reflection on two cases from Israel and Germany. *International Journal of Science and Mathematics Education,10*(3), 581-610.
Mamlok-Naaman, R., Eilks, I., Bodner, G., & Hofstein, A. (2018). Professional development of chemistry teachers: Theory and practice. Royal Society of Chemistry.

Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology, 5*(9), 1315-1316.
Martin, F., Ritzhaupt, A., Kumar, S., & Budhrani, K. (2019). Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation. *The Internet and Higher Education*, *42*, 34-43.
Matcha, W., Gašević, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226-245.
McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276-282.
Miah, S. J., Miah, M., & Shen, J. (2020). Editorial note: Learning management systems and big data technologies for higher education. *Education and Information Technologies, 25*(2), 725-730.
Michaeli, S., Kroparo, D., & Hershkovitz, A. (2020). Teachers' Use of Education Dashboards and Professional Growth. *The International Review of Research in Open and Distributed Learning, 21*(4), 61-78.
Michinov, N., Brunot, S., Le Bohec, O., Juhel, J., & Delaval, M. (2011). Procrastination, participation, and performance in online learning environments*. Computers & Education, 56*(1), 243-252.
Milligan, S. K., & Griffin, P. (2016). Understanding learning and learning design in MOOCs: A measurement-based interpretation. *Journal of Learning Analytics, 3*(2), 88-115.
Milligan, C., & Littlejohn, A. (2014). Supporting professional learning in a massive open online course. *The International Review of Research in Open and Distributed Learning 15*(5), 197-213.
Minogue, J., & Jones, G. (2009). Measuring the impact of haptic feedback using the SOLO taxonomy. *International Journal of Science Education, 31(*10), 1359-1378.
Morland, D. V., & Bivens, H. (2004). Designing instructional articles in online courses for adult learners. *Innovate: Journal of Online Education*, 1(2).
Narayanasamy, S. K., & Elçi, A. (2020). An effective prediction model for online course dropout rate. *International Journal of Distance Education Technologies (IJDET)*, *18*(4), 94-110.
Nawrot, I., & Doucet, A. (2014, April). Building engagement for MOOC students: introducing support for time management on online learning platforms. *In Proceedings of the 23rd International Conference on world wide web* (pp. 1077-1082).
Onah, D. F., Sinclair, J. E., & Boyatt, R. (2014, November). Exploring the use of MOOC discussion forums. *In Proceedings of London International Conference on Education* (pp. 1-4).
Onchiri, S. (2013). Conceptual model on application of chi-square test in education and social sciences. *Educational Research and Reviews*, *8*(15), 1231-1241.
Osborne, J. (2015). Best practices in logistic regression. SAGE Publications, Ltd
Paul, P., Pennell, M. L., & Lemeshow, S. (2013). Standardizing the power of the Hosmer–Lemeshow goodness of fit test in large data sets. *Statistics in medicine, 32*(1), 67-80.
Pappa, E. T., Pantazi, G., Tsaparlis, G., & Byers, B. (2021). Using Static Colored Visual Representations of Chemical Bonding: An Analysis of Students' Responses Using the SOLO Taxonomy. In *Book of abstracts* (p. 22).
Paul, P., Pennell, M. L., & Lemeshow, S. (2013). Standardizing the power of the Hosmer–Lemeshow goodness of fit test in large data sets. *Statistics in medicine*, *32*(1), 67-80.
Pelánek, R., Rihák, J., & Papoušek, J. (2016, April). Impact of data collection on interpretation and evaluation of student models. *In Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 40-47). ACM
Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert systems with applications, 41*(4), 1432-1462.
Pintrich, P. R. (2000). Multiple goals, multiple pathways: the role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544e555.
Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review, 16*(4), 385-407.
Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive

