Mining Activity Tracker Data to Analyse Physical Activity Behaviours and Provide Personalised Feedback in Health Education Programmes

CLAUDIO ESTEBAN DIAZ CIFUENTES

THE UNIVERSITY OF SYDNEY

Supervisor: Kalina Yacef
Associate Supervisor: Corinne Caillaud

A thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy

Faculty of Engineering
The University of Sydney
Australia

26 April 2023

Abstract

The use of activity trackers in health education offers great potential to objectively assess healthy behaviour learning. However, the current models and techniques used to exploit and analyse their data mainly compare highly aggregated amounts of physical activity, thus diluting the fine physical activity patterns contained in the granular tracker data. The extraction of these latent patterns is valuable for health education because they may help to better understand how interventions affect behaviours and how they change over time. Physical activity patterns offer a more detailed view of how participants learn and offer the possibility to encourage healthy behaviours when detected on time.

This thesis proposes data mining models and techniques to extract these finer physical activity patterns that can then be used to analyse physical activity behaviours. This allows performing elaborate intervention assessments and generating timely personalised feedback during interventions.

In Chapter 1, we introduce the research questions and objectives. In Chapter 2, we present a systematic literature review of the current state of data mining techniques that use physical activity sensor data in health education to detect behaviour changes. We discuss common challenges and opportunities to guide future work. In Chapter 3, we propose a data mining method that highlights the nature and timing of behaviour changes for a more insightful assessment of health interventions. In Chapter 4, we describe U-BEHAVED, an unsupervised machine learning technique to detect significant physical activity behaviour changes and to determine whether they become habitual as the health intervention unfolds. In Chapter 5, we model physical activity behaviour changes to provide personalised feedback. Finally, we discuss our work and describe how it answered the research questions.

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes. I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Claudio Esteban Diaz Cifuentes

Statement of Authorship

This thesis contains material from:

C. Diaz, C. Caillaud and K. Yacef, ‘Mining sensor data to assess changes in physical activity behaviors in health interventions: Systematic review,’ *JMIR Medical Informatics*, 2022.
Chapter 2 consists entirely of this paper. I performed the systematic review under the supervision of Corinne Caillaud and Kalina Yacef. I wrote this review article under the guidance of Corinne Caillaud and Kalina Yacef.

C. Diaz and K. Yacef, ‘Detecting behaviour changes in accelerometer data,’ in *Knowledge Discovery in Healthcare Data (KDH@IJCAI-ECAI Stockholm)*, vol. 2148, 2018, pp. 21–26.
Chapter 3 includes material from this publication, and was the foundation of [34]. Specifically, Sections 3.1, 3.4, 3.5.1 and 3.6.1 are based on contents from this article. All results, including figures, were obtained by data analysis using a code that I wrote. I conceived and developed all research ideas and methods under the supervision of Kalina Yacef. I wrote the article under the guidance of Kalina Yacef.

C. Díaz, O. Galy, C. Caillaud and K. Yacef, ‘A clustering approach for modeling and analyzing changes in physical activity behaviors from accelerometers,’ *IEEE Access*, vol. 8, pp. 224 123–224 134, 2020.
Chapter 3 consists entirely of this paper that is an extension of [33]. All results, including the figures, were obtained from data analysis using a code that I wrote. I conceived and developed all research ideas and methods under the supervision of Corinne Caillaud and Kalina Yacef. I wrote the article under the guidance of Olivier Galy, Corinne Caillaud and Kalina Yacef.

C. Diaz, C. Caillaud and K. Yacef, 'Unsupervised early detection of physical activity behaviour changes from wearable accelerometer data,' *Sensors*, vol. 22, no. 21, 2022.
Chapter 4 consists entirely of this paper. All results, including figures, were obtained from data analysis using a code that I wrote. I conceived and developed all research ideas and methods under the supervision of Corinne Caillaud and Kalina Yacef. I wrote the article under the guidance of Corinne Caillaud and Kalina Yacef.

C. Díaz, O. Galy, C. Caillaud and K. Yacef, 'Modelling physical activity behaviour changes for personalised feedback in a health education application,' *Proceedings of the 30th International Conference on Computers in Education*, pp. 369–374, 2022.
Chapter 5 contains parts of this publication, with substantial additions. All results, including figures, were obtained from data analysis using a code that I wrote. I conceived and developed all research ideas and methods under the supervision of Corinne Caillaud and Kalina Yacef. I wrote this article under the guidance of Olivier Galy, Corinne Caillaud and Kalina Yacef.

Attestation of Authorship Statement

As supervisors for the candidature upon which this thesis is based, we can confirm that the authorship attribution statements above are correct.

Kalina Yacef

Corinne Caillaud

Acknowledgements

From the moment I began my PhD journey, I have met extraordinary people who have supported, helped and guided me on this adventure, and whom I would like to recognize.

Firstly, I would like to express my deepest gratitude to my supervisors, Kalina Yacef and Corinne Caillaud. Kalina supported me from the beginning of my doctorate, constantly advising me to make the best decisions, sharing her wisdom, and always having time and patience to answer all my questions. Corinne supported and guided me into the world of health and sports sciences with her invaluable knowledge, constantly motivating me and providing me with unconditional support. I feel tremendously lucky and honoured to have had them both as supervisors.

I am also very grateful to my collaborator Olivier Galy from the University of New Caledonia. Thank you for your continuous support and assertive feedback. I would also like to thank Emilie Paufique, Lauren Ha, Susan Ledger, Gael Clerc and everyone who was involved and participated in the iEngage project.

I would like to thank my friends and research colleagues. First, to my dear friends Jessica McBroom and Augusto Dias. With Jess, we spent innumerable hours chatting about the upside downs of life, our PhD journey, and her enormous love for Quokkas. With Augusto, we shared countless fascinating hallway chats, providing me valuable feedback about my research and which also taught me about the interesting Forro dance. I also would like to thank my friends Soojeong Yoo, Jisu Jung, Lie Min Tang, Chris Wong, Miad Zandavi, Waheeb Yaqub, Bernard Bucalon and Kim Bente - You have all made my PhD an unforgettable experience.

I would also like to thank Irena Koprinska, Judy Kay, Bob Kummerfeld and all the students and academics from the Computer-Human Adapted Interaction (CHAI) Research Group for fueling me with their ideas and different research topics.

I am also very grateful to the administrative staff, especially to Evelyn Riegler, Katie Yang and Cecille Faraizi, who supported me during my candidature. I am also very thankful to Dr. Elisabetta Andermarcher, who professionally proofread all the chapters of this thesis.

Thank you to the Universidad Adolfo Ibañez and CONICYT (Comisión Nacional de Investigación Científica y Tecnológica/Chilean National Commission for Scientific and Technological Research) "Becas Chile" Doctoral Fellowship program Grant No. 72200111 for their support.

And I would also like to especially thank Jorge Villalon and Alejandra Beghelli, who have always supported me from the day I met them, and who gave me crucial advice that allowed me to decide to do the PhD. I am enormously grateful to you both. I would also like to extend this sincere thank you to the late Alejandro Jadresic, who was Dean of the Engineering Faculty of the Universidad Adolfo Ibañez and who always pushed me to follow my PhD and supported me in this endeavour.

I want to show my gratitude to my Chilean and Australian friends. Thank you for all the laughs, smiles and support throughout these years!

And finally, I wish to thank my family. To my grandparents, uncles, aunts and cousins, for your thoroughgoing cherished support from Chile, especially during the lockdowns. To my Australian family: Maria, Francisco and Nicole, for always loving and caring for me and my family. To my parents-in-law, Fernando and Silvana, and my sisters-in-law, Constanza and Fernanda for your immense affection. To my parents, Tito and Luisa, and my brother Francisco, thank you for your endless love and constant encouragement throughout my life. To my wife Francesca, for being my favourite person and my soulmate, who has unconditionally supported all my endeavours since the beginning of our relationship. This PhD would have been impossible to do without you. And to my daughter Isabella, thank you for your pristine love and for inspiring me every day.

Contents

Abbreviations

IQR	Interquartile Range
RQ	Research Question
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
AIED	Artificial Intelligence in Education
U-BEHAVED	Unsupervised Behaviour and Habits Very Early Detector
SVMgs	Gravity-subtracted Signal Vector Magnitudes
GPS	Global Positioning System
MET	Metabolic Equivalents
PAL	Physical Activity Levels
MVPA	Moderate to Vigorous Physical Activity
BD	Baseline Dataset

CHAPTER 1

Introduction

In recent years, technology has been playing an increasing and central role in supporting health education interventions. Physical activity trackers have been widely adopted due to their low cost and ability to unobtrusively and objectively monitor the participants' daily physical activity [84]. As they capture physical activity at a very high level of detail (e.g. 60Hz, second, or minute), the participants' steps and time spent at different physical activity intensities can be accurately measured in free-living conditions. Such rich, objective, and unobtrusive data offer an unprecedented opportunity to better support health education interventions and understand their impact on physical activity.

In many studies, activity tracker data are often highly aggregated to analyse the participants' learning by determining whether they have significantly increased their physical activity, met their physical activity goals, or adhered to physical activity behaviour guidelines. This information is clearly valuable to understand changes in physical activity during health interventions. However, the highly detailed data from these trackers may contain latent insights that become diluted within such high aggregation level. Examples of such insights are physical activity distribution and shape throughout the day or week, the creation/change/deletion of physical activity patterns, or how long behaviours are sustained over time.

Traditional intelligent teaching systems provide adaptable content and personalised automated feedback using the interaction data via user interfaces (e.g. mouse, keyboard, tactile interface). In physical domains, such as health and physical education, sports and any physical human movement learning, these data sources involve various sensor data that provide information about this physical learning (what the student is doing) and possibly map it to their knowledge of the topic [68]. The fields of Educational Data Mining, Artificial Intelligence in Education

(AIED) and Learning Analytics and Knowledge are interested in integrating the physical dimension of learning to support learning by developing data mining techniques for extracting learning behaviours. This concerns, for instance, learning martial arts, dance, and the use of clinical equipment [69]. However, techniques for extracting meaningful information for learning from sensor data focus on learning specific movements relative to a known standard (or expert) movement pattern. In the case of an open learning domain, such as improving physical activity behaviours, where there are no prescribed way or expert way to which the learner's physical activity can be compared, new methods are needed.

Detailed information about learners' physical activity could be used for various applications to support and improve health education. For instance, extracting learners' daily behaviour changes could help to scrutinise the impact of intervention programmes by analysing which programme contents were more effective, and also to tailor such programmes by automatically providing just-in-time relevant physical activity objectives and personalised feedback.

The aim of this thesis is to contribute to the development of data mining techniques and models for extracting physical activity behaviours using the students' physical activity data captured by wearable accelerometer sensors in order to support their learning and assess health education intervention programmes.

1.1 Research Questions

This thesis addresses two key research questions (RQ):

RQ1 Can data mining techniques and models help to assess the effectiveness of physical activity interventions using activity sensor data?

RQ2 Can data mining techniques help to detect behaviour changes and generate personalised feedback to support learning?

These research questions are addressed in four thesis objectives. Each objective is tackled in a chapter of this thesis. Figure 1.1 shows how chapters 2 to 5 contribute to address these research questions and specific objectives.

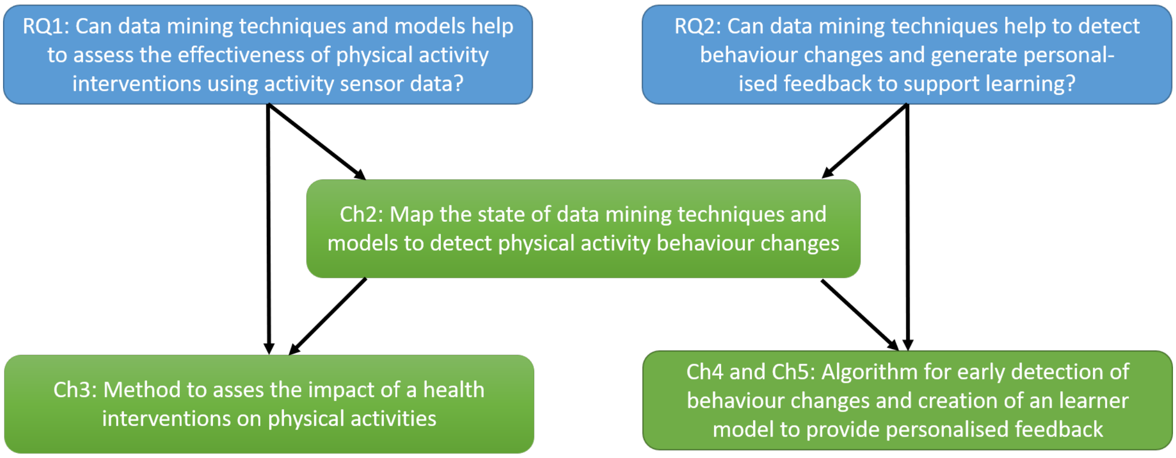

Figure 1.1. Thesis overview

Thesis Objective 1: To map the state of data mining techniques and models to detect physical activity behaviour changes. The detection of behaviour changes is valuable information for health education and behaviour change health interventions. This information allows a detailed analysis of how participants modify their physical activity. This helps to understand how and when participants learn during interventions, thus allowing the objective assessment of the intervention effectiveness. Its importance is reflected by the wide and growing range of artificial intelligence and data mining methods to detect and extract behaviour changes in recent years. However, these techniques and models are diverse and come from different research areas, making it difficult to determine what techniques/approaches are used, what they provide, and what their limitations are.

The contribution of Chapter 2 is a systematic literature review of the current state of data mining techniques used to analyse physical activity sensor data in health education in order to detect behaviour changes. It also discusses common challenges and opportunities to guide future work.

Thesis Objective 2: To propose a data mining method to understand at a finer level the impact of health interventions on physical activities. The impact of an education programme is usually evaluated by comparing aggregated data statistics from participants' accelerometer sensors, thus generating an overall view of the intervention effectiveness.

However, we hypothesise that highly precise sensor data may contain latent information on important and essential aspects of physical activity that could support a much more complete analysis of the interventions. For instance, the duration, intensity, distribution and frequency of certain physical activity behaviours could offer a much more comprehensive examination of how these behaviours were impacted by the intervention.

The contribution of Chapter 3 is a data mining method to assess in more detail the impact of a health intervention. The method extracts the participants' physical activity characteristics from their accelerometer data, creating daily and hourly features. These are then clustered and analysed to understand how participants' physical activity behaviour changes, providing a detailed view of the intervention impact on the participants' behaviour. This method emphasises the behaviour change types and timings, thus complementing the standard and widely used statistical evaluations.

Thesis Objective 3: To develop a data mining technique for the early detection of behaviour changes using periodically streamed physical activity data. The aim of health education interventions is to allow people to learn healthy physical activity behaviours and sustain them. Therefore, it is very important to identify during such interventions whether new physical activity behaviours emerge and whether they are maintained over time. Their early detection can allow reinforcing new healthy behaviour changes and discouraging unhealthy behaviours before they become established. This may increase the intervention adherence and effectiveness. The automated periodic detection of behaviour changes during an intervention would provide key information about the participants' learning that could allow developing artificial intelligence health education systems to support their learning as the intervention unfolds.

The contribution of Chapter 4 is U-BEHAVED (**U**nsupervised **BE**haviour and **HA**bits **V**ery **E**arly **D**etector), an unsupervised data mining algorithm that periodically detects significant changes in physical activity behaviour and whether they become habitual. Using the participants' step data streamed from wearable sensors it compares current physical activity behaviours to recent previous ones using rolling time windows to detect any substantial change and whether it is repeated over time. The automated detection of physical activity behaviour changes offers the opportunity of modelling these behaviours to create personalisation in health education, paving the way to generate and trigger detailed individual feedback to participants and to adapt the health education programmes in function of each participant's specific behaviours.

Thesis Objective 4: To build a learner model to create automated personalised feedback in order to support students' learning on the basis of their physical activity behaviour changes. The widespread use of physical activity tracking sensors in health education opens the possibility of extracting physical and motor ability domain interactions and then incorporating them into AIED systems to drive personalised feedback to support the participants' learning. There are three challenges involved. First, the participants' physical activity behaviour changes must be detected continuously using trackers data. Second, the participants' relevant behaviour change attributes must be integrated into a learner model. Third, timely personalised feedback must be generated and provided.

The contribution of Chapter 5 is the integration of the physical domain into a learner model to provide automated personalised feedback based on behaviours. Physical activity data from the participants' trackers are retrieved and analysed continuously by a set of unsupervised window-based algorithms that feed relevant behavioural-related learner attributes. Based on the participant's learner attribute values, timely and personalised feedback can be generated as a narrative to support learning.

1.2 Thesis Contributions

The contributions of this thesis are as follows:

- Chapter 2: A novel systematic review that maps the current state of data mining techniques used to capture physical activity sensor data in health education in order to detect behaviour changes.

- Chapter 3: A novel method to thoroughly assess the impact of a health education intervention.
- Chapter 4: A novel unsupervised algorithm for the early detection of significant physical activity behaviour changes.
- Chapter 5: A novel method that integrates physical activity behaviour changes into a learner model to generate automated personalised feedback.

CHAPTER 2

Mining Sensor Data to Assess Changes in Physical Activity Behaviours in Health Interventions: A Systematic Review

---

Sensors are increasingly used in health interventions to unobtrusively and continuously capture participants' physical activity in free-living conditions. The rich granularity of sensor data offers great potential for analysing patterns and changes in physical activity behaviours. The use of specialised machine learning and data mining techniques to detect, extract, and analyse these patterns has increased, helping to better understand how participants' physical activity evolves. This chapter aims to identify and present the various data mining techniques employed to analyse changes in physical activity behaviours from sensors-derived data in health education and health promotion intervention studies. This chapter addressed two main research questions: (1) What are the current techniques used for mining physical activity sensor data to detect behaviour changes in health education or health promotion contexts? (2) What are the challenges and opportunities in mining physical activity sensor data for detecting physical activity behaviour changes? We performed a systematic review using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, resulting in 19 articles included for analysis. The main findings are that the main sensors used were accelerometers, sometimes in combination with another sensor (37%). Data were collected over a period ranging from 4 days to 1 year (median: 10 weeks) from a cohort size ranging between 10 and 11,615 (median: 74). Data pre-processing was mainly carried out using proprietary software, generally resulting in step counts and time spent in physical activity aggregated predominantly at the daily or minute level. The main features used as input for the data mining models were descriptive statistics of the pre-processed data. The most common data mining methods were classifiers, clusters, and decision-making algorithms and these focused on personalisation (58%) and analysis of physical activity behaviours (42%). Mining sensor data offers great opportunities to analyse physical activity behaviour changes, build models to better detect and interpret behaviour changes, and allow for personalised feedback and support for participants, especially where larger sample sizes and longer recording times are available. Exploring different data aggregation levels can help detect subtle and sustained behaviour changes. However, the literature suggests that there is still work remaining to improve the transparency, explicitness, and standardisation of the data preprocessing and mining processes to establish best practices and make the detection methods easier to understand, scrutinise, and reproduce. This chapter is based on work published in [30].

2.1 Introduction

Wearable sensors are increasingly employed in health interventions because of their ability to track participants' physical activity in an unobtrusive, continuous, and precise manner under free-living conditions [84]. In the context of health promotion, sensor data are commonly used to objectively assess interventions by monitoring physical activity changes and progress toward compliance with public health physical activity guidelines [16]. The rich data captured by activity sensors contain information about the participants' physical activity, potentially unlocking valuable insights into physical activity behaviours patterns [93], for instance by identifying when and how often they make a physical activity behaviour change and whether the behaviour change is significant and must be encouraged. These insights can help to advance the understanding of how interventions affect physical activity behaviours and how behaviours change, thereby scaffolding the design of future interventions, and enhancing their outcomes, efficacy, and adherence.

In the last decade, a growing number of artificial intelligence and data mining models and techniques have been developed to detect and extract these latent physical activity patterns beyond the typical summaries of pre- and post-intervention daily steps or time spent in various physical activity levels. In this systematic review, we aimed to describe the data mining models and techniques currently used to detect physical activity with a focus on behaviour changes. We discuss their value, identify gaps or challenges, and highlight opportunities. The following research questions (RQs) guided this review:

- RQ1 What are the current techniques used for mining physical activity sensor data to detect behaviour changes in health education or health promotion contexts?
  - RQ1.1 What are the types of sensors used and what data are collected?
  - RQ1.2 How are data preprocessed?
  - RQ1.3 What features are used to detect behaviour changes?
  - RQ1.4 What are the data mining models and techniques used to detect behaviour changes?
  - RQ1.5 What are the interpretation of data mining models used for?
- RQ2 What are the challenges and opportunities in mining physical activity sensor data for detecting physical activity behaviour changes?

The RQ1 subquestions were established following the reasoning and order of the process of knowledge discovery in databases [43]. Figure 2.1 summarises this process and maps each step with the relevant RQ1 subquestion.

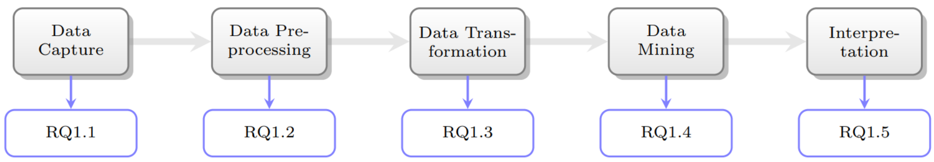

Figure 2.1. Knowledge discovery in database steps (in grey) and RQ1 subquestions (in blue)

2.2 Methods

2.2.1 Design

For this systematic review, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [64] and used the Rayyan QCRI web application [81] to manage the review process. We identified studies by searching the Association for

Computing Machinery (ACM), IEEE Xplore, ProQuest, Scopus, Web of Science, Education Resources Information Center (ERIC), and Springer digital libraries. We also searched Google Scholar to identify grey literature and extracted the first 100 results. For this scholarly reference search, we used the following query: (education OR promotion OR "behaviour change") AND ("data mining" OR "machine learning" OR "artificial intelligence") AND (sensor OR accelerometer OR tracker OR wearable) AND "physical activity" AND health. All extracted scholarly references had been added to the database at the latest on the search day (May 28, 2021).

When building the query, we initially explicitly included the term "Health Education", however, we found scarce references using this specific term, therefore we adapted the term by using the keywords 'education', 'promotion' or "behaviour change". Similarly, a wide range of terms are used to name devices that record physical activity, so we attempt to capture them by using the terms 'sensors', 'accelerometer', 'tracker' and 'wearable'.

We included all references that met the following inclusion criteria:

- Full-length articles
- Peer-reviewed articles in journals or conference papers
- Articles that used data mining techniques for data from physical activity wearable sensors
- Articles that included physical activity data
- Articles on applied health education/promotion or on behaviour change scenarios
- Articles that used well-known data mining techniques, such as classification, regression, clustering, association and sequence algorithms, as well as specific algorithms to model physical activity data

We excluded all articles that met at least one of the following exclusion criteria:

- Use of analytics without data mining
- Studies on animals (e.g. accelerometers on dogs)
- Self-quantification without a health education or health motivation component

- Dissertations and theses, due to lack of a peer review process
- Systematic reviews, reviews and meta-analyses
- Healthcare applications without a health education or motivation for behaviour change component
- Specific movement detection (abnormal gait, falls)
- Aid for sport training (e.g. maintaining heart rate, postures, specific movements)

2.2.2 Search Outcome

The number of references extracted from each electronic database is summarised in Table 2.1.

TABLE 2.1. Number of references extracted from each database

Database	Query Result (N)
ACM	584
IEEE Xplore	12
ProQuest	1678
Scopus	44
Web of Science	16
ERIC	2
Springer	1952
Google Scholar	100

Following the PRISMA methodology, we retrieved 4,388 references from the sources listed in Table 2.1. We then removed 415 duplicates, leaving 3,973 unique references that were screened by reading their titles and abstracts. Using the inclusion/exclusion criteria, we excluded 3,688 references and selected 285 publications. After full-text reading, we excluded 266 references: 33 on activity recognition, 5 on data mining, 24 on systems, 31 on rehabilitation, 39 not on behaviour changes, 54 without data mining, 51 not on health education/promotion, 13 not on physical activity, and 16 reviews. At the end of the selection process (summarised in Figure 2.2), we retained 19 references for this systematic review.

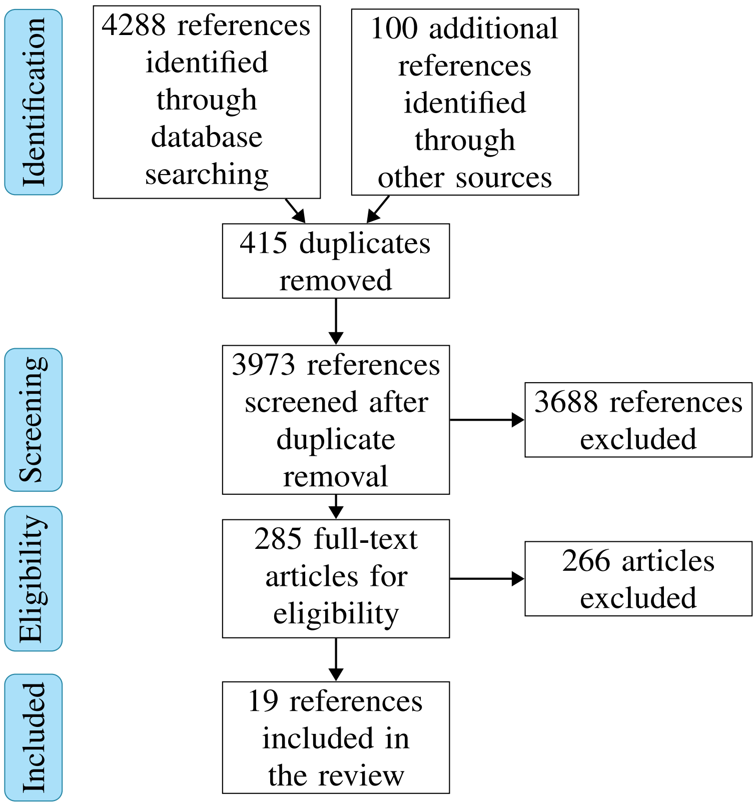

Figure 2.2. Study inclusion flowchart according to the PRISMA methodology

2.3 Results

2.3.1 Overview

The 19 included articles were published between 2013 and 2021. Their number per year increased from one in 2013 to two in 2017 and up to five in 2018. Subsequently, the number of publications decreased to a mean of three per year.

The selected articles were published in conferences and journals focused on five different themes (Table 2.2): Medical and Public Health, Medical and Health Informatics, Human Computer Interactions, Physical Human Behaviour, and Engineering and Science. The three most popular themes were Medical and Health Informatics, Human Computer Interactions, and Engineering and Science (15/19, 79%). Among the included articles, four were published

TABLE 2.2. Conferences proceedings and journals in which the included articles were published (N=19)

Theme	Conference proceedings and journals	Reference
Medical and Public Health	BMJ Open	[4]
	Public Health Nutrition	[62]
Medical and Health Informatics	JMIR mHealth and uHealth	[114] [86] [47]
	JMIR Public Health and Surveillance	[46]
	Journal of Biomedical Informatics	[102]
Human Computer Interactions	Proceedings of the ACM on Human-Computer Interaction	[116]
	User Modeling and User-Adapted Interaction	[48]
	Journal of Ambient Intelligence and Humanized Computing	[12]
	Multimedia Tools and Applications	[5]
	Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization	[98]
Physical Human Behaviour	Journal of Behavioral Medicine	[45]
	Journal of Electromyography and Kinesiology	[53]
Engineering and Science	Applied Sciences	[18]
	Sensors	[36]
	Springer Proceedings in Complexity	[74]
	IEEE Access	[34]
	International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems	[75]

Journal of Medical Internet Research (JMIR) publications: three in JMIR mHealth and uHealth, and one in JMIR Public Health and Surveillance.

2.3.2 Sensor Types and Data Capture

The characteristics of the sensors (e.g. number and type) used to capture physical activity behaviours and of the collected raw data are summarised in Table 2.3.

The length of data recordings varied between 4 days and 1 year, with a median of 70 days. Recording lasted ≤ 7 days in two studies, between 3 and 5 weeks in six studies, between 10 and 16 weeks in eight studies, and ≥ 6 months in three studies.

The number of participants varied between 10 and 11,615, with < 30 in five studies, between 30 and 299 in ten studies, and ≥ 300 participants in four studies.

TABLE 2.3. Number of sensors, device type and model used, and raw data generated

Number of sensors	Sensor type	Device and Model	Raw data	Reference
Single	Accelerometer	ActiGraph GT1M uniaxial	Uniaxial Accelerometry	[62]
		GENEActiv triaxial accelerometer	Gravity-subtracted Signal Vector Magnitudes (SVMgs) per second	[34] [47]
		Generic device from the mobile phone	Acceleration (sample rate not specified)	[4] [114]
		Triaxial accelerometer (HJA-350IT, Active Style Pro, Omron Healthcare Co, Ltd)	Triaxial Acceleration (6Hz)	[46]
		Fitbit	Triaxial Acceleration (sample rate not specified)	[116]
		Fitbit Flex	Triaxial Acceleration (sample rate not specified)	[36]
		Fitbit Charge HR and Fitbit Flex	Triaxial Acceleration (sample rate not specified)	[102]
		Fitbit One	Triaxial Acceleration (sample rate not specified)	[75]
		Not specified	Not specified	[74]
		Smartphone and Actigraph (GT3X model)	Triaxial Acceleration (sample rate not specified)	[98]
Combination	Accelerometer and heart rate monitor	Mix of devices and models	Accelerometry, heart rate monitor, physical activity information and user information (sample rate not specified)	[18]
	Accelerometer, Global Positioning System (GPS), self-log physical activity and food	Smartphone	Smartphone accelerometry, GPS data, physical activity and food logs with sample rate specified	[86]
	Activity tracker, smart scale and smartphone (what they ate and drank in the Fitbit app)	Fitbit Flex 2 activity tracker, Yunmai smart scale, smartphone	Accelerometry, weight and food logs (sample rate not specified)	[45]
	Accelerometer, gyroscope and magnetic compass	ProMove-3D (developed by Inertia Technology)	Accelerometry (sample rate not specified)	[53]
	Accelerometer and GPS	Smartphone	Accelerometry and GPS (sample rate not specified)	[12]
	Triaxial accelerometer, heart rate monitor, GPS, 3-axis gyroscope, digital compass, altimeter, light sensor	Samsung Gear Fit and Fitbit Surge	Accelerometry, heart rate data, GPS, 3-axis gyroscopes, digital compass, altimeter, light sensor (sample rate not specified)	[5]
	Accelerometer, heart rate and smart scale	Nokia, models not specified	Accelerometer, heart rate data and smart scale (sample rate not specified)	[48]

All included studies used accelerometer sensors. We could categorise these devices into three groups:

- Commercial wrist-worn wearable accelerometers that are consumer-grade devices with a sample rate between 30 Hz and 60 Hz, such as Fitbits [45] [102] [36] [75] [116], Samsung Gear [5], and Nokia [48].
- Smartphone accelerometers with a sample rate usually set to 50 Hz and up to 100 Hz, in which data were collected via an application installed in the smartphone [86] [12] [98] [4] [114].
- Scientifically validated wearable accelerometers with a sample rate up to 100 Hz, such as ActiGraph [62], GENEActiv [34] [47], and other devices developed for healthcare [46] [53].

In 7 out of the 19 (37%) selected studies, accelerometers were used with other sensors, such as global positioning system (GPS) tracking [12] [86] [5], compass position tracking [53] [5], heart rate trackers [18] [5], smart scales [48] [45].

The recorded raw data varied in function of the sensor characteristics, including sampling frequency, accuracy, and axis number. Moreover, other sensor features, such as battery duration and storage capacity, affected the recording length. For instance, a long battery life and high storage capacity enable longer recording without interruptions.

Table 2.4 summarises the number of participants and data recording duration for the included studies.

TABLE 2.4. Length of data recording and number of participants among the included studies

Length of recording category	Length of recording	Number of participants	Reference
1 to 7 days	4 days	1714	[62]
	7 days	215 women	[46]
1 to 5 weeks	3 weeks	17	[86]
	3 weeks	48	[116]
	1 month	14	[5]
	4 weeks	24 adolescents	[47]
	4 weeks	74 children	[98]
	5 weeks	87 children	[34]
6 to 20 weeks	10 weeks	11	[98]
	10 weeks	64	[114]
	3 months	10	[53]
	12 weeks	48	[36]
	12 weeks	108	[75]
	3 months	269	[18]
	12 weeks	2472	[74]
	16 weeks	52	[45]
21 weeks to 1 year	6 months	276	[4]
	6 months	500	[12]
	1 year	11615	[48]

2.3.3 Data Pre-processing

Raw data extracted from sensors need to be transformed into variables that will contribute to generate the input features for data mining models to detect physical activity behaviour changes. Table 2.5 provides a summary of the initial transformation and the resulting pre-processed data.

The pre-processing of the raw data from sensors was carried out in two ways. The first approach was to use proprietary programs to transform sensors' data directly into the resulting pre-processed data without specifying whether there was an initial pre-processing such as that used to generate steps, metabolic equivalents (METs), calories, heart rate, or exercise characteristics (type, duration, distance, or frequency). The second approach was to produce intermediate data that were then transformed in the resulting pre-processed data using a custom pre-processing tool. For instance, to generate Physical Activity Levels (PALs), raw data were first transformed into MET, activity classes, or signal vector magnitudes.

TABLE 2.5. Summary of data pre-processing variables.

Resulting/initial pre-processing	Reference
Steps	
Unknown (proprietary program)	[114], [4], [45], [116], [75], [48], [36], [102], [5]
Metabolic equivalents	
Unknown (proprietary program)	[18]
Calories	
Unknown (proprietary program)	[45], [5], [86]
Exercise characteristics*	
Unknown (proprietary program)	[18], [5], [86]
Sleeping time	
Unknown (proprietary program)	[48], [5]
Weight	
Unknown (proprietary program)	[48], [45]
Heart rate	
Unknown (proprietary program)	[18], [5]
Physical activity levels	
Signal vector magnitudes	[34], [47]
Physical activity counts	[62]
Metabolic equivalents	[46]
Activity Classes	[98]
Not specified	[74]
Integrals of the moduli of acceleration signals	
Integrals of the moduli of acceleration signals	[53]
Actual activity level**	
Not specified	[12]

*Type, duration, distance, frequency;**Definition of activity level was not specified.

The resulting pre-processed data were mainly activity characteristics (steps count, PAL, integrals of the moduli of acceleration, activity types, duration, distance travelled, and frequency) and energy expenditure (MET and calories). Step count from smartphones and commercial wrist-worn devices was the most frequent, followed by PAL from research-grade devices.

The resulting pre-processed data were aggregated at different time levels (Table 2.6). Day and minutes were the most frequent time levels of aggregation. Generally, PAL and MET were aggregated per minute. Calories and step counts were calculated per day.

TABLE 2.6. Aggregation level of the resulting pre-processed data

Month	Week	Day	Hour	Minute	Seconds	Not specified	Reference
X	X	X	X				[5]
		X					[114]
		X					[4]
		X					[116]
		X					[75]
		X					[45]
		X					[48]
X				X			[18]
			X				[36]
				X			[62]
				X			[46]
				X			[102]
				X			[98]
					X		[34]
					X		[47]
						X	[74]
						X	[53]
						X	[86]
						X	[12]

2.3.4 Features Used to Detect and Extract Behaviour Changes

The features of the data mining models were mostly generated from the sensors' pre-processed data and, in some cases, from other sources (non-sensor data). Table 2.7 provides the features categorised with respect to the function of their source: accelerometers, other sensors, and non-sensor devices.

TABLE 2.7. Features used for data mining to detect behaviour changes.

From accelerometers	From other sensors	From non-sensor devices	Reference
Number of minutes of activity in the last day, cumulative number of minutes of activity this week, fraction of activity goal, fraction versus expected activity goal at this point in the week	Number of days since each feedback message was sent	Age, gender, language, 8-item Patient Health Questionnaire (depression) score	[4]
Not specified	Not specified	Not specified	[53]
Monthly mean metabolic equivalent of task, effective exercise time, type, frequency	Monthly mean exercise and resting heart rate	Gender, height, weight, age	[18]
Days where physical activity goal is met	Sum of days with self-monitored weight, days with self-monitored eating, days where calorie goal is met, weight loss in pounds	Number of days in the intervention period	[45]
Consecutive daily segments of steps, consecutive daily segments of sleep	-	-	[48]
Actual activity level	-	Desired activity level, intention (attitude, subjective norms, perceived behavioural control), habit and 16 demographic features (e.g. age, gender, marital status)	[12]
Hour of the workday, number of steps for that hour, number of steps in the past hour, total number of steps till that hour, mean number of steps of workdays	-	-	[36]
Physical activity frequency and calories	-	-	[86]
Daily steps and goal	-	-	[114]
Total and mean hourly, daily, weekly and monthly sleep duration, sleep calories, exercise duration, exercise distance, exercise calories, step count, step distance, step calories, body mass index, and basal metabolic rate	-	Height (cm), weight (kg), age, gender	[5]
Hourly and daily frequency and mean time spent in moderate to vigorous physical activity bouts of at least 3, 10 and 30 seconds, and in sedentary bouts of at least 60, 120, 300 seconds	-	-	[34]
Total daily time spent in light/moderate/vigorous physical activity, total daily number of steps, and a binary goal achievement feature	-	-	[47]
Mean metabolic equivalent of tasks per minute, mean moderate to vigorous physical activity per minute	-	-	[46]
24-hour mean physical activity count on weekdays and 24-hour mean physical activity count on weekends	-	-	[62]
Steps, physical activity level and bouts count, mean, percentages, ratios and SD. Circadian rhythm time series statistics and texture features from an image processing technique	-	-	[102]
Impact of online community (sharing my PAL level with peers), target physical activity level and goal achievement	-	-	[74]
Physical activity level per minute	-	-	[98]
Daily steps	-	Psychological questionnaire scores for self-efficacy, barriers, social norm, long-term goals, intentions, satisfaction, outcome expectations	[75]
Daily steps	Motivation to exercise (Likert scale)	Iowa-Netherlands Comparison Orientation Measure-23 (INCOM-23) for social comparison (psychometrics)	[116]

Most of the included articles used descriptive statistics to present the pre-processed data as features, for instance total number of steps per day [114] [5] [47] [75] [116], mean number of steps per day [5], or physical activity count per hour [62]. Other studies created windows or segments of time to calculate physical activity characteristics, including segments of steps or sleep [48] and physical activity bouts [34] [102]. Other articles used the pre-processed data to calculate the participants' step achievements, such as whether they reached their step goal

[114] [45] [47] [74]. [102] used more complex features, such as the ratio between the most active and least active period, or the circadian rhythm strength.

In addition to the features derived from sensors, others were created from measurements carried out during the intervention by scientists, such as the number of days that a person participated in the intervention [45] and anthropometric [4] [18] or psychological [12] [75] [116] characteristics. Data were collected through surveys/questionnaires or interviews with participants.

2.3.5 Data Mining

Table 2.8 summarises the data mining methods and specific algorithms used by the selected articles.

Clustering was the most used method, particularly the K-means algorithm. Indeed, in health interventions, the physical activity performed by each participant varies in duration, form, and intensity. Therefore, an algorithm that clusters physical activity behaviours is required to analyse them. The unsupervised K-means algorithm is suitable for this task. Indeed, due to its simplicity and ease of use, this is one of the most popular options for data mining [110]. Decision-making algorithms and classifiers were the second most used methods. Both rely on supervised algorithms that use physical activity characteristics as a method for predicting when and/or what information must be delivered to individual participants for increasing their physical activity. Other algorithms were also tested to extract physical activity behaviours, such as social cognitive and contagion models, physical activity windows permutations, and recommendation algorithms.

TABLE 2.8. Data Mining Methods and Algorithms

Data mining method	Algorithm	Reference
Classifiers	K-Nearest-Neighbour and Support Vector Machine	[53]
	Random Forest	[36]
	Random Forest and Weighted Score	[98]
	Shallow Neural Networks	[12]
Clustering techniques	K-means	[34], [47], [62], [46]
	Agglomerative	[18]
	Partitioning Around Medoids and Reinforcement Learning	[48]
Decision-Making Algorithms	Multi-Armed Bandit	[86]
	Multi-Armed Bandit Upper Confidence Bound	[45]
	Reinforcement Learning Multi-Armed Bandit	[4]
	Behavioural Analytics Algorithm	[114]
	MAB	[116]
Social Cognitive Model	Social cognitive model for predicting exercise behaviour change	[75]
Social Contagion Model	Social contagion model combined with a linear model	[74]
Physical Activity Change Detection	Small window permutation-based change detection in activity routine	[102]
Recommendation	Genetic algorithms and pareto optimality	[5]

Classifiers. [53] used a k-nearest-neighbour model and a support vector machine to determine whether a specific time of the day was suitable for sending a motivational message to optimize adherence to the intervention. [36] used a tree and tree-based ensemble algorithm classifiers to predict whether users will achieve their daily physical activity goal. On the basis of this prediction, a personalised physical activity coaching program was proposed. [98] developed gamified personalized feedback using a score model depending on the physical activity change detected from accelerometer data. [12] predicted the likelihood that the physical activity level of a given patient was too low. They also predicted which patients were at higher risk of not adhering to the prescribed therapy to optimise their physical activity.

Clustering Techniques. [62] grouped participants in two clusters on the basis of their step counts (one more active than the other), and analysed them to better understand these physical activity patterns. [34] used a clustering-based approach for a more insightful analysis of the participants' physical activity behaviour and of the nature of the physical activity behaviour changes, if present. [47] clustered physical activity levels and daily step goal achievement to assess the adherence to a health programme. [46] identified physical activity clusters to analyse and compare sociodemographic features and cardiometabolic risks among participants belonging to these clusters. [18] clustered the participants' physical activity and then established a system to adapt the exercise programme for the next week as a function of the individual physical activity behaviour change. [48] clustered the participants' physical activity to generate groups of habits recommended by a system to the participants with the objective of changing their physical activity to obtain weight loss effects.

Decision-Making Algorithms. [86] generated personalised suggestions in which users were asked to continue, avoid, or make small changes to their existing physical activity behaviours in order to help them reach their physical activity goals. In [45] developed an algorithm, that could personalise and optimise the physical activity levels during the intervention as a function of the amount of physical activity performed. [4] generated personalised messages for participants in the intervention to increase their physical activity and consequently, the intervention effectiveness. [114] adapted the step goal settings of the intervention depending on the physical activity behaviour change. [116] personalised social comparison among participants to motivate them toward improving their physical activity behaviour.