engagement in classroom academic tasks. *Student perceptions in the classroom*, *7*, 149-183.
Rabin, E., Kalman, Y. M., & Kalz, M. (2019). An empirical investigation of the of learner-centered outcome measures in MOOCs. *International Journal of Educational Technology in Higher Education 16*(1), 14.
Raju, D., & Schumacker, R. (2015). Exploring student characteristics of retention that lead to graduation in higher education using data mining models. *Journal of College Student Retention: Research, Theory & Practice*, *16*(4), 563-591.
Rakes, G. C., & Dunn, K. E. (2010). The Impact of Online Graduate Students' Motivation and Self-Regulation on Academic Procrastination*. Journal of Interactive Online Learning, 9(1).*
Ramírez-Gallego, S., Krawczyk, B., García, S., Woźniak, M., & Herrera, F. (2017). A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing*, *239*, 39-57.
Rap, S., & Blonder, R. (2016). Let's face(book) it: Analyzing interactions in social network groups for chemistry learning. *Journal of Science Education and Technology. 25*(1), 62-76.
Rap, S., & Blonder, R. (2017). Thou shall not try to speak in the Facebook language: Students' perspectives regarding using Facebook for chemistry learning. *Computers & Education, 114*, 69-78.
Read, D., & Lancaster, S. (2012). Unlocking video: 24/7 learning for the iPod generation. *Education in chEmistry, 49*(4), 13-16.
Richards-Babb, M., Curtis, R., Ratcliff, B., Roy, A., & Mikalik, T. (2018). General chemistry student attitudes and success with use of online homework: Traditional-responsive versus adaptive-responsive. *Journal of chemical education*, *95*(5), 691-699.
Roco, M. C. (2001). International strategy for nanotechnology research. *Journal of Nanoparticle Research, 3*(5), 353-360.
Rodrigues, M. W., Isotani, S., & Zarate, L. E. (2018). Educational Data Mining: A review of evaluation process in the e-learning. *Telematics and Informatics*, *35*(6), 1701-1717.
Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics, 2*(1), 7-12.
Romero, C., Romero, J. R., & Ventura, S. (2014). A survey on pre-processing educational data. *In Educational data mining* (pp. 29-64). Springer, Cham.
Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, *Part C (Applications and Reviews), 40*(6), 601-618.
Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1355.
Salmon, G., Gregory, J., Lokuge Dona, K., & Ross, B. (2015). Experiential online development for educators: The example of the Carpe Diem MOOC. *British Journal of Educational Technology, 46*(3), 542-556.
Sakhnini, S., & Blonder, R. (2015). Essential concepts of nanoscale science and technology for high school students, based on a Delphi study by the expert community. *International Journal of Science Education. 37(11)*, 1699–1738.
Sakhnini, S., & Blonder, R. (2016). Nanotechnology applications as a context for teaching the essential concepts of NST. *International Journal of Science Education*, *38*(3), 521-538.
Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of educational research*, *77*(4), 454-499.
Shaked, N., Sahar-Inbar, L., Peleg, G., Einhoren, B., & Eyal, H. (2020). Assessing Online Teaching: A Peer-Review Methodology in a Multidisciplinary Setting. *Proceedings of the 16th Chais Conference for the Study of Innovation and Learning Technologies: Learning in the Digital Era*.I. Blau, A. Caspi, Y. Eshet-Alkalai, N. Geri, Y. Kalman, T. Lauterman (Eds.), Ra'anana, Israel: The Open University of Israel
Shea, P., & Bidjerano, T. (2019). Effects of Online Course Load on Degree Completion, Transfer, and Dropout Among Community College Students of the State University of New York. Online Learning, 23(4). *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(6), 601-618.

Shelton, B. E., Hung, J. L., & Baughman, S. (2016). Online graduate teacher education: Establishing an EKG for student success intervention. *Technology, Knowledge and Learning*, *21*(1), 21-32.
Sgouros, G., & Stavrou, D. (2019). Teachers' professional development in Nanoscience and nanotechnology in the context of a *Community of Learners. International Journal of Science Education, 41*(15), 2070-2093.
Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard educational review, 57(1)*, 1-23.
Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, *57*(10), 1380- 1400.
Siemens, G., & Baker, R. S. D. (2012, April). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254).
Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses. *Journal of Computer Assisted Learning, 35*(3), 378-389.
Soffer, T., Kahan, T., & Livne, E. (2017). E-assessment of online academic courses via students' activities and perceptions. *Studies in Educational Evaluation, 54*, 83-93.
Stephens, M., & Jones, K. M. (2014). MOOCs as LIS professional development platforms: Evaluating and refining SJSU's first not-for-credit MOOC. *Journal of Education for Library and Information Science*, 345-361.
Taitelbaum, D., Mamlok-Naaman, R., Carmeli, M., & Hofstein, A. (2008). Evidence for teachers' change while participating in a continuous professional development programme and implementing the inquiry approach in the chemistry laboratory. *International Journal of Science Education. 30*(5), 593-617.
Tan, K., Dawson, V., & Venville, G. (2008). Use of cognitive organisers as a self regulated learning strategy. *Issues in Educational Research*, *18*(2), 183-207.
Taranto, D., & Buchanan, M. T. (2020). Sustaining lifelong learning: A self-regulated learning (SRL) approach. *Discourse and Communication for Sustainable Education, 11*(1), 5-15.
Trowler, V. (2010). Student engagement literature review. *The higher education academy*, *11*(1), 1-15.
Tsaparlis, G., Pappa, E. T., & Byers, B. (2018). Teaching and learning chemical bonding: research-based evidence for misconceptions and conceptual difficulties experienced by students in upper secondary schools and the effect of an enriched text. *Chemistry Education research and practice*, *19*(4), 1253-1269.
Tømte, C. E. (2019). MOOCs in teacher education: institutional and pedagogical change? *European Journal of Teacher Education*, *42*(1), 65-81.
Tuvi-Arad, I., & Blonder, R. (2019). Technology in the service of pedagogy: Teaching with chemistry databases. *Israel Journal of Chemistry*, *59*(6-7), 572-582.
Wang, W., Guo, L., He, L., & Wu, Y. J. (2019). Effects of social-interactive engagement on the dropout ratio in online learning: insights from MOOC. *Behaviour & Information Technology*, *38*(6), 621-636.
Warner, J., Doorenbos, J., Miller, B., & Guo, P. J. (2015, March). How High School, College, and Online Students Differentially Engage with an Interactive Digital Textbook. In *EDM* (pp. 528-531).
Watted, A., & Barak, M. (2018). Motivating factors of MOOC completers: Comparing between university-affiliated students and general participants. *The Internet and Higher Education, 37*, 11-20.
Wolters, C. A., & Taylor, D. J. (2012). A self-regulated learning perspective on student engagement. In *Handbook of research on student engagement* (pp. 635-651). Springer, Boston, MA.
Yair, G., Rotem, N., & Shustak, E. (2020). The riddle of the existential dropout: lessons from an institutional study of student attrition. *European Journal of Higher Education, 10*(4), 436-453.
Yonai, E., & Blonder, R. (2020). Scientists suggest insertion of nanoscience and technology into middle school physics. *Physical Review Physics Education Research, 16*(1), 010110.
Yoo, J., Lee, J., & Lee, D. (2020). A verification of motivations for over-the-top binge and short viewing of audio-visual content. *New Review of Hypermedia and Multimedia*, 1-30.

Yoon, M., Lee, J., & Jo, I. H. (2021). Video learning analytics: Investigating behavioral patterns and learner clusters in video-based online learning. *The Internet and Higher Education*, *50*, 100806.
You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education, 29*, 23-30.
You, J.W. (2015). Examining the effect of academic procrastination on achievement using LMS data in e-learning. Educational Technology & Society, 18(3), 124–134.
You, J.W., & Kang, M. (2014). The role of academic emotions in the relationship between perceived academic control and self-regulated learning in online learning. *Computers & Education, 77*, 125–133.
Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, *27*, 44-53.
Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. *In Handbook of self-regulation* (pp. 13-39). Academic Press.
Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American educational research journal, 45*(1), 166-183

8. Appendices

8.1 Appendix 1: Interview Protocol for Chapters 4-6

General questions:

1. Why did you choose to take this course?
2. Do you have previous experience in online learning?
3. Would you choose to take an online course again in the future? If yes, why?
4. What are your thoughts on the course?
5. Did you attend the live session sessions? If yes, were you active in them (e.g., asking questions, participating in discussions)?
6. Tell me about your learning throughout the course. For example, did you submit the assignments? Did you follow the schedule?
7. What were the advantages and disadvantages of the online course platform?
8. Did you skip certain parts of the course?