Social Cognitive Model. [75] developed a model that simulates changes in physical activity levels over two to twelve weeks to optimise the participants' health outcome.

Social Contagion Model. [74] used a social contagion model to explain the physical activity level dynamics in a community.

Physical Activity Windows Permutations. [102] proposed a window-based algorithm to detect changes in segments of users' physical activity behaviour to motivate progress toward their goals

Recommendation Algorithms. [5] used Genetic Algorithms and Pareto Optimality to compare the participants' and peer community's data to help participants interpret physical activity data and to generate personal lifestyle improvement recommendations.

2.3.6 Interpretation of the Data Mining Models

The resulting data mining models detecting PA behaviour changes were used for several purposes, as summarized in Table 2.9 and below.

TABLE 2.9. Main uses of the resulting data mining models

Main use	Reference
Personalised feedback	[4], [53], [98], [48], [12], [86]
Personalised programme	[18], [45], [36], [114]
Support for self-reflection	[5]
Cohort analysis of the intervention impact on physical activity	[34], [47], [46], [62], [102]
Analysis of the social component effects on physical activity	[74], [75], [116]

Personalised Feedback. The physical activity behaviour changes extracted from participants' data were used to promote physical activity by creating and sending personalised messages that reported the behaviours and gave suggestions for achieving the previously established physical activity goals. For instance, [4] built a system that detects the participants' physical activity behaviour changes and generates personalised daily text messages with custom timing, frequency, and feedback about their step count/goal and motivational content. [53] built a system that chooses the best suitable time to send a message with personalised intention, content, and representation. [98] created an application with gamified feedback where different avatars are awarded based on the participant's daily physical activity behaviour. [48] suggested personalised physical activity patterns based on the participants' physical activity patterns. [12] detected the participants' physical activity behaviour while commuting and suggested how to increase it. [86] generated personalised simple physical activity suggestions (continue, avoid, or make small changes).

Personalised Programmes. The physical activity intervention programme and objectives are adapted to each participant's needs. For instance, [18] created a guided exercise prescription system that adapts as the participants' physical activity behaviour changes. Similarly, [45] changed the participant's exercise intensity suggestion depending on their physical activity behaviour achievements. On the basis of each participant's step count progress, [36] suggested new daily step objectives. [114] used push notifications to deliver daily step goals.

Support for Self-Reflection. Algorithms can help participants to interpret their physical activity behaviour changes. For example, [5] used an algorithm to assist in the interpretation of the participant's physical activity data by comparing them with those of the peer community and to generate personalised recommendations to achieve their daily goals.

Cohort Analysis of the Intervention Impact on Physical Activity. These algorithms detect physical activity behaviour changes in participants that allow analysing the intervention impact. For example, [46] determined physical activity patterns in women throughout the day that could help to develop more personalised interventions and guidelines. [34] analysed the changes in physical activity behaviour (bouts and frequency) during an intervention. [47] tracked the participants' adherence to the physical activity international recommendations during an intervention. [62] identified physical activity patterns associated with specific subgroups of people who participated in an intervention. [102] analysed the participants' physical activity changes during an intervention by comparing multiple time windows.

Analysis of the Social Component Effect on Physical Activity. These algorithms analyse the psychosocial influences on the participants' physical activity. For example, [74] analysed the physical activity dynamics in a community using a social contagion model. [75] analysed the physical activity dynamics in a networked community using social cognitive theories, and [116] personalised social comparison during an intervention to increase the participants' physical activity.

The main uses can be classified in two groups. The first group, composed of around 11 out of the 19 (58%) selected studies, aimed to generate personalised feedbacks/physical activity programmes to scaffold and support physical activity behaviour changes among participants. Indeed, researchers seem inclined to generate greater personalisation because it increases the intervention efficiency, effectiveness, enjoyment and reliability [19]. The second group, composed of 8 out of the 19 (42%) selected studies, sought to analyse the impact of interventions on the participants' physical activity. Specifically, these studies analysed the intervention impact on physical activity at the cohort level to assess health education interventions, and analyse participants' physical activity to show them their behaviours and help to understand them. The main objective of both groups was to explore how physical activity behaviour patterns relate to the intervention effectiveness, which can add new evidence on how to create more effective interventions [108].

2.4 Discussion

We found 19 articles about data mining models and techniques to detect physical activity behaviour changes in health education or promotion studies, and their number has progressively increased over time. In this section, we discuss the principal findings, identify opportunities and challenges for future research directions, and present the limitations of this systematic review. The discussion is structured using the research questions as a guide.

2.4.1 Principal Findings: Opportunities and Challenges

Sensor types and data capture. All selected studies used accelerometer sensors to capture physical activity behaviours. While 7 out of the 19 (37%) studies utilised accelerometers exclusively, the rest employed them with other sensors. Non-accelerometer sensors capture additional information that may be relevant to physical activity (such as work/school schedule, itineraries, and sleep patterns [99]) and could yield auxiliary features for the data mining models. For instance, GPS sensors provide the number of kilometres and location of physical activity performed.

The median number of participants in the selected studies was 74, and participants were mainly young or middle-aged adults. This low number of participants and the skew toward adults may have generated biased data mining models that can detect and find behaviour changes only in a specific population. Different population groups behave differently and should be studied independently. For instance, physical activity behaviours are different in children and adults [16]. Some of the studies focused on groups with specific physical activity behaviours, such as children [98] [34], adolescents [47] and women [46]. However, some population groups with distinctive physical activity patterns, such as pregnant women [14] and people with health conditions or disabilities [103], may need custom detection models.

In 15 out of the 19 (79%) included studies, data were recorded for less than 3 months. Therefore, the current methods for detecting physical activity behaviour changes have been developed mostly for capturing short-term patterns, making the conclusions valid only for short periods. To detect medium- and long-term physical activity behaviour changes, studies with more extended recording periods are needed, such as the study by [48] based on data collected during 1 year. Moreover, new methods to detect extended (e.g. annual or seasonal) physical activity patterns are required to study how the participants' behaviour and habits change over time. An increase in the participants' number and recording length will lead to new challenges related to big data analysis, such as efficient data management and data mining processing speeds.

Data pre-processing. Many of the selected studies used commercial accelerometers that allow only the retrieval of aggregated pre-processed data using proprietary software (i.e. number of steps per minute), without being transparent on how data were pre-processed (i.e. how steps were calculated from the accelerometry data). This data pre-processing black box makes it impossible to determine the quality of the captured physical activity data and makes the data mining results scientifically irreproducible. Conversely, in studies that used medical-grade accelerometers, the accelerometry data were explained in detail and the pre-processing steps were documented and referenced.

We found a lack of standard procedures for data pre-processing that made it challenging to compare the study results and conclusions. Indeed, if data are not pre-processed correctly, this could cause the transfer of incorrect information to the features and then to the data mining models. This could lead to the creation of inaccurate models, thus limiting the study validity. Data cleaning is a good example of this issue. Indeed, the best procedure to eliminate the non-wearing time remains unclear along with the impact on the accuracy of the resulting. If non-wearing time is poorly removed, features can generate a physical activity underestimation by recognising non-wearing time as sedentary behaviour when it is not. Moreover, if sensor data concerning changes in accelerations while commuting by car or bus are not completely removed, they will be erroneously classified as steps, thereby overestimating physical activity in the model and in the conclusions. Similarly, sedentary activities could be overestimated if sleep time is not correctly removed.

Most of the selected studies aggregated information by day or minute. Although data aggregation is useful when comparing general features of physical activity behaviours, such as daily steps, this procedure may overlook subtle behavioural changes that can be crucial for detecting major physical activity behaviour changes. For instance, if a person who walks every morning decides to change their behaviour and starts to walk at night, the sum of daily steps will be the same, but this new behaviour will not be detected. Conversely, it could be detected if the aggregation level is changed to the hour. To detect these and other subtle behaviour changes, physical activity should be analysed simultaneously at different aggregation levels, and new time frames should be created to match daily habits and behaviours, such as periods of the day (e.g. morning, afternoon) or participants' office hours.

Features used to detect and extract behaviour changes. Most of the pre-processed data were transformed into features that are simple descriptive statistics, such as the total time spent at a specific PAL or the mean number of steps. These features are valuable to detect behaviour changes, but they mainly capture the physical activity intensity and the physical activity presence or absence. Yet, physical activity has more valuable characteristics that vary during physical activity behaviour changes and that can help to detect such behaviour changes, such as the length of physical activity levels bouts or the amount of time spent doing physical activity. These physical activity characteristics can be extracted from current sensor data. For instance, [47] explored different moderate to vigorous physical activity bout lengths, and [102] assessed the circadian rhythm. International physical activity guidelines can serve as inspiration to identify new physical activity features. For instance, according to the World Health Organisation recommendations, adults should perform muscle strengthening activities (involving all major muscle groups) at moderate or higher intensity at least twice per week [16]. This calls for the creation of features that capture the muscle activity type, intensity, and frequency. Moreover, most of the included studies used only physical activity-derived features to detect behaviour changes, and did not consider relevant non-physical activity data associated with physical activity changes, such as the participants' weight and quality of sleep. Some studies captured non-physical activity data, but they did not use them to detect physical activity changes. For instance, [86] used only physical activity-derived data (physical activity frequency and calories burned) to detect behaviour changes, although they also recorded the participants' food intake, thus excluding their caloric intake that is closely related to weight and the amount of physical activity participants are likely perform.

The use of simple descriptive statistics as features and the exclusion of non-physical activity data associated with behaviour changes indicate that sensor data were underexploited and that the features used to detect physical activity behaviour changes are still underdeveloped. Including new physical activity characteristics and new non-physical activity features could help to better understand the nature of physical activity changes and how these features influence physical activity behaviour changes, ultimately increasing the model detection accuracy.

Data mining models and techniques. Most studies used off-the-shelf classifiers, clusters, and decision-making algorithms to detect physical activity behaviour changes. We expected to find tailor-made algorithms because in health education settings, it is important to find specific physical activity patterns in participants of different classes who follow learning modules with different contents, and with different physical activity goals. Moreover, we noticed that most authors did not explain how they chose the algorithms and did not specify the efficiency and accuracy of the models used for detecting physical activity behaviour changes, raising uncertainty about how good they are at this task. This suggests that more efficient and accurate algorithms could be created and calls for more transparency in the algorithm choice process. Therefore, authors should explicitly describe the steps and methodology of new algorithms, and share their source codes to be scrutinised and to compare their detection accuracy. The creation of open accelerometry databases is also needed to enable benchmarking.

Interpretation of the resulting data mining models. The main uses of the data mining models focused on personalisation, support for self-reflection, and analysis of physical activity behaviours. Model interpretation focused on generating personalisation and support for promoting behaviour changes. Personalised feedback and intervention programmes were based mostly on the participants' physical activity data. The inclusion of additional information that may influence behaviour changes (e.g. contexts, schedules, social constraints, motivation, and weather) would allow for better interpretation and use of the detected behaviour changes. Systems could exploit these additional data to improve the feedback delivery time and content, with positive effects on the effectiveness of health education programmes and interventions. For instance, with the current models, a participant could receive an automatised personalised behaviour change message that suggests taking a short walk, although it is snowing outside. This would decrease the likelihood of following the suggestion. However, if the system could be aware of the weather, the participant would receive this suggestion only after the weather conditions have improved, or a different suggestion that is more likely to trigger a behaviour change at that point in time. Moreover, as the models relied mainly on physical activity features to model and interpret the behaviour changes, only the physical dimension of the learning process in health education was incorporated in the models and their interpretation, leaving aside the knowledge dimension of the learning process. Learning Management Systems and Intelligent Tutoring Systems already capture the knowledge dimension. Their integration would help to understand in a comprehensive way how participants learn and would enable the real-time monitoring of how physical activity behaviour changes align with the intervention purpose. This would allow adapting each participant's content and learning objectives in real-time, thereby improving instructions and learning, ultimately increasing the programme or intervention effectiveness.

Most of the included studies generated complex output models that require detailed knowledge of how they were created to interpret the resulting patterns, making them difficult to understand for health scientists and any other scientist not familiar with machine learning. This is a common problem in interdisciplinary teams; however, an effort can be made to create more readable, intuitive, and easy-to-understand algorithms and methods, a goal that exists in related machine learning areas, such as Explainable Artificial Intelligence (XAI) [2].

2.4.2 Limitations

Studies on wearable machine learning devices to detect changes in physical activity in health education have only started to be published in the last decade. As research is advancing, keywords are changing, and new terms are created. Although we used a wide range of keywords in our query to include sensors, physical activity and health education, we may have left some keywords out, and thus we may have missed some references. This may have also affected the initial reference screening process by title and abstract. We minimized this issue by testing several queries before starting our systematic review until we found the one we ultimately used. Another possible limitation in our search is that we might have omitted references listed only in other peer-reviewed databases (we searched only the most popular databases in engineering and computer science), such as medical databases (i.e. PubMed). We mitigated this risk by including grey literature in our systematic review (see Methods section).

Regarding the research sub-questions and the review structure, we created research sub-questions in line with the usual data mining process steps, but we certainly left some topics unaddressed. For instance, we did not address ethics, privacy, and security issues, or how data are filtered during pre-processing (e.g. sleeping time or sensor non-use). Although these are common sub-steps during the data mining process and including them would have made this systematic review more comprehensive, we preferred to limit this review only to the critical steps.

2.4.3 Conclusions

In the last ten years, different methods have been developed to detect behaviour changes in health education or health promotion contexts. These methods have been tested in small populations, are based on short data-recording periods, and rely mainly on accelerometry data. Incorporating information that is complementary to the participants' physical activity data would allow for creating more precise detection models, better interpreting these models, and understanding how participants learn and what triggers new behaviours. Exploring other data aggregation levels, in addition to days and minutes, could help to detect more subtle and long-term behaviour changes. Fully describing the data pre-processing methods and the efficiency and accuracy of the behaviour change detection models would help to better understand, scrutinise, and compare studies. Detection models were mainly used to generate personalised feedback and to provide support for promoting or maintaining behaviour changes, but did not integrate the knowledge dimension of the learning process. Adding the knowledge dimension and creating easier-to-understand models could facilitate the interpretation of participants' behaviour changes in a more comprehensive way, opening the way toward better and deeper analyses and personalisation.

2.5 Chapter Contributions and Relationship to the Thesis Research Questions

This chapter contributes and is related to both research questions:

RQ1 Can data mining techniques and models help to assess the effectiveness of physical activity interventions using activity sensor data?

RQ2 Can data mining techniques help to detect behaviour changes and generate personalised feedback to support learning?

(1) We comprehensively analysed the current literature on data mining algorithms used in health education to detect behaviour changes using physical activity sensor data. This review analysed current techniques by dividing them into the sub-processes of data capture, data pre-processing, data transformation, data mining, and data interpretation. The review helps to address both RQ because it summarises the available approaches, how they work, and how they relate to each other. This helps to determine what data mining techniques can be used for physical activity assessment (RQ1) and for supporting learning (RQ2). This information can also help health and computer scientists to better understand how the current data mining

techniques work and how the detected behavioural changes can be used to improve health education.

(2) We provided a discussion on the common challenges and opportunities based on the review findings. We identified literature gaps on data mining methods to detect behaviour changes using physical activity data in health education interventions. It addresses both RQ because it helps us to guide and justify our work for physical activity assessment (RQ1) and for supporting learning (RQ2). This information may also guide the future work of health and computer scientists interested in creating the next generation of data mining methods.

CHAPTER 3

A Clustering Approach for Modelling and Analysing Changes in Physical Activity Behaviours from Accelerometers

---

To evaluate the impact of health interventions for promoting physical activity, researchers typically conduct pre- and post-assessments using accelerometers. Aggregated metrics, such as daily counts, daily steps and time spent at various intensity levels, are commonly used. On the other hand, very few studies have exploited the richness of data that are often collected with a very fine granularity. In this chapter, we investigated the benefit of a deeper analysis of wrist accelerometry data to understand physical activity behaviours throughout the day, and how they may change over time. To analyse physical activity behaviour changes, we propose a methodology that extracts physical activity bouts characterised by their activity levels and duration, and then uses them as features to cluster the participants' daily and hourly behaviours. Then, we compared these clusters to assess changes following an intervention for promoting physical activity in children. We demonstrated that this approach provides a more insightful analysis of the physical activity behaviours because it highlights the nature and timing of behaviour changes, when present. We illustrated this methodology using data from research-grade activity trackers (GENEActiv) and explained the insights discovered in the context of an intervention to educate school children about healthy behaviours. This chapter is based on work published in [33] and [34].

3.1 Introduction and Background

Evidence shows that children are not sufficiently physically active worldwide. A recent global health study [22] revealed that only 9% of boys and 2% of girls meet the World

Health Organisation daily recommendations of 60-min moderate to vigorous physical activity (MVPA). This is worrying because physical activity plays an important protecting role against many metabolic diseases. To tackle this issue, a growing number of health interventions have been deployed, especially in schools. However, only 56% of school-based interventions have led to a significant, but very small physical activity behaviour change [25]. This highlights the need to improve the design of such interventions and to better understand their impact on physical activity. Indeed, the capacity to precisely assess the effectiveness of interventions, health promotion or health education programmes is crucial from a public health perspective and also from an economic perspective

It has been shown that the use of technology in health promotion or education interventions can improve health behaviours and provide insights into how enhance their effectiveness [60], particularly when physical activity is concerned. Due to the increasing availability of wearable technologies, researchers now routinely use activity trackers with triaxial accelerometers to asses physical activity unobtrusively, objectively, and continuously [84] [47]. These devices can capture the daily activity in real life, thus replacing or complementing self-reported data that are often inaccurate and coarse, especially in studies on children where self-reported data and/or parental reports are not very accurate [58]. Activity trackers provide a better measure of the achievement of public health recommendations, such as performing 12,000 daily steps [3] and 60 minutes of MVPA per day [52]. In addition, the richness and objectiveness of accelerometer sensor data open the way to deeper physical activity analyses, in terms of occurrence and distribution, relative to the currently established health guidelines. Researchers can address more effectively and accurately research questions, such as whether the physical activity bout continuity and duration (e.g. frequent, short bouts vs less frequent, longer bouts) play an important role in achieving health guidelines. They can also gain insights into how physical activity is distributed during the day, for instance in relation to the child's typical schedule. In addition, this can help to build evidence for determining the optimal way to accumulate physical activity and MVPA [79] in the longer term and for informing policies and recommendations.

The most frequent use of accelerometers is to quantify daily physical activity, particularly in response to physical activity-promoting interventions. However, they capture much more detailed information to be used for more specialised analyses, such as activity recognition [89] and changes in every day physical activity [102]. Extracting behaviour changes in response to learning, and investigating the evolution of knowledge or skills from user data is not new. Specialised data science fields, such as Educational Data Mining [8] and Learning Analytics [101] have been developing techniques to exploit various types of learner data. These techniques have been recently expanded to use physical movement data captured by accelerometers. There is now an emerging interest in using sensors to capture fine grained physical behaviours in education contexts: for example, for learning kinaesthetic skills in martial arts [96], dance [26], for properly using clinical equipment [69], or for learning precise hand movements in engineering activities [109]. Generally, these education domains are focused on learning specific movements where a known gold standard (or expert) movement pattern exists, making supervised learning techniques appropriate.

On the other hand, assessing the impact of health education interventions is open-ended, and thus data must be analysed using unsupervised learning methods [63] to determine, at a meaningful granularity level, the physical activity shape and patterns and to detect physical activity behaviour changes after the intervention, especially in relation to what has been learned. Here, we tackle the problem of modelling and comparing physical activity behaviours between accelerometer datasets that were captured before and after the learning and behavioural change intervention, with the objective of understanding the extent and nature of its impact. For instance, an important question is whether any additional MVPA observed after an intervention occurs in longer bouts of activity (which may suggest more sustained intentional activity), or scattered in very small amounts throughout the day (which is more likely to be just incidental). Another question is to determine when new MVPA bouts occur (e.g. during school time, when the children are in their school environment or after school when they are with family and friends).

The contribution of this chapter is a clustering-based approach for a more insightful analysis of the participants' physical activity behaviour and of the nature of physical activity behaviour changes, if present. The two research questions (RQ) were:

- (RQ1) Can we extract daily physical activity behaviour changes from accelerometer data, at the appropriate granularity that allows a meaningful assessment of the health intervention impact relative to the daily recommended targets?
- (RQ2) What technique can be used to precisely analyse a physical activity behaviour change in function of the time of the day when it occurred?

This chapter builds on our earlier work [33] to understand the impact of a learning intervention by analysing physical activity behaviour changes at the daily level. We developed a more comprehensive approach for a deeper analysis of daily and also hourly behaviours that we then tested using a large dataset.

Chapter 3 is structured as follows. Section 3.2 discusses existing work on sensor data mining to extract health-related behaviours. Section 3.3 introduces the health education programme, the data collected, and the rationale. Section 3.4 describes our methodology and its application to extract daily and hourly behaviours. Section 3.5 describes the resulting physical activity behaviour clusters. Section 3.6 analyses the behaviour changes using these behaviour clusters to answer the two research questions (RQ1 and RQ2). Section 3.7 concludes the chapter. Lastly, Section 3.8 outlines the chapter contributions and relationship to the thesis research questions.