Video Session recordings

9. Did you watch the video session recordings?
10. How do you watch the lectures?
11. Did you watch the recording in its entirety? Do you watch the lecture after a live session or before submitting an assignment?

12. What made you stop watching the session before it ended?
13. How many times do you watch a particular session?

Course website

14. Beyond the recordings that appear on the course website, what do you think about the other learning materials: presentations, forums, books, links. Did you use them? How?
15. Did you try using all the learning materials in the course?
16. How do you navigate the course website - by type of activity - video, links, forum, or by course chapters?
17. Can you give an example from the course where a technological tool was particularly effective for presenting content? (e.g., experiment, animation, simulation as part of the video)
18. In addition to the course website and meetings, the course was accompanied by printed books. Did you use them? If so, how? How did you divide the learning time between the printed and digital materials?
19. Did you use additional learning materials other than those offered in the course (books, other websites, etc.)?

Course communication

20. Did you study individually or in collaboration with other students in the course?
21. Who did you turn to when you did not understand something in the course?

Difficulties

22. Were there any difficulties in the course? If yes, what were they?
23. What helped you complete the course?
24. Do you know anyone who has not completed the course? If so, do you have any thoughts as to why?
25. Have you encountered any technical difficulties? If yes, can you give an example? Can you explain how you solved it?

Additional questions for teachers only:

26. Have you experienced teaching remotely?

27. Following the course, will you be interested in integrating elements of online teaching in your classes? If so, which features (for example, pre-recorded lectures, discussion forums, etc.)?
28. Following the course, will you be interested in integrating topics covered during the course into your teaching materials? If so, which?

8.2 Appendix 2. Models Evaluation

8.2.1 Logistic Regression for First Semester in 2020 at the OUI

Model A for semester 2020 A at the OUI: The results of model A are presented in Table 8.1. The logistic regression model for the 154 students was found to be statistically significant $\chi^2(6) = 44.374$, $p < .001$. The submission rate of the first optional assignment ($p< 0.01$) was found to be a significant parameter for predicting the final course success status ($p< 0.05$). The model explains 33% (Nagelkerke R^2) of the variance in the courses' success and correctly classifies 75.2% (see table 8.2) of the cases. The model is well fit data according to the Hosmer-Lemeshow test. These results suggest that starting at the 5th week, when students submit their first optional assignment, we can determine the probability that a specific student will complete the course.

Model B for semester 2020 A at the OUI: Model B's results, which are based on the SCOP variable as a predictor, are presented in the two rightmost columns of Table 8.1. It was found to be statistically significant, $\chi^2(6) = 20.297$, $p < .001$, suggesting that one can identify the probability to succeed in the courses based on the SCOP at the 7th week ($p< 0.01$). The model explains 16.2% (Nagelkerke R^2) of the variance in the courses' success and correctly classifies 64.3% (see table 8.2) of the cases. The model is well fit data according to the Hosmer-Lemeshow test.