3.2 Related Work

3.2.1 Non-Data Mining Approaches

Physical activity changes after health interventions are traditionally evaluated using statistical methods and data aggregated over the test period. For example, [91] reported the daily mean amounts of MVPA (min/day), light physical activity (min/day), accelerometer counts

(counts/min), and step counts (steps/day). [107] used accelerometers, questionnaires and a series of fitness tests (Eurofit) to evaluate physical activity and fitness. They compared pre- and post-intervention data by computing the percentage of time spent at various physical activity intensities per day and their mean values, standard deviations, and main effects (F-values). Such analyses provide a big picture of changes; however, these measures are highly aggregated and do not provide information about physical activity patterns and whether/how they may have changed.

3.2.2 Supervised Techniques

Supervised mining of activity tracker data has been essentially used to assess the movement quality or to identify activities; however, this analysis type does not give insight into physical activity patterns and characteristics. Nevertheless, relevant accelerometer features and data mining techniques can be found in these works. For example, [24] assessed the quality of physical activity movements using smartphone accelerometer data features, such as mean acceleration, standard deviation, and interval between acceleration peaks, and a decision trees to classify them. For activity recognition, [61] used triaxial acceleration data from smartphones to generate acceleration data features (mean, standard deviation, difference, and average). They exploited these features to build classifiers for six pre-labelled activities (walking, jogging, walking upstairs, walking downstairs, sitting, standing) using decision trees (J48), logistic regression and multilayer neural network models that showed up to 90% of accuracy. [9] used the mean, energy, entropy, and correlation features extracted from five biaxial accelerometers as input to recognise twenty different pre-labelled activities with $>$80% of accuracy. [70] exploited features extracted from accelerometers and other sensors (light, temperature sensor and microphone data) to classify six different activities (sitting, standing, walking, ascending stairs, descending stairs, and running) using a C4.5 Decision Tree classifier with $>$80% of accuracy.

Health monitoring assistance and support systems make use of various sensors combined into a wireless body area network. [113] proposed a decision support system using a Support Vector Machine analysis for classification and regression analysis to extract features from accelerometers, global positioning system (GPS), heart rate, blood pressure and weight data. [57] used a range of data sources (electrocardiogram, electromyography, electroencephalography, blood pressure, tilt, breathing, movement and a "smart sock") combined with external data (e.g. weather forecast data) to classify the users' states and activities in the domain of computer-assisted physical rehabilitation. Such techniques have been used also for educational purposes. For instance, [45] created a decision-making algorithm to optimise a reinforcement learning system for a weight loss intervention programme using pre-classified data from Fitbit sensors and the intervention responses. [98] developed a gamified intervention for motivating children to increase their daily physical activity in which smartphone accelerometer data were mined with Support Vector Machine and Random Forests techniques to (i) classify the children's physical activity (sitting, standing, walking, jogging, walking upstairs, walking downstairs, and intense physical activity) and (ii) calculate a score based on the amount of time spent in these activities.

3.2.3 Unsupervised Techniques

Unsupervised data mining has been used to detect changes and to analyse physical activity. For the detection of physical activity behaviour changes, time window-based techniques have been developed. For example [102] adapted the Permutation-based Change Detection in Activity Routine (PCAR) algorithm [23] to create the Physical Activity Change Detection (PACD) algorithm that uses the distance between physical activity time series windows to detect a significant physical activity change. Using synthetic and physical activity data (Fitbit data), they found that the PACD algorithm detected more changes compared with other window-based change detection algorithms (RuLSIF, texture-based, PCAR, and a virtual classifier). The authors' aims was to detect physical activity changes from accelerometer data, but they focused on changes in daily activity-based routines to identify and quantify changes in these routines, rather than comparing two different datasets. Furthermore, the physical activity data were represented by a single feature (sum of step time series), whereas other important and relevant characteristics, such as the physical activity intensity that is important to meaningfully assess the impact of a health intervention, were not considered.

For physical activity analysis, [62] used a K-means clustering algorithm to group people according to their physical activity levels extracted from uniaxial hip accelerometer data in an observational study on family health, happiness and harmony. Once transformed into MVPA and light physical activity, they separated datapoints into weekdays and weekend, and averaged them by hour. Then, they used the obtained 48-hour counts of MVPA and light physical activity as features for clustering and then extracting the participants' weekday and weekend behaviours. The authors identified two clusters: one active and one much less active. Then, they analysed demographic, lifestyle, physical activity level, and health characteristics in function of the cluster. [37] clustered data from three accelerometers and a heart rate monitor using Hierarchical Clustering Analysis and K-means algorithms with time and frequency domain as features to provide patients and caregivers with a more accurate overview of their physical activity. The studies by [62] and [37] are the most related to our work, but their aim was to highlight general temporal physical activity patterns in the population without addressing the needs of an intervention, particularly the assessment of the pre- and post- intervention behaviour changes concerning physical activity and sedentary time.

Recently, [47] analysed data from wrist activity trackers worn by children during a school-based health education intervention (iEngage) by creating daily behaviour clusters. These clusters were built using the following features: sum of minutes spent on different physical activity levels (light, moderate, and vigorous), number of steps per day, and weekly consistency of the children's adherence to international guidelines as the intervention unfolded. The resulting clusters allowed a deeper analysis of physical activity progress during the educational intervention and showed that children improved their physical activity behaviours over time, particularly those who were the least active at the intervention start.

In conclusion, data mining techniques have been used successfully to identify activities, characterise daily patterns, and assess behavioural changes during an intervention. However, methods for analysing data at a fine granularity are required in order to 1) characterise fine patterns of engagement in physical activity, and 2) identify detailed behaviour changes using physical activity characteristics extracted from accelerometer data that are consistent with the previous studies mentioned. This is important because it may help to detect early changes, predict trajectories, and help to better tailor interventions.

3.3 Context and Background

3.3.1 The iEngage Study

The accelerometry data were collected before and after a 5-week school-based digital health education intervention (iEngage) [111]. The experimental group (N=61 children) followed the iEngage learning sessions over 5 weeks, but not the control group (N=26 children). The iEngage programme aims to develop health knowledge and skills to allow 9-11 year-old school children to acquire healthy behaviours, with a focus on reaching the recommended levels of daily physical activity (at least 60 minutes per day of MVPA [55]) and decreasing sedentary time.

3.3.2 Accelerometer Datasets

We carried out pre- and post-intervention tests by measuring unobtrusively and continuously the children's physical activity using research-grade activity trackers (GENEActiv [1]) for five consecutive school days. This allowed generating two five-day datasets per child. We placed the GENEActiv accelerometers on the wrist of their non-preferred hand to capture acceleration in three axes (x,y,z) with a sample frequency of 60 Hz. Additional information about the study and the accelerometer data can be found in [111].

To ensure that all daily records had the same length, we excluded records with missing data. The result of this filtering step is summarised in Table 3.1.

TABLE 3.1. Dataset after filtering.

Group		**Total** Hours	**Total** Days	**N** 5-days	**N** Complete	**N** Pairs
Experimental	Pre	7080	295	59	54	45
	Post	6360	265	53	50	
Control	Pre	3120	130	26	26	24
	Post	3000	125	25	25	

Total Hours = Number of hours recorded. Total Days = Number of days with records. N 5-days = Number of children with five-day datasets. N Complete = Number of children with five-day datasets after filtering. N Pairs = Number of children with both pre- and post-intervention datasets, after filtering.

3.3.3 Rationale

A precise analysis of the total time spent in physical activity every day in the control and experimental groups from the first iEngage study [111] showed that before the intervention, both groups spent similar time at each physical activity intensity level (p-value = 0.63, 0.62, 0.76, and 0.29 for sedentary, light, moderate and vigorous intensity, respectively). Conversely, after the intervention, physical activity time was significantly increased in the experimental group, especially for moderate and vigorous intensity physical activity (p-value = 0.003 and 0.017, respectively), whereas it did not change in the control group. This indicates that the overall desired effect (at least in the short term) was reached in this population. Nevertheless, we wanted to know how this activity was distributed throughout the day, and how it changed over time (i.e. RQ1 and RQ2). To answer these questions, we must analyse fine grained information about the salient structure of physical activity, particularly physical activity bout intensity, duration (length) and frequency. Data on bout duration are needed for the meaningful analysis of the collected physical activity data because very short bouts will contain much noise and longer bouts may dilute MVPA time if mixed with lighter physical activity. Furthermore, physical activity sessions that last at least 5 minutes and that include at least 80% of MVPA bouts might be required for effective positive health outcomes [67]. The next important aspect was to allow the meaningful comparison of physical activity behaviours. For instance, a period of mostly sedentary time followed by a period of mostly MVPA is comparable to a period of mostly MVPA followed by sedentary time. Conversely, a period with MVPA scattered among sedentary moments is different from a period with MVPA scattered within light activity. This led us to model behaviours over a certain period of time to eliminate meaningless differences while maintaining the overall shape of the activity during that interval. Once these behaviours are modelled, they can be compared.

3.4 Methodology for Extracting Physical Activity Behaviours

The methodology is summarised in Figure 3.1. We pre-processed both datasets (pre- and post-intervention) and transformed them into sequences of physical activity intensity level vectors. Then, we compressed these vectors into physical activity bouts per intensity level and used their characteristics as features to cluster the data and identify physical activity behaviour types during the relevant period (day or hour). We will describe these steps in detail in the next sub-sections. The methodology was carried out by scripts written in R [85]. The scripts are accessible on the Open Science Framework website and are divided into two packs; one has the scripts for the first four steps [32], while the other has the scripts for the final two steps [28].

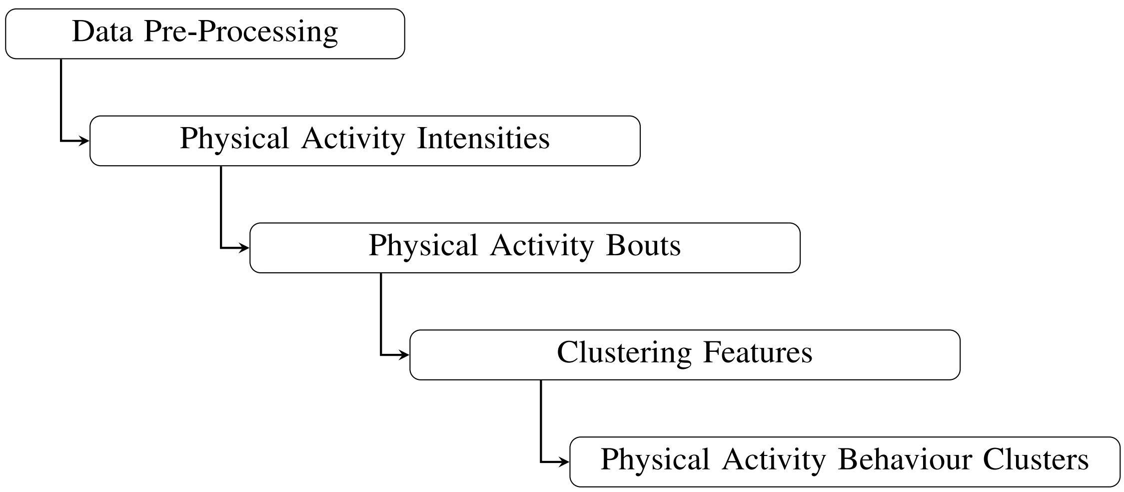

Figure 3.1. Methodology for extracting physical activity behaviours

3.4.1 Data Pre-Processing

We pre-processed both datasets using a specific R library for handling GENEActiv trackers data [42]. We converted the accelerometer binary files to accelerometer data frames, and applied the data cleaning explained in sub-section 3.3.2. Then, we translated the three-dimensional 60 Hz acceleration data frame into gravity-subtracted Signal Vector Magnitudes (SVMg) [40] within a 1-second epoch (see equation 3.1). Each data was time-stamped.

$$SVMgs = \sum_{i=1}^{60} |\sqrt{x_i^2 + y_i^2 + z_i^2} - g| \tag{3.1}$$

We chose the 1-second epoch because children engage in moderate to vigorous activities mostly in very short bouts. For example, it has been reported that 80% and 93% of moderate and vigorous physical activity are performed in bouts lasting less than 9s and 4.7s, respectively [7] [10]. This indicates that children's behaviour, particularly physical activity patterns, could have been incorrectly assessed and that smaller epochs are needed to accurately detect activity.

3.4.2 From SVMgs to Physical Activity Intensity Levels

We categorised each 1-sec SVMgs into a physical activity intensity level according to scientifically validated cut-offs for assessing physical activity intensity in children [83] (Table 3.2).

TABLE 3.2. SVMgs Cut-Off Levels

Physical Activity Intensity Levels	**SVMgs Cut-Off**
Sedentary	[0, 4.5[
Light	[4.5, 16.5[
Moderate	[16.5, 42[
Vigorous	≥ 42

Using these cut-offs, we coded each second with a letter, as follows: S for sedentary time, L for light physical activity, M for moderate physical activity, and V for vigorous physical activity (Figure 3.2). For example, a string of 5 seconds may be coded as LLLMVV, which can be read as 3 seconds of light activity, followed by 1 second of moderate activity, and 2 seconds of vigorous activity.

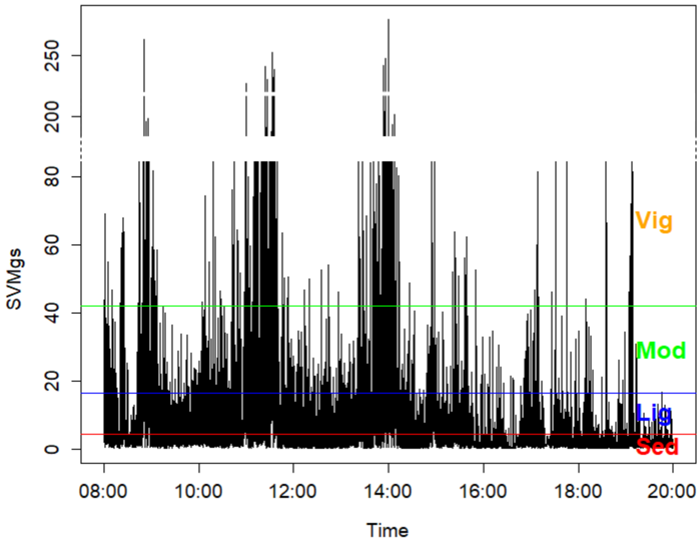

Figure 3.2. SVMgs time series of one child during one day. The red, blue and green horizontal lines represent the cut-offs from sedentary to light, from light to moderate, and from moderate to vigorous physical activity, respectively. The figure is truncated between 80 and 200 SVMgs for easier reading.

3.4.3 Physical Activity Bouts

As mentioned earlier, we analysed the data using the notion of physical activity bouts with a focus on MVPA and sedentary time bouts during daytime. First, we need to introduce some definitions.

- A *bout* is a continuous physical activity episode at a specific intensity level range.
- The *length of a bout* is the number of seconds spent in that bout.
- The *bout frequency* is the number of occurrences of all bouts of a certain length during a day.

As the daily recommendations are expressed in terms of MVPA (thus combining moderate and vigorous physical activity), we merged M and V into the "MVPA" category. For instance, a sequence of 11 seconds in M, 8 seconds in V, and 12 seconds in M, preceded and followed by L, would generate one MVPA bout of 31 seconds (see another example in Figure 3.3).

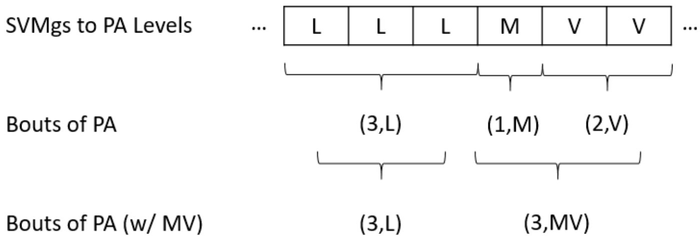

Figure 3.3. From physical activity (PA) levels to PA bouts (and merging of Moderate and Vigorous levels)

Before using physical activity bouts as features for clustering, we rapidly analysed their distribution by computing the total time spent in MVPA, done in bouts of at least x number of seconds. Equation 3.2 shows the reverse cumulative sequence, where t is the bout length, and b is the number of seconds spent in bouts of at least t length. For t=1, this is equivalent to the total number of seconds spent in 1-second MVPA bouts. For t=2, this is the total number of seconds spent in $\geq$2-second bouts (therefore, excluding the time spent in 1-second bouts), and so on.

$$Bouts\,Cum\,Sum_t = \sum_{i=t}^{n} b_i \tag{3.2}$$

Figure 3.4 shows a sample of the result of these calculations, where every line indicates the mean daily cumulative MVPA bout length for a child. After 10 seconds, the curves start to flatten because the bout length increases.

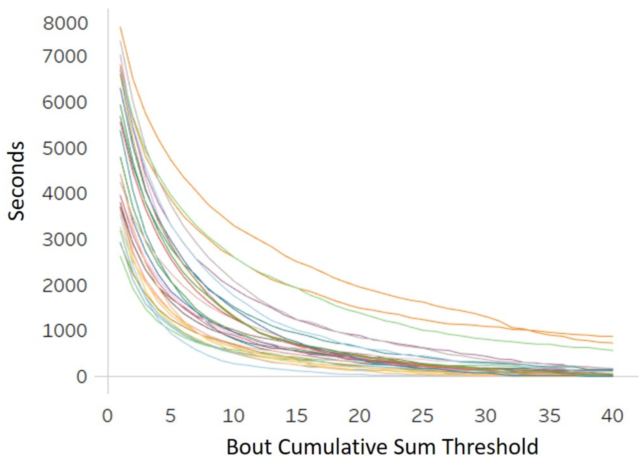

Figure 3.4. Reverse Cumulative Bout Lengths

This figure highlights the presence of much potential noise in the 1-2-second bouts, and an exponentially decreasing function that flattens after 30 seconds for MVPA. We used the same method for sedentary bouts and found that the curves flattened after 300 seconds (not shown).

We compared the pre- and post-intervention cumulative series using a paired T-test and found that in the experimental group, overall, MVPA bout length (p value=6.883e-10) and frequency (p value=0.008) were increased, whereas sedentary bout length (p value=2.2e-16) and frequency (p value=2.2e-16) were decreased. Therefore, such analysis brings new information to characterise how an intervention influences behaviours.

3.4.4 Clustering Features

We selected various bout length cut-off values on the basis of literature data [97] and the results of the previously described exploratory analysis. Specifically, we found that meaningful MVPA can be detected from GENEActiv data starting from 3 seconds, because any shorter activity is likely to be agitation rather than physical activity. Therefore, for MVPA and sedentary bouts we selected lengths of at least 3, 10, and 30 seconds, and of at least 60, 120, and 300 seconds, respectively. Then, for each length and each physical activity intensity level, we calculated two indicators: the bout frequency, and the accumulated time spent (in seconds). Table 3.3 lists the resulting twelve clustering features.

TABLE 3.3. Clustering Features

Feature Number	Measure
1	Total MVPA time $\geq$ 3 seconds in length
2	Total MVPA time $\geq$ 10 seconds in length
3	Total MVPA time $\geq$ 30 second in length
4	Number of MVPA bouts $\geq$ 3 seconds in length
5	Number of MVPA bouts $\geq$ 10 seconds in length
6	Number of MVPA bouts $\geq$ 30 seconds in length
7	Total sedentary time bouts $\geq$ 60 seconds in length
8	Total sedentary time bouts $\geq$ 120 seconds in length
9	Total sedentary time bouts $\geq$ 300 seconds in length
10	Number of sedentary bouts $\geq$ 60 seconds in length
11	Number of sedentary bouts $\geq$ 120 seconds in length
12	Number of sedentary bouts $\geq$ 300 seconds in length

3.4.5 Behaviour Clustering

After feature normalisation and principal component analysis dimensionality reduction, we used a k-means algorithm. We selected this algorithm as first step because it is a widely used algorithm, relatively easy to implement and suitable for our analysis, given the reduced number of dimensions. We also tested DBSCAN, a non-spherical clustering technique [41], but by projecting the features we observed it generated a less cohesive cluster model than k-means. We acknowledge that other clustering methods could have been tried [92], especially more complex, state-of-the-art algorithms, such as ensembles [82], weighting [77], or alternative clustering models [50], because the ratio between the amounts of MVPA and of sedentary behaviours may require different distance metrics.

The clusters were generated by combining the children's behavioural features pre- and post-intervention, implying that children can be present in up to two clusters: one before and one after the intervention. This method simplifies the analysis of behaviour changes by matching and comparing each child's pre- and post-intervention behaviour clusters (see section 3.6). In contrast, creating one set of clusters using pre-intervention behaviours and another set of clusters using post-intervention behaviours would have resulted in clusters with different centroids and descriptions, making matching each child clusters impractical and increasing the complexity of the cluster comparison and analysis.

The resulting clusters should capture the pre- and post-intervention physical activity behaviours of all children, and help to understand the behaviour changes in the whole cohort.

3.5 Physical Activity Behaviours: Results

Our methodology characterises physical activity behaviours at a coarse level, but allows capturing essential elements to understand how physical activity is distributed throughout the day. The idea is to identify the physical activity distribution types that are present in the cohort data, and to distinguish these distributions. We investigated two approaches:

– The Daily approach (subsection 3.5.1) to identify daily activity patterns. Indeed, two days (for two different children, or for the same child) can show the same total MVPA quantity (e.g. 40 minutes), but one will contain much sedentary time and long MVPA sessions, and the other more scattered MVPA sessions with less sedentary time (hence, more light activity).
– The Hourly approach (subsection 3.5.2) to detect hourly patterns for a more fine grained analysis of when, during the day, activity occurs and changes

3.5.1 Physical Activity Daily Behaviour Clustering

For the daily physical activity behaviour analysis, we only considered the awake time during full school days. Therefore, we excluded the days when the bracelets were installed and removed, and filtered out sleeping times. This resulted in 104 physical activity records of up to ∼45 hours of daytime activity (three 15-hour days from 7:00am to 9:59pm) (Table 3.4).