Table 8.1 Models of logistic regressions of courses' success – 2020a. (N = 157). (Chapter 3)

	Model A (optional assignment submission)		**Model B (SCOP)**	
Variable	**Wald**	**Sig.**	**Wald**	**Sig.**
Advanced diploma	0.788	0.375	1.604	0.205
Course	1.713	0.191	1.590	0.207
District of residence (SES)	1.678	0.195	0.873	0.350
First course at the OUI	0.043	0.836	0.555	0.456
Gender	1.143	0.285	1.555	0.212
SCOP at week 7	----	-----	12.919	**0.000
Submission of the first optional assignment at week 5	32.231	** 0.000	----	----

**p<0.01

Table 8.2 Actual and predicted classifications of course completion – 2020a. (N = 157). (Chapter 3)

	Model A Predictions			**Model B Predictions**		
Actual Status	**Improbable to Complete**	**Probable to Complete successfully**	**Correct Predictions (%)**	**Improbable to Complete**	**Probable to Complete successfully**	**Correct Predictions (%)**
Did not Complete	43	26	62.3	31	38	44.9
Completed successfully	13	75	85.2	18	70	79.5
Overall Percentage			75.2			64.3

8.2.2 ROC Curves

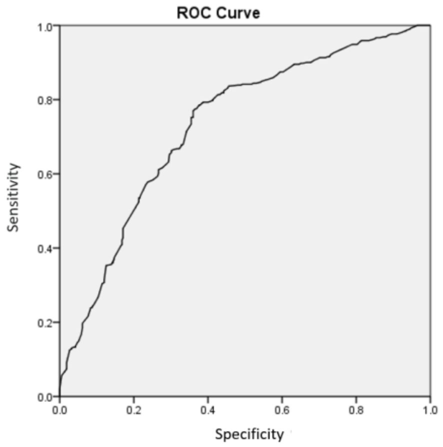


Figure 8.1. ROC curve for model A - Chapter 3. The area under the curve is 0.731

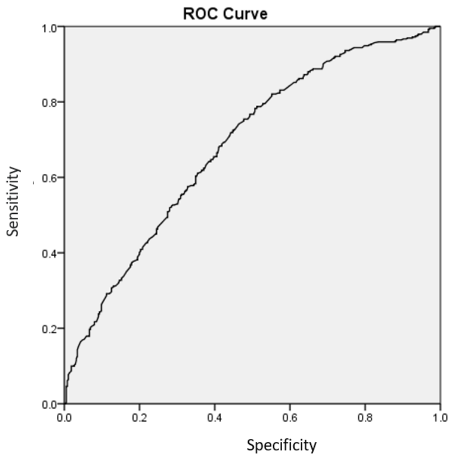


Figure 8.2. ROC curve for model B - chapter 3. The area under the curve is 0.683

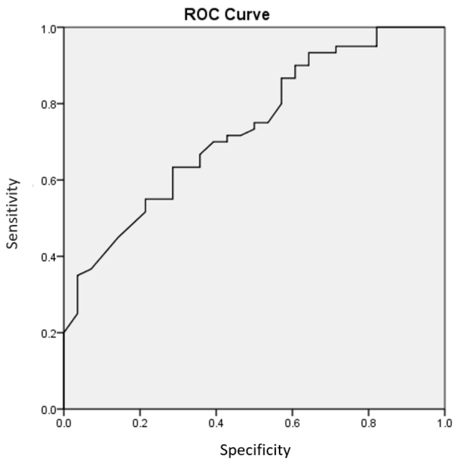


Figure 8.3. ROC curve for the Model in chapter 5. The ROC curve for Model B. The area under the curve is 0.735

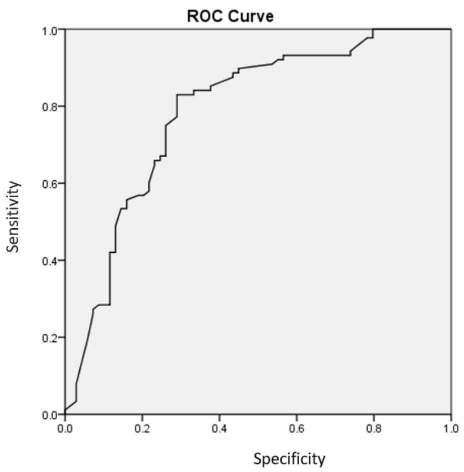


Figure 8.4. ROC curve for model A – 2020 A – the OUI - The area under the curve is 0.786