TABLE 3.4. Daily records after filtering

Group	**Days**			**Total**
	3	**2**	**1**	
Pre	44	10	0	54
Post	46	4	0	50

We computed the daily behaviour vectors using the 12 daily features shown in Table 3.3 and averaged the daily values over the three days. The resulting 104 (54 pre- and 50 post-intervention) vectors characterise the average daily physical activity, pre- or post- intervention, of a child.

We standardised the 12 daily features and then used principal components analysis to reduce the number of features and maximise their variance. We retained the first three principal components based on three criteria [56]: cumulative percentage of total variation $\leq$80%, principal component variance size (eigen value $>$1.0), and scree plot (see Figure 3.5).

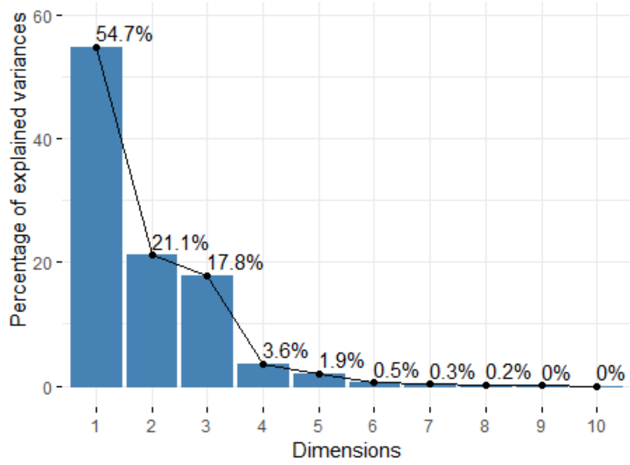

Figure 3.5. Scree plot of the principal components

Then, we used the K-means unsupervised algorithm [66] with k=6. Indeed, we found that adding another cluster did not improve the total within-cluster sum of squares.

Behaviour clustering identified six daily physical activity behaviour types (Table 3.5). The daily cluster 1 (D1) contained days with the smallest number of MVPA bouts and the highest number of sedentary bouts. D6 contained the highest number of MVPA bouts and the second lowest number of sedentary bouts. The other clusters presented intermediate characteristics. It is clear that the most active daily clusters (D4, D5, D6) were also the least sedentary. In all clusters, the time spent in and the number of MVPA bouts $\geq$30 seconds were drastically shorter and fewer than the time spent in and the number of shorter MPVA bouts. This can be explained by the fact that most of the spontaneous activity in 9-11 year-old children is shorter than 30 seconds, with high recovering capacity [87]. This is also consistent with previous studies showing that during one or several repeated high-intensity exercise bouts, young children fatigue less than adults [88].

TABLE 3.5. Daily Cluster Centroids

Sedentary Intensities	**Measure**	**1 (N=14)**	**2 (N=22)**	**3 (N=11)**	**4 (N=17)**	**5 (N=25)**	**6 (N=15)**
$\geq$ 60 seconds							
	Total time (min)	354.4	247.5	455.6	263.4	160.1	184.8
	Bout number	129.1	89.7	57.2	104.7	66	68.7
$\geq$ 120 seconds							
	Total time (min)	247.1	165.9	406.4	171.7	96.7	120.9
	Bout number	51.2	29.4	21	37.4	19	21.9
$\geq$ 300 seconds							
	Total time (min)	132.9	100.3	369.2	86	53.2	71.6
	Bout number	12.5	6.6	8.1	8.1	3.4	5

MVPA Intensities	**Measure**	**1 (N=14)**	**2 (N=22)**	**3 (N=11)**	**4 (N=17)**	**5 (N=25)**	**6 (N=15)**
$\geq$ 3 seconds							
	Total time (min)	29.8	35.3	37.5	54.6	68	92.4
	Bout number	251.9	370.1	351.2	448.2	655.1	729.7
$\geq$ 10 seconds							
	Total time (min)	14.1	11.1	14.9	26.7	25.4	47
	Bout number	38.3	39.6	49.4	77.6	89.2	136.8
$\geq$ 30 seconds							
	Total time (min)	6	2.1	4	9.9	5.5	17.1
	Bout number	6.7	3	5.2	11	7	20

On the basis of these observations, we ordered the clusters in increasing level of MVPA amounts, from the lowest (D1) to the highest activity daily cluster (D6), and characterised them using 3-second bouts (see Table 3.6).

TABLE 3.6. Daily Cluster Description.

Daily Cluster	**Description**
D1	Less active cluster (half of the recommended MVPA amount) and frequent sedentary bouts.
D2	Not very active cluster (approximately half of the recommended MVPA amount) combined with an average amount of sedentary time.
D3	Low MVPA (approximately half of recommended MVPA, but in slightly longer bouts) and highest sedentary time, including in longest bouts.
D4+	Borderline active cluster (centroid 5 min below the recommended MVPA levels, but with longer and more frequent bouts than the previous clusters) combined with frequent, short sedentary time bouts.
D5*	Active cluster (meets the recommended MVPA amount), frequent number of short MVPA bouts, and lowest sedentary times and frequencies.
D6*	Active cluster, with the highest amount of MVPA, but slightly longer sedentary times and bouts than D5.

* Clusters meeting the daily MVPA recommendation.
+ Clusters almost meeting the daily MVPA recommendation.

3.5.2 Physical Activity Hourly Behaviour Clustering

For the hourly physical activity behaviour analysis, we used the whole cleaned datasets. This included all nights and the days when the bracelets were installed and removed, resulting in 10,496 hours of data for the 104 physical activity records (Table 3.7).

TABLE 3.7. Hourly records after filter

Group	Total children	N Hours
Pre	54	5386
Post	50	5110

We computed the hourly behaviour vectors using the 12 daily features from Table 3. The resulting 10,496 vectors (up to 121 hours x 104 participants pre- and post-intervention) characterise a child's physical activity during that specific hour.

After feature normalisation, we retained the first three principal components (Figure 3.6).

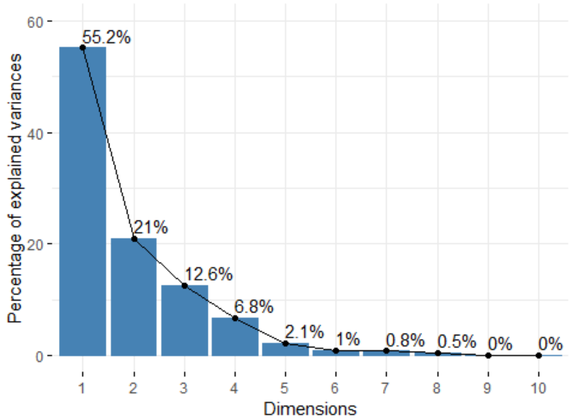

Figure 3.6. Scree plot of the principal components

Then we used the K-means unsupervised algorithm with k=4. We determined the number of clusters using the total within-cluster sum of square method, as before.

Calculation of the centroids of the hourly clusters (Table 3.8) indicated that in the hourly cluster 1 (H1), children spent the whole hour on long sedentary (SED) bouts with negligible MVPA. Conversely, in H4, children spent ∼15 minutes doing high bursts of MVPA, with hardly any sedentary time. H2 and H3 displayed intermediate features.

TABLE 3.8. Hourly Cluster Centroids

SED Intensities	**Measure**	**1 (N=4,340)**	**2 (N=2,359)**	**3 (N=3,268)**	**4 (N=529)**
$\geq$ 60 seconds					
	Total time (min)	57.6	26.9	5.6	5
	Bout number	4.8	11.4	3.2	2.6
$\geq$ 120 seconds					
	Total time (min)	56.6	17.2	2.1	2.5
	Bout number	4.1	4.4	0.6	0.7
$\geq$ 300 seconds					
	Total time (min)	52.6	6.6	0.6	0.8
	Bout number	2.9	0.8	0.1	0.1

MVPA Intensities	**Measure**	**1 (N=4,340)**	**2 (N=2,359)**	**3 (N=3,268)**	**4 (N=529)**
$\geq$ 3 seconds					
	Total time (min)	0.1	0.9	4.1	15
	Bout number	0.6	10	41.7	95.5
$\geq$ 10 seconds					
	Total time (min)	0	0.2	1.4	9.3
	Bout number	0	0.7	5.1	24.3
$\geq$ 30 seconds					
	Total time (min)	0	0	0.2	4.2
	Bout number	0	0	0.3	4.6

We then ordered the hourly clusters in function of their MVPA amount, from the lowest activity (H1) to the highest activity (H4) hourly cluster (Table 3.9).

TABLE 3.9. Hourly Cluster Description

Hourly Cluster	**Description**
H1	Sedentary cluster/sleep. Long bouts of sedentary time, negligible MVPA. These are mostly sleeping hours.
H2	Not very active cluster. Infrequent brief bouts of MVPA (1-min long) and 26 min spent in sedentary bouts of various length.
H3	Lightly active cluster. Occasional short MVPA bouts combined with infrequent and brief sedentary bouts, implying that most of the time is spent doing light activity.
H4	Active cluster. High amount of MVPA bouts of various length combined with infrequent and short bouts of sedentary time.

3.6 Analysing Physical Activity Behaviour Changes

This novel hourly and daily clustering approach offers a unique opportunity to analyse individual and cohort data and to obtain new insights into how participants (here children) respond to the knowledge and skills they learn about healthy physical activity behaviours. Indeed, using this approach, we can characterise individuals, days, or hours based on physical activity descriptors and also on sedentary behaviour.

The daily and hourly clusters capture salient characteristics of the children's physical activity behaviours, such as intensity, duration of bouts of various intensities, frequency of physical activity engagement and time of the day. This approach brings novel information on physical activity behaviour because previous studies on physical activity, including in children (e.g. [91], [107]), used cumulative and sectorised analyses of physical activity that allow only a global interpretation of the results. On the other hand, with this novel cluster analysis we can identify each salient behaviour characteristic, thus highlighting differences at the cohort and at the individual level, relative to the international recommendations (with the daily clusters) and to the activity distribution throughout the day (with the hourly clusters). In the following sections, we will illustrate how these clusters can be used to address the two initial research questions.

3.6.1 Can We Extract Daily Physical Activity Behaviour Changes from Accelerometer Data, at the Appropriate Granularity that Allows a Meaningful Assessment of the Health Intervention Impact Relative to the Daily Recommended Targets (RQ1)?

The **daily behaviour clusters** (Table 3.6) can help to determine whether and to which extent the experimental population changed behaviours by comparing their pre- and post-intervention clusters. These clusters take into account the fact that behaviour can be described in different ways, according to the physical activity structure. This is highly relevant because the physical activity structure can have different effects on the children's health according to their personal objectives, physical fitness level, and other physical activity-associated environmental factors [38].

In the case of the iEngage programme, the analysis of pre- and post-intervention accelerometry data revealed the nature and importance of the changes in relation to the minimum daily recommendations (Table 3.6). The Daily Cluster movement matrix (Table 3.10) displays the cross-tabulation of the number of children (n=45, see Table 3.1) who changed from one cluster pre-intervention to another cluster post-intervention, allowing to assess each salient characteristic simultaneously as captured by the six daily clusters. It showed that 47% of children already belonged to an active cluster before the intervention and remained in the same cluster or moved to a more active cluster after the programme. Moreover, 13% of children were in a cluster below the recommended MVPA guidelines before the intervention, but moved to a more active cluster afterwards. Conversely, 35% of children were in a cluster that did not meet the recommended MVPA guidelines and remained or moved to a cluster that was slightly less active but also less sedentary. Lastly, 4% of children were in an active cluster and moved to a not active cluster.

TABLE 3.10. Daily cluster movement matrix

		To Daily Cluster					
		1	**2**	**3**	**4**	**5**	**6**
From Daily Cluster	**1**	4	2	0	1	0	0
	2	2	3	2	1	2	0
	3	0	2	1	1	0	1
	4	0	1	0	4	0	1
	5	0	0	1	1	7	2
	6	0	0	0	0	2	4

We can visually observe behaviour changes according to various elements of the physical activity structure. For instance, the movement matrix from before to after the intervention (Table 3.10) shows that behaviour changes can be followed in relation to the MVPA guidelines. The green area shows desirable movements (move to or stay in a cluster that meets the recommendation), whereas the blue and red areas indicate no improvement (stay in the same cluster that does not meet the MVPA recommendation) and negative movements (move from a cluster that meets the recommendation to one that does not), respectively. If we are interested in the sedentary times, a slightly different ordering of the clusters would yield a slightly different matrix (D3 would become the most sedentary cluster). This new method allows the reliable overall interpretation of the intervention impact at the cohort and individual levels, and helps to answer other research questions. As the clusters are ordered in increasing MVPA levels, moving to a higher cluster is generally better, but this is not always the case. For instance, a movement from D2 to D3 might seem good at first glance because children in D3 spend a little more time in longer MVPA than in D2. However, D3 is characterised by longer sedentary bouts. Therefore, clusters may not be always comparable in function of their number, and a partial order may exist. Nevertheless, as D4 (borderline), D5 and D6 meet the MVPA recommendations, we coloured the behaviour changes based on this information.

3.6.2 What Technique Can Be Used to Precisely Analyse a Physical Activity Behaviour Change in Function of the Time of the Day when It Occurred (RQ2)?

For RQ2, the **hourly behaviour clusters** provide more specific insights into the period of the day where such potential behaviour changes occurred.

We examined the distribution of time spent in each of the four hourly clusters throughout the school day by calculating the percentage of children in each hourly cluster. Figure 3.7 displays the proportion of time spent in the four clusters during a typical pre-intervention day.

Most children (90%) spent time in active clusters (H3 and H4) between 8am and 3pm. After 3pm, only $\sim$25% of children remained active. Moreover, we observed the highest percentage of children in the more active cluster (H4) before class start (8-9am) and at the morning and afternoon recess times (10am and 2pm). Only fewer than 5% of children spent time in H4 after school. This shows that schools remain the place where children engage in active behaviours compared with the outside school hours.

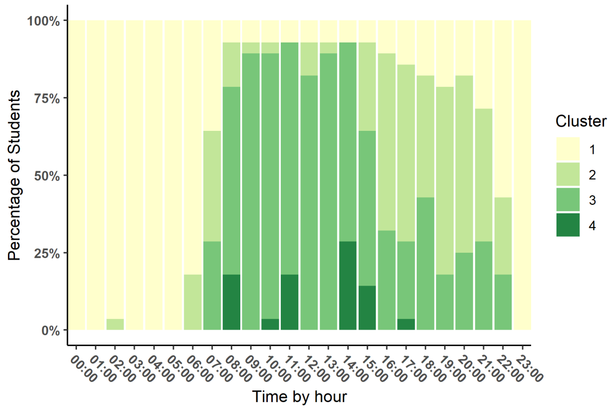

Figure 3.7. Distribution of hourly clusters during one pre-intervention day (24 hours). The x axis shows the hour of the day, and the y axis shows the percentage of children present in each cluster. Cluster numbers are ordered from the least active (H1) to the most active (H4) cluster

Therefore, to address RQ2 and analyse physical activity during specific periods of time, we used the following time windows: Before School (7-9am), During School (9am-3pm), After School (3-7pm), Evening (7-9pm), and Night (9pm-7am). For each time window, we calculated the percentage of time participants spent in the relevant activity clusters, before and after the intervention. Overall, we observed a shift towards higher activity clusters and less time spent in low activity clusters after the intervention (Table 3.11).

Notably, the time spent in the most sedentary clusters (H1 and H2) *Before School*, *After School* and *Evening* decreased, while the time spent in the very active cluster (H4) more than doubled *After School*. The time spent in the activity clusters *During School* was similar, although this time became slightly more active after the intervention. This suggests that participants were more likely to engage in more active behaviour in their own time, outside school hours, especially after school.

TABLE 3.11. Percentage of time spent in each hourly cluster by period of time

Period of Time	**Intervention**	**H1**	**H2**	**H3**	**H4**
Before School					
	Pre	28.2	24.4	40	7.3
	Post	20.5	30.4	38.6	10.5
During School					
	Pre	7	14.4	67.9	10.7
	Post	6.7	11.4	68.8	13.2
After School					
	Pre	10.8	42.2	43	3.9
	Post	7.4	39.8	44.4	8.4
Evening					
	Pre	16.7	52.5	26.6	1.3
	Post	10.2	56.3	31.3	2.3
Night					
	Pre	85.9	11.3	2.8	0.1
	Post	83.7	11.5	4.6	0.2

A more detailed analysis of the *After School* period highlighted a significant (Mann Whitney U test, p-value <0.05) increase by 4-fold of the time spent in H4, evenly distributed between 4pm and 7pm, and a decrease of the time spent in H1 and H2, especially after 5pm (Table 3.12).

The increase of the time spent in the most active cluster (H4) after the intervention indicates a more active behaviour. Then, the cluster centroids (see Table 3.8) allowed assesing in more details the physical activity changes in the cohort. This suggested that these children had a behaviour with few sedentary bouts, particularly almost none over long periods ($\geq$300-second bouts). They also had frequent MVPA bouts, not only short ones ($\geq$3 seconds) that were at least 4 times more frequent in H4 than in the other clusters, but also longer ones ($\geq$30 secs) that were mostly present in H4. This indicated a sustained, purposeful period of MVPA, such as exercise or physical games.

TABLE 3.12. Percentage of time spent on each hourly cluster *After School*

Time	Intervention	H1	H2	H3	H4
3pm					
	Pre	4.4	16.6	68.3	10.7
	Post	3.4	18.9	65.5	12.3
	p-value	0.64	0.48	0.56	0.68
4pm					
	Pre	10	48	40.3	1.9
	Post	8	45.8	38	8.2
	p-value	0.80	0.76	0.67	**0.02**
5pm					
	Pre	13.6	54.8	29.9	1.7
	Post	8.9	47.5	36.9	6.8
	p-value	0.87	0.28	0.20	**0.04**
6pm					
	Pre	15.2	49.6	33.6	1.6
	Post	9.5	47	37.3	6.2
	p-value	0.90	0.62	0.42	**0.02**

The result of this analysis suggests that children became progressively more active after school until dinner time. This is consistent with research showing that the child's environment, including parental/community support, access to facilities, and feeling of security, is a positive factor for engaging in physical activity [94]. Indeed, each child in the iEngage programme was assigned missions involving family, friends, or support persons.

3.7 Conclusion

In this chapter, we presented a methodology to extract salient aspects of children's physical activity behaviour and to determine how these behaviours changed after a health education intervention. We modelled the physical activity into bouts of physical activity intensities, using their length and frequency as features to cluster the daily and hourly physical activity behaviour in relation to international recommendations. We then used these clusters to examine the pre- and post-intervention differences.

This approach is complementary to descriptive statistics (e.g. ANOVA) that measure the overall pre- and post-intervention behaviour differences. It provides an aggregated analysis (via clusters), but also captures important and essential aspects of the activity (length and frequency of physical activity bouts). It also allows a fine grained analysis to understand the intervention impact at the general and individual level, starting from the whole day and then zooming into more specific parts of the day, as illustrated by the *After School* time example.

Such analysis may also help to refine health guidelines concerning physical activity because the current recommendations are mostly based on the health outcomes of exercise intervention studies. This novel physical activity analysis approach allows experts to extract a different type of information that considers the structure of free-living activities. This could pave the way for adopting new methodologies, particularly when physical fitness and cardiovascular or metabolic health markers are available.

Our methodology can easily be adapted to other time aggregation types, besides day and hour, and can also include a wider range of relevant features, such as sleeping. Cluster movements (Table 3.10) can be interpreted relative to different guidelines (e.g. minimal recommended daily minutes spent on vigorous physical activity or maximal amounts and lengths of sedentary behaviour bouts). This flexibility is crucial because the guidelines for physical activity and sedentary behaviours have changed over the years due to new findings on the links between specific physical activity behaviours and health. Future work could also determine whether other clustering algorithms and distance metrics (e.g. [50], [77], [82]) could be used to improve the physical activity cluster quality, and consequently physical activity pattern discovery.

Our methodology was subsequently adapted and used for the iBounce study [51], a digital health intervention to educate and engage childhood cancer survivors in physical activity. In this work, the cluster analysis of participants' physical activity intensities, durations, frequencies, and bouts served to assess the viability and acceptability of iBounce.

3.8 Chapter Contributions and Relationship to the Thesis Research Questions

This chapter contributes and is related to the first research question:

RQ1 Can data mining techniques and models help to assess the effectiveness of physical activity interventions using activity sensors data?

(1) We presented a novel method to extract daily behavioural changes from physical activity data to meaningfully assess the impact of health interventions relative to the recommended daily goals. To measure changes in the participants' behaviour relative to the recommendations and health guidelines, the mean duration and frequency of bouts of physical activity levels were extracted from physical activity tracker data and analysed at the day level using clusters. This approach is valuable because it allows assessing the health education intervention effects in more detail by determining how many participants changed their daily behaviour, whether they now perform more physical activity, and how pronounced the change was.

(2) We described how our proposed technique can be used to further analyse physical activity behaviour changes in relation to the part of the day when it occurred. Clusters of the participants' mean duration and frequency of bouts of physical activity levels were analysed at the hour level. An hourly analysis of behaviours is valuable for assessing any intervention because it allows determining when behaviour changes occur during the day and what kind of physical activity change is made. The analysis of hourly behaviour changes might help to refine physical activity guidelines and recommendations based on the period or time of day that is more likely to generate a behaviour change.