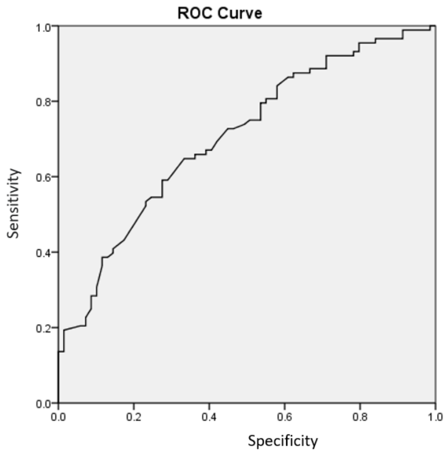


Figure 8.5. ROC curve for model B – 2020 A – the OUI - The area under the curve is 0.705

8.3 Appendix 3: Pre-Post Questionnaire for Chapter 5

1. What is a nanometer? Can you give an example of an object that is nanometer in size?
2. What are nanotubes?
3. Explain the concept of wave interference.
4. What is the Scanning Electron Microscope (SEM)?
5. What is the Transmission Electron Microscope (TEM)?
6. What is the Atomic Force Microscope (AFM)?
7. What is the Scanning Tunneling Microscope (STM)?
8. Explain the phenomenon of Tunneling?
9. What is meant by wavy and particle properties of an electron?
10. What is the color of material?
11. What is fullerene?
12. What is Self-Assembly?
13. How are nanomaterials produced?
14. What properties of a material depend on its size? Explain.
15. Indicate applications of or new developments in nanotechnology
16. Indicate risks involved in the development of nanotechnology
17. How do the topics taught in the course relate to your teaching in the classroom?

8.4 Appendix 4: Binge Analysis for the OUI

11 (1%) of the students (n=954) in our sample exhibited the binge pattern, 8 of whom successfully completed the course. 3 of the students who binged the course skipped between 1 and 2 video sessions.

8.5 Appendix 5: Video Index

Table 8.3 Index of video numbers in the Introduction to materials and nanotechnology course

Lesson	Video number	Description
1	1	What are nanotechnologies?
2	2	Size and scale
3	3	Size-dependent properties
4	No Video at this lesson	
5	5A	How it all started - quantum mechanics
	5B	The photoelectric effect-quantum mechanics
	5C	Interference - quantum mechanics
6	6A	Schrödinger equation – Part A
	6B	Schrödinger equation – Part B
7	7A	The Hamiltonian
	7B	Particle in a box
	7C	Particle in a box – examples
	7D	Particle in a box – tunneling
	7E	Orbitals and chemical bonding
8	8	Quantum dots
9	9A	Introduction – to see nano
	9B	Atomic force microscopy – AFM
10	10A	Scanning electron microscope – Part A
	10B	Scanning electron microscope – Part B
11	11A	Fabrication- preparation approaches to nanoparticles
	11B	Fabrication- self assembly
12	12A	Innovation: nanoparticles with antibacterial properties
	12B	Silver nano-particle
13	13	Classification of nanomaterials

8.6 Appendix 6: Publications and Conference Presentations during the PhD Research

Paper Published in Peer-Reviewed Journal

Feldman-Maggor, Y., Barhoom, S., Blonder, R., & Tuvi-Arad, I. (2021). Behind the scenes of educational data mining. *Education and Information Technologies, 26*(2), 1455-1470.

Rap, S., **Feldman-Maggor, Y**., Aviran, E., Shvarts-Serebro, I., Easa, E., Yonai, E., Waldman R & Blonder, R. (2020). An applied research-based approach to support Chemistry teachers during the COVID-19 pandemic. *Journal of Chemical Education, 97*(9), 3278-3284.