CHAPTER 4

Unsupervised Early Detection of Physical Activity Behaviour Changes from Wearable Sensor Data

---

Wearable accelerometers record physical activity with high resolution, potentially capturing the rich details of behaviour changes and habits. Detecting these changes as they emerge is valuable information for any strategy that promotes physical activity and teaches healthy behaviours or habits. Indeed, this offers the opportunity to provide timely feedback and to tailor programmes to each participant's needs, thus helping to promote the adherence to and the effectiveness of the intervention. This chapter presents and illustrates U-BEHAVED, an unsupervised anomaly detection algorithm that periodically scans step data streamed from activity trackers to detect physical activity behaviour changes to assess whether they may become habitual patterns. Using rolling time windows, current behaviours are compared with recent previous ones, identifying any significant change. If sustained over time, these new behaviours are classified as potentially new habits. We validated this detection algorithm using a physical activity tracker step dataset (N=12,798) from 79 users. The algorithm detected 80% of behaviour changes of at least 400 steps within the same hour in users with low variability in physical activity, and of 1,600 steps in those with high variability. Based on a threshold cadence of approximately 100 steps per minute for standard walking pace, this number of steps would suggest approximately 4 and 16 minutes of physical activity at moderate-to-vigorous intensity, respectively. The detection rate for new habits was 80% with a minimum threshold of 500 or 1,600 steps within the same hour in users with low or high variability, respectively. This chapter is based on [31].

4.1 Introduction

Wearable accelerometry-based activity sensors are widely used to objectively, continuously, and unobtrusively record and monitor the physical activity of subjects in free-living conditions and produce a quantitative assessment of physical activity [112]. They provide detailed physical activity data that may hold valuable information on physical activity behaviours and habits. By retrieving activity sensor data regularly, behaviour changes can be detected as they emerge, helping to identify new behaviours in real time and to determine whether they are repeated over time, suggesting new physical activity habits. Detailed real-time information may be crucial to provide support and feedback, or to understand physical activity behaviour changes in the context of interventions or programmes promoting physical activity or teaching healthy behaviour habits. These can be critical in the fight against overweight and obesity, because physical activity contributes to weight control. Increasing physical activity is considered a progression toward healthy behaviours.

The early detection of recent changes in physical activity behaviour enables timely encouragement and re-enforcement of the newest healthy behaviours while discouraging unhealthy behaviours. The real-time automated detection of behaviour changes and/or new habits is needed to trigger timely personalised feedback that can allow self-reflection, the adjustment of physical activity goals to each participant, and the real-time individualisation of the contents and pace of a health intervention programme.

Many of the existing methods for detecting physical activity behaviour changes have been designed to analyse historical data, focusing on the detection of significant changes that occurred between specific moments (e.g. pre- and post-intervention). Others use specific physical activity thresholds to flag a pre-defined achievement (e.g. when the number of daily steps becomes higher than 10K). However, these methods cannot detect subtle but significant behaviour changes in real time. Indeed, behaviour changes can occur in small increments over a certain period.

This chapter describes the development of **U**nsupervised **BE**haviour and **HA**bit **V**ery **E**arly **D**etector (U-BEHAVED), a novel unsupervised anomaly detection machine learning technique that detects significant day-to-day changes in physical activity behaviour using step data from wearable sensors. The research questions (RQ) were:

RQ1: Can physical activity behaviour changes be periodically detected using step data from activity trackers?

RQ2: Can we detect whether these physical activity behaviour changes are sustained over time (suggesting a new habit)?

The chapter is structured as follows. Section 4.2 presents previous work on the detection of physical activity behaviours by sensor data mining. Section 4.3 discusses the challenges of detecting physical activity behaviour changes. Section 4.4 describes and illustrates the U-BEHAVED algorithm. Section 4.5 reports the U-BEHAVED algorithm's accuracy in detecting physical activity changes from sensor data. Section 6 concludes the chapter.

4.2 Related Work

A diverse number of machine learning techniques use data from wearable sensors to detect physical activity behaviour changes. The techniques can be grouped into supervised, unsupervised, and semi-supervised methods to identify changes during physical activity promotion interventions (real time) or after the interventions (not real time).

4.2.1 Supervised Machine Learning Techniques

Supervised machine learning techniques have been used to classify sensor-derived physical activity data by training a model with pre-classified sensor-derived physical activity data. In educational scenarios, researchers rely on this approach to identify the improvement or correction of physical movements towards expert standards, for example, specific body movements in martial arts [96] and dance [26], or of physical interactions with clinical equipment [69]. However, these techniques are not suitable to detect physical activity behaviour changes when the objective is to compare current physical activity behaviours with past ones. As physical activity behaviours differ between individuals, a specific volume of physical activity increases may represent a behaviour change for one individual but not for another. For instance, for a sedentary person, walking ten extra minutes in the morning can be a significant physical activity increase, thus signalling a physical activity behaviour change. Conversely, for a very active person, walking ten extra minutes may be insignificant compared with their normal volume of physical activity, and may not be considered a physical activity behaviour change. Therefore, the classification of behaviour changes must be adapted for each person and is unique at each point in time because behaviours constantly evolve.

4.2.2 Unsupervised Machine Learning Techniques

Unsupervised machine learning techniques have been used to create sets of similar physical activity behaviours that are analysed to identify physical activity behaviour changes. We review them under two sub-groups: (i) those for analysing pre- and post-intervention data and (ii) those for real-time detection.

(i) A first group of studies analysed health education behaviour changes after an intervention (post-intervention analysis, not in real-time). For instance, [47] analysed the impact of an intervention using a k-means algorithm to group the time spent at different physical activity levels (PAL) and step goal achievements. [62] assessed physical activity patterns during an intervention using k-means clustering to group the participants' hourly steps. [34] evaluated the impact of an intervention by grouping the participants' physical activity bouts using a k-means algorithm. [46] analysed physical activity patterns in women throughout the day using k-means clustering to group their daily metabolic equivalent (MET) per minute. [75] assessed the impact of sharing personal physical activity behaviours in an online community using agents. Lastly, [102] developed a window-based algorithm to detect and analyse changes in participants' behaviours after an intervention using time series of participants' steps captured by wearable sensors. All these studies detected changes in physical activity behaviours; however, they focused on physical activity behaviour changes after the intervention (post-intervention analysis) and used the participants' behaviours during the whole intervention to determine significant behaviour changes. These techniques are not suitable for our purpose because they work once all the physical activity data recording period is available, and they aim to identify the most prominent behaviour changes among all behaviours over the recording period. In contrast, we aim to detect significant behaviour changes as they appear. Such early and timely detection of behavioural changes is key to provide quick support and feedback for promoting physical activity.

(ii) A second group of unsupervised techniques analyse the participants' physical activity behaviour changes during interventions to promote physical activity (real-time analysis). For instance, [21] helped participants to examine their daily behaviour by grouping their physical activity using a k-means algorithm to cluster the mean heartbeat and oxygen saturation values. [4] generated daily personalised text messages with custom timing, frequency and feedback about their step count/goal and with motivational content to support reflection using a multiarmed bandit (MAB) algorithm and the number of minutes spent in physical activity. [5] used a genetic algorithm and pareto-optimality, and the participants' daily sleep duration, steps, calories, exercise duration, exercise distance, exercise calories, step count, step distance, and step calories to analyse the wearer's data and make personal lifestyle improvement recommendations. [18] recommended physical activity to participants using an agglomerative cluster technique that grouped MET values, heart rate, gender, height, weight, age, and exercise time, type and frequency. [45] personalised the intervention for each participant using an MAB to model the participant's days in the intervention, physical activity goal compliance, weight, food intake, and calories. [74] explained the PAL dynamics in a community using a social contagion model to model the steps and results of a psychological questionnaire on self-efficacy, barriers, social norms, long-term goals, intentions, satisfaction, outcome expectations and models. [86] generated personalised suggestions to help users reach their physical activity behaviour goals using MAB and physical activity frequency and calories. [115] adapted the step goal settings of the intervention for each participant using a behavioural analytics algorithm, the daily steps, and a goal. [116] used an MAB for step modelling, in addition to motivation and psychometric data to personalise the social comparison among participants with the aim of motivating them towards increasing their physical activity behaviour. Finally, [48] developed a reinforcement learning recommendation system that used clustered daily segments of participants' steps and sleep behaviours to provide personalised suggestions of physical activity patterns to achieve weight-loss effects.

This second group of unsupervised techniques used a threshold to detect when a physical activity change is important: a step goal, a cluster (sorted by physical activity), or any other measure.

4.2.3 Semi-Supervised Machine Learning Techniques

Some studies used thresholds to train classification algorithms, creating semi-supervised techniques to detect physical activity behaviour changes. For instance, [12] provided personalised daily activity recommendations using shallow neural networks to process physical activity and demographics, attitudes, intentions and habit data from questionnaires. [98] created gamified personalised feedback using a random forest technique and a weighted score to model the physical activity change from accelerometry data. [36] delivered personalised feedback to participants about their progress to help them achieve their personal step goal using a random forest technique to model the hour when physical activity work was performed, the hourly number of steps for that hour, the number of steps made in the past hour, the cumulative number of steps up to that hour, and the mean number of steps on workdays. Lastly, [53] personalised feedback time and content using k-nearest-neighbour and support vector machine techniques to model physical activity physical variables (not specified).

4.2.4 Summary of Related Work

Although unsupervised and semi-supervised algorithms detect real-time individual behaviour changes in physical activity promotion contexts, they rely on two different methods: comparing the aggregated physical activity data or determining whether a predefined physical activity objective is met. When using aggregated physical activity data, all details of any significant behaviour change below the aggregated level are diluted, and cannot be detected. For example, at the daily level, it would be possible to detect a significant physical activity increase in a given day, but without details on whether, when, and how many different behaviour changes occurred during that day. Similarly, when using a physical activity objective, any physical activity change below that goal is not detected. For instance, if a sedentary participant adds a new physical activity during a day (e.g. going for a walk), but does not reach the step goal, this behaviour change is not detected. Conversely, our objective is to detect all significant behaviour changes, including small and subtle behaviour changes, because future habits are progressively built on past habits, generating notable behaviour changes.

In conclusion, there are successful algorithms to detect behaviour changes. Conversely, methods to identify real-time progressive behaviour changes in the framework of physical activity promotion interventions and healthy behaviour teaching are lacking. Here, we extended these previous works on detecting physical activity behaviour changes by creating an unsupervised machine learning technique that identifies hourly outliers in participants' physical activity behaviour changes. We also propose a method to recognize sustained physical activity behaviour changes that suggest new habits.

4.3 Challenges of Detecting Physical Activity Behaviour Changes in Health Education

Physical activity behaviour changes are indicated by differences in the step number between days; however not all step differences are true behaviour changes. Indeed, step differences between days can be observed due to the natural variability of the usual daily physical activity [11] [54], making the step difference magnitude to be considered an undefined behaviour change.

Physical activity variations are highly individual. Therefore, the first challenge is to identify significant step differences while taking into account the variations of each participant.

Even people who maintain regular daily physical activity are expected to show some variability in their execution time and step number. For instance, a person who runs every morning will not run exactly at the same time and perform exactly the same number of steps. The second challenge is to avoid flagging habitual physical activity with shifted execution time and similar step number as a behaviour change.

When a physical activity behaviour change is detected, it could either be a transient variation not sustained in time or the beginning of a new behaviour. Therefore, the third challenge is to flag physical activity behaviour changes that are maintained over time, suggesting that they are potentially becoming habits.

To exploit the patterns detected in a health education programme, it is important that they are detected in a timely manner, which means shortly after they occur. The fourth challenge is processing large amounts of physical activity data almost in real time to detect physical activity behaviour changes briefly after they occur.

4.4 Methods

The U-BEHAVED algorithm uses continuous and streaming physical activity data from activity trackers. Data are pre-processed, resulting in a time series of hourly steps coarse enough to detect intra-day changes and avoid mislabelling any energy burst (e.g. a short sprint) as a behaviour change. To detect the behaviour changes as they occur, the algorithm first computes the usual behaviour by building a rolling time window that creates a time series of the mean number of steps per hour. Next, to measure the magnitude of the behaviour change, the algorithm calculates the difference between the hourly steps of each day and the mean number of steps per hour. Outliers of the hourly step difference are classified as behaviour changes using a second rolling time window to avoid incorrectly flagging habitual physical activity with negligible changes in execution time and step number as changes. Finally, to flag a continuous behaviour change as a new habit, the algorithm moves the outlier limits from the day when the initial behaviour change is detected to the subsequent days.

The algorithm extends the work discussed in Section 4.2 and relies on time series anomaly detection of residuals [13] and interquartile ranges (IQR) [59]. The algorithm was implemented in R [85] and the source code is available at the Open Science Framework [29]. In the next section, we explain in detail the data requirements to serve as input to U-BEHAVED, how data must be pre-processed, the algorithm steps, and the resulting outputs.

4.4.1 Dataset Requirements

The input dataset must be a discrete time series of steps, regularly spaced at the input sampling rate, and synchronised with U-BEHAVED at regular intervals. We used an hourly interval as a sensible time interval to identify changes in human activity. However, other regular intervals can also be used. Datasets from any physical activity tracker device that can be transformed into number of steps with $\leq$ 1-hour sampling can be used, such as step data from (or computed from) commercial wearable devices (e.g. Fitbit), smartphone devices, and research-grade devices (e.g. GENEActiv).

The completeness and accuracy of the dataset is important because periods with missing or erroneous PA tracker data (for instance caused by non-wear time) could result in erroneous behaviour-change detection (false positives) or lack thereof (false negatives). In the event of missing data, U-BEHAVED would skip the detection for that period. We will explain this process further in the next section.

4.4.2 Data Pre-Processing

When the discrete time series of steps is streamed to U-BEHAVED, the sum of the steps is stored for that corresponding hour. In case the data is incomplete for any reason, i.e., there is not one full hour worth of data, the sum is not calculated and NULL is stored. The U-BEHAVED data pre-processing produces two vectors, one containing information about the day and hour of the steps recorded, and the other about the total number of steps performed during that day and hour (See Table 4.1). It could possibly be NULL if data were missing for that hour.

TABLE 4.1. Example of total number of steps per hour

Day Hour	**Total Number of Steps**
1 November 2021 08:00	390
1 November 2021 09:00	564
1 November 2021 10:00	1046
...	...

4.4.3 U-BEHAVED Algorithm

The purpose of the algorithm is to detect significant behaviour changes and new habits as they emerge. Using the two vectors updated during data pre-processing (Section 4.4.2) it compares current behaviours with recent ones using rolling time windows. The width of the rolling time windows are adjustable to reflect the period used as recent behaviours. The algorithm is executed every hour and is divided into five steps: i) calculation of the mean number of steps per hour using a rolling time window, ii) calculation of the step difference per hour, iii) definition of the upper and lower limits using moving IQR from the step number difference, iv) classification of step difference outliers as behaviour changes, and v) classification of consecutive outliers as new habits. If the vector contained NULL (due to missing data, as explained in Section 4.4.2), the detection for that hour is entirely skipped. The algorithm steps for non-missing data are summarised in Figure 4.1 and explained below.

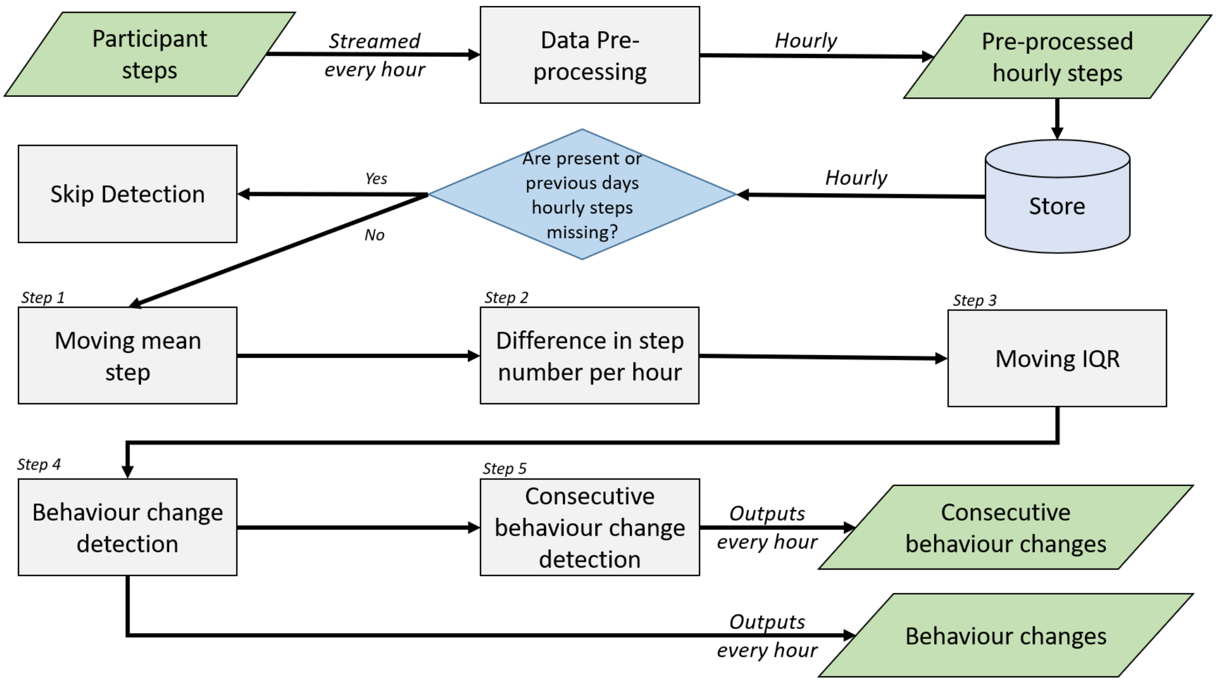

Figure 4.1. Diagram of U-BEHAVED algorithm steps

Step 1: Moving mean step number per hour. The moving mean number of steps per hour is calculated by rolling a time window of w days over the hourly steps per day (the two vectors obtained from the pre-processed data) and using Equation 4.1 where $AvgSteps_{d,h}$ is the mean number of steps per hour at day d and hour h, and $S_{d,h}$ is the number of steps per hour at day d and hour h. This results in a vector of the mean step number per hour. If any of $S_{d,h}$ contain a NULL value, the algorithm exits this cycle, skipping the detection.

$$AvgSteps_{d,h} = \sum_{i=d-w}^{d} \frac{S_{i,h}}{w} \quad \text{where } d > w \tag{4.1}$$

Step 2: Difference in step number per hour. The difference in step number per hour is calculated by subtracting the hourly number of steps per day (vector from the pre-processed data) from the rolling windowed mean number of steps per hour (vector from Step 1) using Equation 4.2, where DS is the difference in step number per hour for day d and hour h. This process results in a vector of hourly step differences.

$$DS_{d,h} = S_{d,h} - AvgSteps_{d,h} \tag{4.2}$$

A positive DS value means that the participant did more steps per hour in the present day compared with the mean number for the w days. Conversely, a negative value means that the participant did fewer hourly steps in the present day than in the w days.

Step 3: Moving IQR. The moving IQR from the last w days are calculated by rolling a time window over the difference in step number per hour (vector from Step 2) and using Equation 4.3, where the IQRs for day d are obtained by subtracting from the 75$^{\text{th}}$ percentile of the difference in step number per hour (DS) in the last w days the 25$^{\text{th}}$ percentile of the difference in step number per hour (DS) in the last w days.

$$IQR_d = p75_{DS_{d-w},\ldots,DS_{d-1}} - p25_{DS_{d-w},\ldots,DS_{d-1}} \tag{4.3}$$

Then, the upper limit (UL) and lower limit (LL) of day d are calculated using the last w days (Equations 4.4 and 4.5, respectively). This results in two new vectors that contain the daily upper and lower limits.

$$UL_d = p75_{DS_{d-w},\ldots,DS_{d-1}} + 1.5 * IQR_{d-1} \tag{4.4}$$

$$LL_d = p25_{DS_{d-w},\ldots,DS_{d-1}} - 1.5 * IQR_{d-1} \tag{4.5}$$

Step 4: Behaviour change detection.

As expressed in Equation 4.6, differences in step number per hour above the daily upper limit (from Step 3) are classified as positive behaviour changes, and differences of steps per hour below the daily lower limit (from Step 3) are classified as negative behaviour changes.

$$Class_{d,h} = \begin{cases} Positive\,Behaviour\,Change & \text{if } DS_{d,h} > UL_d \\ Negative\,Behaviour\,Change & \text{if } DS_{d,h} < LL_d \\ No\,Behaviour\,Change & \text{otherwise} \end{cases} \tag{4.6}$$

Step 5: Consecutive behaviour change detection.

For each detected behaviour change, Equation 4.7 is used to calculate the difference between the hourly step number from the pre-processing data and the mean step number per hour from when the behaviour change was initially detected.

$$DS_{d,h,d_{detection},h_{detection}} = S_{d,h} - AvgSteps_{d_{detection},h_{detection}} \tag{4.7}$$

Then, this difference is compared with the limits when the behaviour change was initially detected using Equation 4.8. If it is consecutively higher than the upper limit, the behaviour change is classified as a positive habit. If it is consecutively lower than the lower limit, it is classified as a negative habit.

$$Habit_{d,h,d_{detection},h_{detection}} = \begin{cases} Yes & \text{if } DS_{d,h,d_{detection},h_{detection}} > UL_{d_{detection}} \text{ and consecutive} \\ No & \text{if } DS_{d,h,d_{detection},h_{detection}} < LL_{d_{detection}} \text{ or non-consecutive} \end{cases} \tag{4.8}$$

4.4.4 Output

The algorithm output is represented by two data frames. The first data frame (Table 4.2) contains in each row a detected behaviour change, and in the columns it contains the day and hour when the behaviour change occurred, the number of additional (or fewer) steps made, and if the behaviour changes were positive or negative. The second data frame (Table 4.3) contains in each row a detected behaviour change sustained over time, and in each column it contains the day and hour when the sustained behaviour change was first detected, the day when it was last detected, and whether the sustained behaviour change were positive or negative.

TABLE 4.2. Example of data frame with the detected behaviour changes

Day and Hour	**Step Number Difference**	**Type of Behaviour Change Detected**
11 November 2021 11:00	-956	Negative
15 November 2021 15:00	2300	Positive
15 November 2021 17:00	1549	Positive
...	...	...

TABLE 4.3. Example of data frame with the detected sustained behaviour changes

First Day and Hour	**Last Day**	**Type of Habit**
16 November 2021 13:00	18 November 2021	Positive
18 November 2021 14:00	19 November 2021	Positive
21 November 2021 08:00	25 November 2021	Negative
...	...	...

4.4.5 Illustration

To illustrate the U-BEHAVED steps, we used data of one participant from a previous health education intervention [111]. We defined the width of the rolling time window to three days

($w = 3$) because the intervention had educational content delivered every three days. Figure 4.2 shows the first four steps of the U-BEHAVED algorithm. In Step 1, the time window was rolled through the hourly steps (blue box) to calculate the mean step number per hour (red box). In Step 2, the hourly steps were subtracted from the mean number of steps per hour (yellow box) to generate the hourly differences in step number (green box). In Step 3, the IQR limits were calculated (purple box), and in Step 4 they were used as thresholds to detect significant behaviour changes (green box).

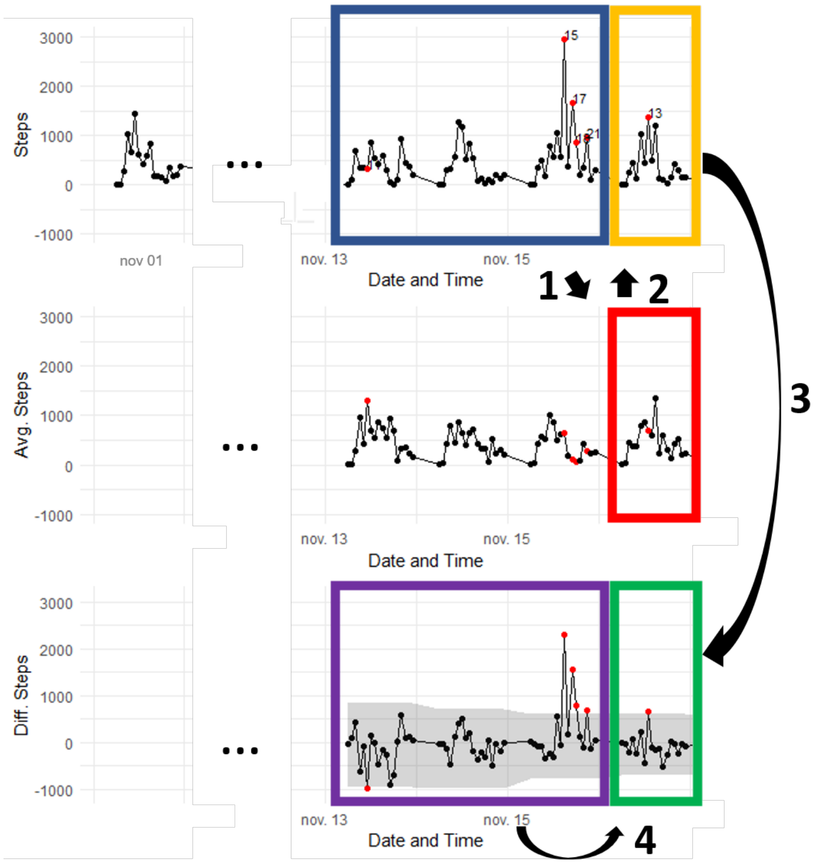

Figure 4.2. Illustration of first four steps of U-BEHAVED algorithm. Arrows and numbers represent algorithm steps and coloured boxes highlight relevant data of each step. X−axis represents date and time (in hours). Y−axis represents total number of steps (upper graph), mean number of steps using 3-day window (middle graph), and difference between present day and mean number of steps (lower graph). Black dots: data points; red dots: detected behaviour changes; and grey area: IQR.

Step 5 of the U-BEHAVED algorithm is described in Figure 4.3. It scans the step number difference (pink box) using the IQR of the previously detected behaviour changes (orange box). If the step number difference is outside the IQR limits of previously detected behaviours, the behaviour change is sustained and is labelled as a habit.

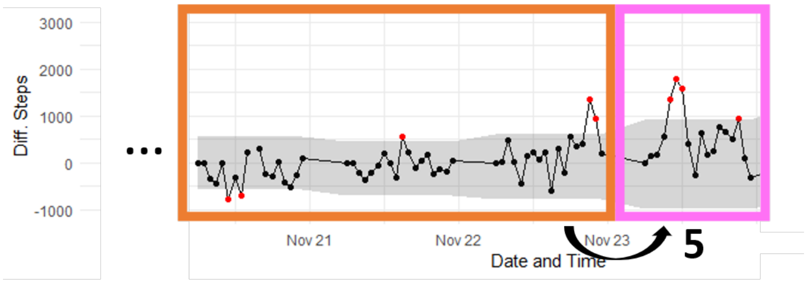

Figure 4.3. Illustration of Step 5 of U-BEHAVED algorithm. Arrow and number represent Step 5 and coloured boxes highlight relevant processes. X−axis indicates time (in hours). Y−axis is difference in step number.

4.5 Evaluation

We used physical activity tracker data with pre-labelled physical activity behaviour changes and habits to calculate the accuracy of the U-BEHAVED algorithm. It was constructed from real physical activity tracker data collected as part of health studies (Real Raw Data) where all significant physical activity changes were smoothed out to remove all behaviour changes and habits (Baseline Dataset). We set the width of the algorithm rolling window to a short period ($w = 3$ days) and added different magnitudes of physical activity behaviours (Baseline Dataset + Controlled Changes) and habits (Baseline Dataset + Controlled Changes + Controlled Habits) in a controlled manner. We simulated an hourly stream of data to the algorithm, and for each added magnitude of physical activity behaviour, we calculated the algorithm accuracy as the ratio of detections relative to all added behaviours. The evaluation method is summarised in Figure 4.4.

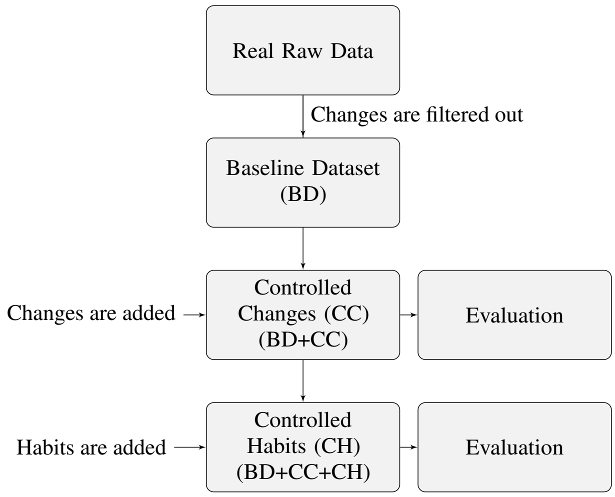

Figure 4.4. Method to evaluate U-BEHAVED accuracy in detecting behaviour changes and habits

4.5.1 Construction of the Evaluation Dataset

The real raw data contained records of 79 participants who used a physical activity tracker that continuously recorded their number of steps per minute. We divided the 79 participants into two groups. The first group included 30 children from a private school-based health education programme [17], [111] (median number of hourly steps=320, mean hourly step coefficient of variation=50%), and each wore a Misfit wearable sensor [73]. The second group included 49 adults from two public datasets [15] [104] (median number of hourly steps=218, mean hourly step coefficient of variation=140%), and each wore a Fitbit wearable sensor [44].

Using the Real Raw Data, we built the Baseline Dataset (BD) free of behaviour changes and habits. We first pre-processed the data for each participant (Section 4.4.2) and calculated the mean step number per hour for each participant. Then, we replicated this nine times for each participant, simulating nine equal days of physical activity. We needed nine days of physical activity because we used the first six days as baseline (the width of the algorithm rolling window was set to three days), and then we added behaviour changes in the last three days to evaluate the algorithm. We simulated the participants' natural physical activity variability [11] [54] by introducing (or removing) steps per hour based on each participant's hourly step coefficient of variation, calculated from the pre-processed Real Raw Data. We manually inspected the resulting dataset of 12,798 data points (79 participants x 18 hours x 9 days) for sharp increases or decreases in the number of steps per hour to confirm that behaviour changes were not involuntarily introduced.

4.5.2 Evaluation of Behaviour Change Detection

We incorporated two behaviour changes to the BD of each participant on day seven: one positive behaviour change at 8 am by adding physical activity (steps), and one negative behaviour change at noon by removing physical activity (steps) (Figure 4.4). We chose these specific times because they are times when participants may generate behaviour changes, such as walking to school/workplace in the morning and using a computer or smartphone at lunchtime. We also varied the number of steps.

Figure 4.5 shows the detection accuracy for various amounts of steps. As expected, as more steps were added or removed to simulate behaviour changes, the accuracy of the algorithm increased sharply.

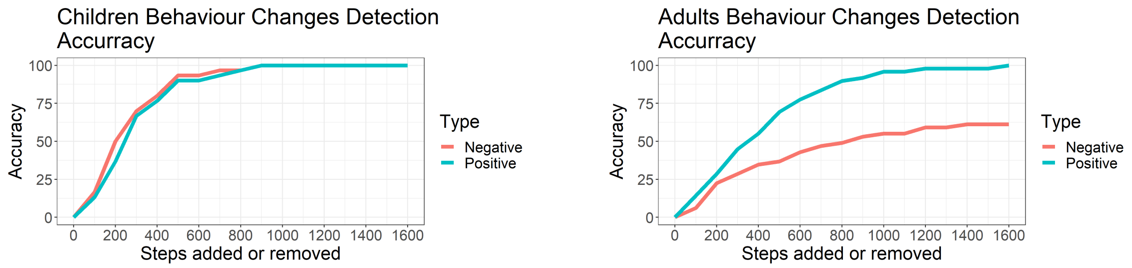

Figure 4.5. Left and right graph show behaviour-change detection accuracy of children and adults, respectively. X−axis indicates number of steps added at 8 am as positive behaviour changes or removed at noon as negative behaviour changes. Y−axis indicates percentage of behaviour changes detected relative to all changes added. Red and blue lines indicate percentage of negative and positive behaviour changes detected, respectively.

For the children's dataset, the algorithm detected ∼80% of the positive and negative behaviour changes when the added changes corresponded to at least 400 steps per hour. This detection accuracy reached 100% when the added changes corresponded to at least 900 steps per hour.

For the adults' dataset, the algorithm detected ∼80% and 100% of positive behaviour changes when the added changes corresponded to at least 600 steps and 1,600 steps per hour, respectively. The detection of negative behaviour changes by the algorithm reached 61% when the added changes corresponded to at least 1,400 steps per hour. Overall, 80% of all (positive and negative) added behaviour changes were detected at 1,600 steps.

4.5.3 Habit Detection Evaluation

Similar to the behaviour change detection evaluation, we incorporated two habits for each participant from day 7 to day 9: one positive habit at 8 am by adding the same amount of physical activity (steps) during the three consecutive days, and one negative habit at noon by removing the same amount of physical activity (steps) during the three consecutive days (Figure 4.4). The habit was labelled as detected only if it was consecutively detected from day 7 to day 9.

Figure 4.6 shows the detection accuracy at various amount of steps. Before the addition of the two habits (i.e. addition = 0), the algorithm did not detect any habit, as expected. Then, the number of detected habits increased sharply as these became more pronounced.

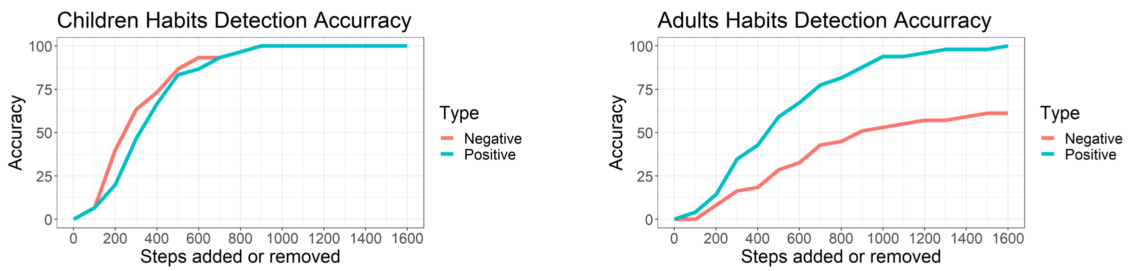

Figure 4.6. Left and right graphs show habit detection accuracy for children and adults, respectively. X−axis indicates number of steps added at 8 am (positive behaviour changes) or removed at noon (negative behaviour changes). Y−axis indicates percentage of habits detected in function of number of steps added. Red and blue lines indicate percentage of negative and positive habits detected, respectively.

For the children's dataset, the algorithm detected ∼80% of the positive and negative habits when the added habits corresponded to at least 500 steps per hour. This detection accuracy reached 100% when the added habits corresponded to at least 900 steps per hour.

For the adults' dataset, the algorithm detected ~80% of positive habits when at least 700 steps per hour were added and increased to 100% following addition of at least 1,600 steps per hour. The detection of negative habits by the algorithm reached 61% upon addition of at least 1,500 steps per hour. Overall, 80% of all (positive and negative) added habits were detected at 1,600 steps.

4.6 Discussion and Conclusion

This chapter presents U-BEHAVED, an unsupervised anomaly detection algorithm for detecting physical activity behaviour changes and new habits as they appear. The algorithm identifies significant changes in current behaviours by comparing them to recent past behaviours using rolling time windows of participants' step count data captured in real-time using wearable physical activity trackers. The detection of physical activity behaviour changes and new habits as they appear represents valuable information for a physical activity promotion strategy because it can help to increase its effectiveness and the participants' adherence by enabling personalisation [65]. For instance, by detecting and understanding how participants change their behaviours in real-time, physical activity promotion strategies can be adjusted to each participant's needs, and relevant personalised feedback can be generated. It can also help to assess and better understand how the programme or intervention influences participants by analysing their behaviour changes and new habits after the physical activity promotion strategy ends. It may also reveal how the participants' behaviour changes affect other lifestyle behaviours, such as sleep and diet.

The evaluation of U-BEHAVED detection accuracy using data from 79 users showed that it can successfully detect significant physical activity behaviour changes and new habits, even when subtle, in the general population with a non-pathological gait. In the children's dataset (lower physical activity variability), the algorithm detected 100% of behaviour changes and new habits when a difference of at least 900 steps per hour was added. In the adult dataset (higher physical activity variability), U-BEHAVED detected 80% of behaviour changes when a difference of at least 1,600 steps per hour was added. The difference in the number of steps needed to detect behaviour changes and new habits was lower in the children's than in the adults' dataset because the change detection by U-BEHAVED is based on the participants' physical activity variability, which was lower in the children's than adults' dataset. The algorithm can detect behaviour changes and habits as subtle as a 100 steps difference per hour if they are significant, for instance for a sedentary participant with low physical activity variability. The difference in steps would approximate 9 and 16 minutes of physical activity at moderate-to-vigorous intensity for the children and adult datasets, respectively, based on a threshold cadence of around 100 steps per minute, which represents a habitual walking pace [76], [78], [106]. Examples of behaviour changes and habits that can be detected in children are walking with their parents to school in the morning [20] [71] and playing an active game at lunchtime [90], and for adults, examples include active commuting [100] and exercising at lunchtime [49].

The U-BEHAVED algorithm can be easily implemented in the framework of any strategy that promotes physical activity or teaches healthy behaviours and relies on physical activity trackers because it uses as input data the number of steps performed by participants (i.e. the output data of most commercial, smartphone, and medical-grade physical activity trackers). The algorithm can also be easily adapted to detect behaviour changes at any programme content delivery pace because the algorithm rolling window can be set to any length. For instance, if the health intervention delivery pace is weekly, the rolling window can be set to 7 days. The flexibility to adjust the window length enables comparisons of past behaviour in the short, medium, and long term. For instance, to detect changes in current behaviour compared with the previous month, the rolling window can be set to 30 days.

U-BEHAVED uses a step aggregation level per hour to detect physical activity behaviour changes as they occur. This aggregation level was selected because it is coarse enough to detect intra-day changes and avoid mislabelling any insignificant variation in the step number as a change. Other aggregation levels could be used to detect other behaviour change types; however this would lead to new misclassification issues. For instance, a lower aggregation level, such as per minute, might allow for identifying more subtle behavioural changes but it might mislabel small physical activity bursts as behaviour changes when they are not. Additional methods need to be developed for each aggregation level to avoid misclassification.

We note that while steps captured by physical activity trackers are acknowledged as being a reliable high-level indicator of a person's total amount of physical activity [105] with a high step activity recognition rate [6], they may not capture all physical activities. U-BEHAVED focuses on detecting behaviour changes in physical activity that can be defined as steps and excludes other non-ambulatory activities, such as cycling [72]. Future work can explore the inclusion of user-defined activities into the algorithm, as well as the inclusion of additional data that may help in detecting the presence of a non-ambulatory activity, such as intensity levels. This would also address the issue of missing PA data caused by the removal of the tracker (non-wear time) or by a technical error during recording or synchronisation.

Furthermore, the detection of other types of habits can be explored. We flagged as new habits behaviour changes that were performed consecutively every day at the same hour; however, different behaviour changes can be performed at different times of the day, non-consecutively or with a broader difference of time, such as doing sport only on Mondays. This suggests that many types of habits remain to be detected.

4.7 Chapter Contributions and Relationship to the Thesis Research Questions

This chapter contribute and is related to the second research question:

RQ2 Can data mining techniques help to detect behaviour changes and generate personalised feedback to support learning?

(1) We presented U-BEHAVED, a novel unsupervised anomaly detection data mining algorithm for detecting significant changes in physical activity behaviours as soon as they appear. This information is valuable for any health education intervention because it provides real-time information on when behaviour changes happen and

whether they are positive (or not). Real-time detection of behaviour changes enables the development of timely and personalised feedbacks in response to each participant's requirements, such as modifying the content of a health education programme or setting custom physical activity objectives.

(2) We showed that U-BEHAVED can detect physical activity behaviour changes that are sustained over time, suggesting a new habit. This information is valuable for health education interventions because it informs whether participants learned a new habit, when they learned it, and whether it is healthy (or not). This information can allow the development of timely and personalised feedback in response to each participant's behaviour with the aim of strengthening healthy habits, discouraging unhealthy habits, and re-establishing healthy habits that have been abandoned.

CHAPTER 5

Modelling Physical Activity Behaviour Changes for Personalised Feedback in a Health Education Application

---

Activity sensors have made their way into Intelligent Tutoring Systems to support physical learning. Open-ended domains, where the focus is not about learning specific expert movements but about adopting healthy physical activity behaviours, require the use of algorithms and artificial intelligence in education techniques for modelling evolving patterns from physical activity sensor data to enable feedback and personalisation. Specifically, this requires (1) techniques to detect and monitor undetermined shapes of physical activity changes from accelerometer data, (2) modelling these detected behaviour changes and integrating them into the learner model, and (3) updating them periodically for meaningful and timely feedback. We present a suite of window-based algorithms that detect physical activity changes aligned with learning objectives from accelerometer data. These are translated into learner model attributes and used to generate timely feedback. We illustrate our method in the context of iEngage, a health education programme that teaches adolescents about healthy physical activity behaviours through an application connected to a wrist-worn activity tracker. We present the feedback generated by our algorithms and report on the qualitative evaluation with four experts. We conclude that the automated feedback is useful, important, and timely to leverage adolescents' physical activity learning. Two of the four algorithms presented are part of the U-BEHAVED algorithm previously described in Chapter 4. This chapter is based on [35].

5.1 Introduction

Artificial Intelligence in Education (AIED) systems have focused on the students' interactions with traditional computer interfaces; however, the recent availability of affordable sensors has opened opportunities to capture new interaction types in physical and motor skill domains [69] [95], creating the need of novel learner modelling techniques to drive personalisation. In a traditional AIED system, feedback may be based on the knowledge and mistakes made by the student who executes some tasks using a computer. On the other hand, in physical learning systems, feedback may focus on the physical activity behaviours detected by the sensors. This means that the system needs to extract and assess relevant information from all sensor data in real-time, in addition to the more standard forms of computer-based interactions.