Under Review

Feldman-Maggor, Y., Blonder, R., & Tuvi-Arad, I. Let Them Choose: Optional assignments and online learning patterns as predictors of success in online general chemistry courses, **Under Review.**

Feldman-Maggor, Y., Tuvi-Arad, I & Blonder, R. A Multi-Dimensional Course Evaluation Framework for Online Professional Development of Chemistry Teachers, **Under Review.**

Blonder, R; **Feldman-Maggor Y** & Rap, S. Online Instruction by advanced degrees' lecturers in the natural sciences during the COVID-19 breakout: Development of TPACK and self-efficacy, **Under Review**

Presentations at conferences

Paper Presentation, **Feldman-Maggor, Y;** Blonder, R. & Tuvi-Arad I. "Identifying significant indicators that predict success in online general chemistry courses", Eurovariety 2021, to be held online, 7-9 July 2021.

Paper Presentation, **Feldman-Maggor, Y**, Tuvi-Arad I, &. Blonder, R. "Design Principles and Evaluation of an Online Nanotechnology Professional Development Course for Teachers," NARST 21 A global organization for improving science education through research, Online Conference, 7-10 April 2021

Paper Presentation, **Feldman-Maggor, Y**., Blonder, R., & Tuvi-Arad, I. "Examining MOOC courses for teacher professional development". Presented at the 18th Annual MEITAL National Conference: New Technologies and their Evaluation in Online Teaching and Learning (held online), 1July 2020.

Paper Presentation, **Feldman-Maggor, Y**, Tuvi-Arad I, &. Blonder, R. "Self-regulated learning as a supportive tool for online learning". Presented in the Annual Israeli Chemistry Teachers Conference (held online), 29 June 2020.

Paper Presentation, **Feldman-Maggor, Y**., Barhoom, S., Blonder, R., & Tuvi-Arad, I. (2020). "Behind the scenes of educational data mining". Proceedings of the 15th Chais, Conference for the Study of Innovation and Learning Technologies: Learning in the Digital Era, Ra'anana, Israel: The Open University of Israel,11 February 2020.

Paper Presentation, **Feldman-Maggor, Y**, Tuvi-Arad I, &. Blonder, R. "Online nanotechnology courses for teachers: learning evaluation and learning patterns", ESERA 19: European Science Education Research Association, Bologna, Italy, 26-30 August 2019.

Poster Presentation, Feldman-Maggor, Y, Tuvi-Arad, I & Blonder R. "The Participation Patterns of Chemistry Teachers in an Online Nanotechnology Course: Learning Evaluation and the use of course materials". Presented in: Nanao.IL.2018, Jerusalem, 9 October 2018.

Paper Presentation, **Feldman-Maggor, Y,** Tuvi-Arad, I & Blonder R. "The Participation Patterns of Chemistry Teachers in an Online Nanotechnology Course". Presented in the 83 Annual Meeting of the Israel Chemical Society held at David intercontinental hotel, Tel-Aviv, 13-14 February 2018.

פרסומים בעברית:

פלדמן-מגור י';טובי-ערד ע; בלונדר ר'. (התקבל). מאפיינים של קורסי MOOC והפוטנציאל הטמון בהם לפיתוח מקצועי של מורים. בתוך ע' כהן, ג' רביד, ר' בלונדר, א' פוקוש-ברוך, ח' מישר-טל (ערוכים). **טכנולוגיות למידה בהשכלה הגבוהה.** בהוצאת מיט"ל.

בלונדר ר'., פלדמן-מגור י'; רפ ש'. (התקבל). הוראה מקוונת של מרצים לתארים מתקדמים במדעי הטבע בתקופת הקורונה: התפתחות ידע ומסוגלות עצמית. בתוך ע' כהן, א' ברונשטיין (עורכות) ממשבר להזדמנות - למידה והוראה במשבר הקורונה בהשכלה הגבוהה בישראל.

9. Scholarships and Funding

- The Zalman Aran Scholarship for Outstanding Doctoral Students, *Awarded by the Chief Scientist, Israeli Ministry of Education (2018-2021)*
- Israeli Ministry of Immigration Research Scholarship (2017)
- Supported the Open University of Israel research fund (grant No. 507441)
- Supported by the i-PEN Project (Erasmus + Programme of the European union)