Our goal was to create techniques for extracting relevant physical activity behaviour change data from activity tracker time-series to determine whether students learn and improve their physical activity behaviours in the context of a health education programme dedicated to 10-12-year-old adolescents called iEngage. This programme is run in Australian schools for five weeks. iEngage teaches health knowledge and skills related to physical activity goal setting and self-assessment of achievements, and aims to promote participation in physical activity [111] [47] [17]. The learning activities are delivered through an application connected to a wrist-worn activity tracker that records steps continuously throughout the programme running period. The activity tracker helps adolescents to monitor their physical activity towards personal step goals, and supports knowledge acquisition through experiential learning activities. The learning programme delivers evidence-based content about physical activity and health, complemented by practical activities and quizzes for knowledge consolidation. The activity tracker captures steps and feeds them back to the application. This information is then presented to the adolescents in the application, thus allowing them to self-assess their achievements against their individual goals. Indeed, learning is not just about how well students know the theoretical health facts, but also about how they translate this knowledge into actual physical activity behaviours in their daily life. For instance, after learning how to recognise the various physical activity intensities and the importance for their health of performing 60 minutes of moderate to vigorous physical activity (MVPA) each day, the activity tracker data can tell them whether they have actually changed their behaviour accordingly (such as by increasing their daily time spent doing MVPA), and whether they have acquired new habits (such as exercising after school, or walking to school). As the detected patterns will inform the feedback, it is important to capture, model and analyse patterns of physical activity changes in response to the health education programme. This allows determining whether new physical activity behaviours are observed in response to what they have learnt, whether these new behaviours are positive or negative, when they occur, and, importantly whether these new behaviours are maintained over time. These data are crucial to provide individualised feedback to the students.

The task is three-fold: (1) to detect changes in physical activity behaviours using the student's step data continuously recorded by the activity tracker, (2) to model and integrate them into the learner model, and (3) to update these data periodically to give timely personalised feedback. To this aim, we propose a framework and algorithms that we tested using step data and personal PA goals extracted from the e-learning platform from students who previously participated in the iEngage programme. We also validated the value of the automated feedback with experts.

Chapter 5 is structured as follows. Section 5.2 outlines related work in which physical activity sensors have been used to support physical activity learning, and algorithms to detect behaviour changes. Section 5.3 presents our methodology, including the algorithms for detecting and extracting physical activity behaviour changes, the learner model attributes that capture relevant physical activity behaviours, and how these attributes are used to generate personalised feedback. Section 5.4 describes the evaluation of the automated feedback by experts. Section 5.5 concludes the chapter and, finally, section 5.6 outlines the chapter contributions and relationship to the thesis research questions.

5.2 Related Work

The emerging use of sensors to support physical activity and psycho-motor learning has led to the development of specific data analysis techniques for tracking physical behaviour patterns, modelling the learners' behaviour, and providing personalised feedback. For instance, [27] used a smartphone application that tracks the movement of students while practising dance exercises to synchronize the student's body movements with the song, and to extract relevant learning characteristics, such as rhythm, pauses, and step size. At the end of the learning session, these characteristics are presented to the students as a feedback on their performance in the form of numbers, graphics, and texts. In the context of Aikido training, [96] used accelerometers and cameras to capture the students' body motion while they perform specific Aikido movements, in order to extract a range of relevant movement features. These features were then compared with recorded expert movements to trigger vibrotactile real-time feedback via actuators built in the student's training uniform. For healthcare training, [39] built a training simulation system using a manikin and sensors (accelerometers and movement sensors) to support nurse practice training. The algorithms extracted relevant features (e.g. the order of the tasks done by nurses) that were represented visually to help teachers identify and judge the students' correct and incorrect behaviours.

On the other hand, learning about healthy physical activity behaviours does not involve learning prescribed postures or movements patterns, but achieving an overall goal represented by the recommended physical activity level [80]. Moreover, physical activity changes are specific to each learner and are influenced by their fitness levels and daily schedule. Importantly, physical activity does not follow a particular pattern, occurs at various times of the day, with various characteristics of frequency, speed, duration and intensity. This implies using techniques to capture physical activity changes for each learner during the intervention programme instead of specific activity recognition supervised techniques.

Various unsupervised techniques exist for detecting physical activity behaviour changes in cohorts or in individuals after an intervention [62] [37] [34]. However, these methods are not suitable for monitoring physical activity behaviour changes in real time. To provide feedback during a health education programme or intervention, live streaming data need to be collected, and analysed using window-based techniques. For instance, [23] used semi-supervised algorithms to detect physical activity changes in behavioural routines during functional health assessments. They used data collected by a range of smart home sensors (motion, temperature and accelerometers) and a trained activity recognition algorithm to label activities and produce activity curves. They could detect dissimilarities using a Permutation-based Change detection in Activity Routine (PCAR) algorithm and the Kullback-Liebler (KL) divergence distance metric for detecting dissimilarities between activity curves. Further expanding this work, [102] created an unsupervised Physical Activity Change Detection framework, an adapted version of PCAR that calculates KL distances between moving windows of physical activity time series to produce a change score to track behaviour changes from wearable sensor data. Our work extends these ideas by capturing the characteristics of the students' physical activity behaviour changes that consolidate their knowledge acquisition.

5.3 A Learner Model for iEngage

A learner model for physical learning environments needs to include knowledge-specific attributes (i.e. what the student knows) and physical activity behaviour-specific attributes (i.e. what the student actually does). In iEngage, the knowledge-specific attributes rely on the interactions with the application (quizzes, answers, and set goals), while the physical activity behaviour-specific attributes are constructed from the data collected by the activity tracker continuously worn by the student.

Using the physical activity data from the activity trackers and the knowledge-specific data from the e-learning application, we generated learning model attributes using algorithms based on rolling window procedures. Then, we used these attributes to produce personalised summative and formative feedback content. The method is illustrated in Figure 5.1.

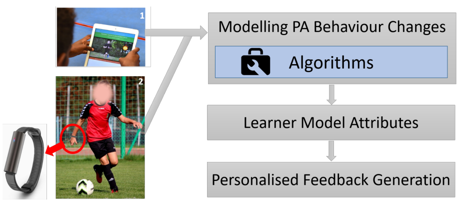

Figure 5.1. Schematic illustration of the method. Quizzes, answers, and goals are captured by the application (top-left photograph), and real-life physical activity by the wrist activity tracker (bottom-left photograph). These data are processed by algorithms and stored in the learner model as attributes to generate personalised feedback

5.3.1 Algorithms for Modelling Physical Activity Behaviour Changes

Detecting physical activity behaviour changes requires comparing present and past behaviours to identify significant differences. As physical activity behaviours are individual and dynamic, the developed algorithms use moving time windows that iteratively go through time frames of each student's physical activity behaviour and analyse the presence and nature of the changes.

Four algorithms were built to detect physical activity behaviour changes from step data: Difference Scanner, Rolling Window Linear Regression, Window-Based Interquartile Range (IQR), and Multi-Shifted Window-Based IQR, the latter two of which are components of the U-BEHAVED algorithm described in Chapter 4. The fourth algorithm relies on the results of the third, while the first three algorithms are independent of one another. Each algorithm has a specific aim linked with the iEngage themes and the learner model attributes (Table 5.1).

5.3.1.1 Difference Scanner Algorithm

The aim of this algorithm is to detect whether a student performed an hourly physical activity behaviour change by comparing it with the cohort's hourly physical activity behaviour (baseline). Using the hourly steps, it indicates when students do less physical activity than their peers. Specifically, as illustrated in Figure 5.2, the student's number of steps per hour (in red) is compared with the cohort's median number of steps per hour (in blue). This allows identifying the hours when the student does fewer steps than the cohort (black dots). The algorithm returns a data frame with two vectors: the day and the time when the difference is detected.

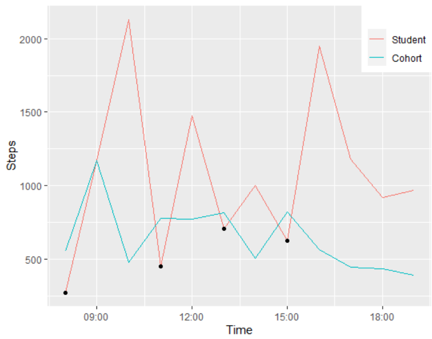

Figure 5.2. Illustration of the results produced by the Difference Scanner algorithm using data from one day. The X-axis display the time (in hours) and the y-axis the number of steps. The red line represents the student's steps number and the blue line the cohort's median steps numbers. Black dots represent the hours when the student is doing fewer steps than the cohort's median

5.3.1.2 Rolling Window Linear Regression Algorithm

The aim of this algorithm is to detect recent daily physical activity behaviour changes. Using the number of steps per day, the algorithm measures the student's physical activity trend during the previous days. Specifically, it uses a 3-day window to scan the students' time series of daily steps, and calculates iteratively the linear regression for each rolling window. Figure 5.3 shows the 3-day moving window at two instants (day 3 in red and day 6 in green) with their calculated regression coefficients at the top. It returns a data frame with two vectors: the day where the window is located and the resulting linear regression coefficients.

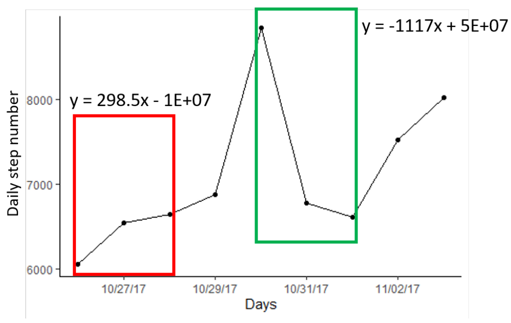

Figure 5.3. Illustration of the Rolling Window Linear Regression algorithm using the sum of one student's steps during 8 days. The x-axis displays the day and the y-axis the number of steps per day. The red and green boxes show a moving 3-day time window at day 3 and day 6, respectively, with the regression coefficients on top

5.3.1.3 Window-Based Interquartile Range (IQR) Algorithm

The aim of this algorithm is to detect recent physical activity behaviour changes in each student individually and independently from their peers. Using the hourly steps, it labels the student's latest significant physical activity differences. Is composed of the first four steps of the U-BEHAVED algorithm explained in Chapter 4 and in [31]. Specifically, as illustrated in Figure 5.4, a time window rolls through the number of steps and time spent doing moderate to vigorous PA (MVPA) per hour (step 1, red box). For each rolling window, the algorithm calculates the mean step number and MVPA per hour (step 2, blue box). Then, it subtracts the present-day hourly steps and MVPA from the hourly mean step number and MVPA of the windows, to generate the hourly differences of step number and MVPA (step 3, yellow box). These differences are box-plotted (step 4, green box) and the IQR limits are used to detect outliers that are labelled as significant behaviour changes (step 5, purple box). It returns a data frame with eleven vectors: the date and time when behaviour changes were detected, the mean hourly step number and MVPA, the difference between the detected behaviour changes and the hourly mean step number and MVPA, the upper and lower IQR limits, and the length of the rolling window used.

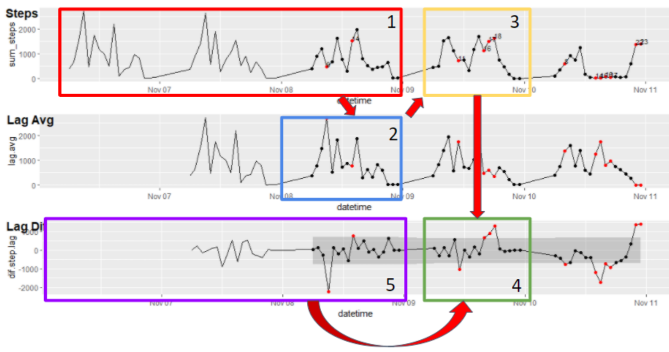

Figure 5.4. Schematic illustration of the Window-Based IQR algorithm using data from one student during one week. The algorithm steps are numbered inside the coloured boxes. The x-axes represent the time in hours. The y-axis display the sum of steps (upper graph), the mean sum of the steps from the 3-day window (middle graph), and the difference between the present day value and the last 3-day window mean value (bottom graph). Red dots represents the detected behaviour changes and the dark grey area represents the IQR limits

5.3.1.4 Multi-Shifted Window-Based IQR Algorithm

The aim of this algorithm is to detect physical activity behaviour changes that are sustained over time. Using the previously detected hourly behaviour changes, it labels as habits the changes that are repeated consecutively. This algorithm is composed of the last step of the U-BEHAVED algorithm explained in Chapter 4 and in [31]. Briefly, as illustrated in Figure 5.5, it scans the upper and lower limits of the previously detected behaviour changes from the Window-Based IQR algorithm through the time series of the hourly step difference (box 1, purple). If the hourly steps difference is outside the limits of the previously detected behaviours (box 2, green, and box 3, yellow), then the behaviour change is sustained. It returns a data frame with three vectors: the sustained behaviour change date, time, and the number of days of uninterrupted behaviour change.

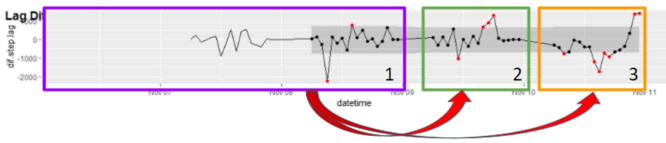

Figure 5.5. Schematic illustration of the Multi-Shift Window-Based IQR algorithm using data from one student during one week. The different algorithm steps are numbered inside the coloured boxes. The x-axis shows the time in hours. The y-axis represents the difference between the present day and the last 3-day window mean step number (same as in Figure 5.4)

5.3.2 Learner Model Attributes

The main topics of iEngage were related to the learning model attributes to capture knowledge and specific patterns of physical activity behaviour. These attributes use outputs from the algorithms described in the previous section. Table 5.1 lists the attributes and maps them to the iEngage topics and to the algorithms. For instance, the first and second attributes of the table make use of particular patterns of physical activity behaviour detected by the Difference Scanner algorithm using physical activity data from activity trackers, as well as each student's daily step and MVPA goal established in the e-learning application.

5.3.3 Personalised Feedback Generation

We used the learner model attributes to generate timely personalised feedback to support the students' learning, motivation, and healthy physical activity behaviour changes.

In detail, after the values of the learning model attributes are calculated, a set of rules triggers specific feedback for each of the learning model attributes. Then, the specific feedback generated is filtered by another set of rules before being delivered to the learner. Table 5.2 maps out which values of the learner model attributes trigger which type of specific feedback, and Table 5.3 displays the rules used to filter the specific feedback generated from Table 5.2.

5.3.3.1 Illustration

To illustrate the personalised feedback each student receives, we sampled data from a real participant in a previous intervention to generate the feedback received on a particular day. In [111] additional information regarding the intervention and the data used is provided. The values of each learner model attribute were calculated by our algorithms using the sampled data to trigger the feedback rules and to build the personalised feedback. This is presented to the participant as a narrative in an educational, constructive and adolescent-friendly way (Figure 5.6).

TABLE 5.1. Map of the attributes to the iEngage topics, the algorithms used, how they are calculated, and their range of values. All values are calculated for the last 3 days (because the algorithms use a rolling window of 3 days).

ID	Attribute	iEngage Topic	Algorithm Used	Description	Range of Values
1	Step Goals Achieved	Recommended Physical Activity	Difference Scanner	Number of days when the step sum $\geq$ daily step goal	From 0 to 3
2	MVPA Goals Achieved	Recommended Physical Activity	Difference Scanner	Number of days when the sum of minutes spent in MVPA per day $>$ daily MVPA goal	From 0 to 3
3	Step Change	Be Physically Active	Difference Scanner	Indicates whether the daily number of steps has increased since the previous module	1, if step number increases or is maintained, 0 otherwise
4	Step Trend	Be Physically Active	Rolling Window Linear Regression	Trend (β_1) of the step number	Can be 0, positive or negative
5	Positive Step Behaviour Changes	Healthy and Sedentary Behaviours	Window Based IQR	Number of days in which positive step behaviour changes are detected	From 0 to 3
6	Negative Step Behaviour Changes	Healthy and Sedentary Behaviours	Window Based IQR	Number of days in which negative step behaviour changes are detected	From 0 to 3
7	Positive Behaviours, Step Increase	Healthy and Sedentary Behaviours	Window-Based IQR	Difference between the step number in positive steps behaviour changes and the mean step number	≥ 0 (0, if no positive behaviour is detected)
8	Negative Behaviours, Step Loss	Healthy and Sedentary Behaviours	Window Based IQR	Difference between the step number in negative step behaviour changes and the mean step number	≥ 0 (0, if no negative behaviour is detected)
9	Daily MVPA Change	Physical Activity Intensity Levels	Difference Scanner	Indicates whether the daily minutes spent in MVPA has increased since the previous module	1 if minutes increase or are maintained, 0 otherwise
10	Daily MVPA Trend	Physical Activity Intensity Levels	Rolling Window Linear Regression	Trend (β_1) of time spent in MVPA	Can be 0, positive or negative
11	Positive MVPA Behaviour Changes	Physical Activity Intensity Levels	Window-Based IQR	Number of days in which positive MVPA behaviour changes are detected	From 0 to 3
12	Negative MVPA Behaviour Changes	Physical Activity Intensity Levels	Window-Based IQR	Number of days in which negative MVPA behaviour changes are detected	From 0 to 3
13	Positive Behaviours, MVPA Increase	Physical Activity Intensity Levels	Window-Based IQR	Difference between the minutes spent in MVPA during positive MVPA behaviour changes and the mean minutes spent in MVPA	≥ 0 (0, if no positive behaviour is detected)
14	Negative Behaviours, MVPA Loss	Physical Activity Intensity Levels	Window-Based IQR	Difference between the minutes spent in MVPA during negative MVPA behaviour changes and the mean minutes spent in MVPA	≥ 0 (0, if no negative behaviour is detected)
15	Longest Positive Habit	Healthy and Sedentary Behaviours	Multi-Shifted Window Based IQR	Positive step behaviour change that has been maintained for the highest number of days	Initial time (day and hour) and number of days maintained
16	Opportunity to Increase Steps	Recommended Physical Activity	Difference Scanner	Day of maximum difference in step number between the student and the cohort	Time (hour) of the day when the student can continue increasing their physical activity
17	Follow-up	Recommended Physical Activity	Difference Scanner	Number of days the adolescent did more steps than recommended (Attribute ID 16)	From 0 to 3

TABLE 5.2. Mapping of attributes, their values, and the general and specific feedback.

Feedback-triggering attribute	**General feedback**	**Attribute rules and specific feedback**
Step goals are achieved (ID.1 value = 3)	A-congratulate and encourage	You have achieved your step goal each day in the last 3 days
	B-find positive changes to reinforce	If ID.3 > 0, you have increased the daily steps since the last module If ID.3 = 0, you did fewer steps than before the last module
Step goals are mostly achieved (ID.1 value = 1 or 2)	C-congratulate and encourage	You almost achieved your steps goals every day in the last 3 days
	B-find positive changes to reinforce	If ID.3 > 0, you increased the daily steps since the last module. If ID.3 = 0, you did fewer steps than before the last module.
	D-find positive trends and reinforce them	If ID.4 = 1, your step trend was positive in the last 3 days
	E-identify positive behaviours to reinforce them	If ID.5 = 3, you had positive step behaviour changes in the last 3 days. These amount to ID.7 additional steps over the last 3 days. If ID.5 = 1 or 2, you had positive step behaviour changes in ID.5 of the last 3 days. These amount to ID.7 additional steps over the last 3 days.
	F-congratulate for achieving a positive habit	If ID.15 present, congratulate for creating a positive habit
	G-promote following opportunity, suggestions	If ID.17 = 3, 2, or 1, congratulate for following the advice in the last 3 days and keep up the good work If ID.17 = 1, encourage to do it more often If ID.17 = 0, highlight the benefit of following advice
	H-suggest opportunities for more physical activity during the day	Suggest ID.16 as a possible time window for being more active, compared with peers
Step goals are not achieved (ID.1 value = 0)	P -state non-achievement	You did not achieve the step goals on any day in the last 3 days
	E-identify positive behaviours to reinforce them	If ID.5 = 3, you showed positive step behaviour changes in the last 3 days. These amounted to ID.7 additional steps over the last 3 days If ID.5 = 1 or 2, you showed positive steps behaviour changes in some of the last 3 days. These amount to ID.7 additional steps over the last 3 days
	F-congratulate for achieving a positive habit	If ID.15 present, congratulate for creating a positive habit
	I-find negative behaviours to fix	If ID.6 $\geq$ 2, there were ID.6 days in which you did fewer steps compared with a typical day, losing ID.8 steps
	G-promote following opportunity, suggestions	If ID.17 = 3, 2, or 1, congratulate for following the advice in the last 3 days and keep up the good work If ID.17 = 1, encourage to do it more often If ID.17 = 0, highlight the benefit of following advice
	H-suggest opportunities for more physical activity during the day	Suggest ID.16 as a possible time window for being more active, compared with peers
MVPA goals are achieved (ID.2 value = 3)	J-congratulate and encourage	You achieved your MVPA goal each day in the last 3 days
	K-find positive changes to reinforce	If ID.9 $>$ 0, you increased your daily MVPA since the last module If ID.9 = 0, you did less MVPA than before the last module
MVPA goals are mostly achieved (ID.2 value = 1 or 2)	L-congratulate and encourage	You almost achieved the MVPA goal ID.2 days out of 3.
	K-find positive changes to reinforce	If ID.9 $>$ 0, you increased your daily MVPA since the last module If ID.9 = 0, you did less MVPA than before the last module
	M-find positive trends and reinforce them	If ID.10 = 1, your MVPA trend was positive in the last 3 days
	N-look for positive behaviours to reinforce them	If ID.11 = 3, you did positive MVPA behaviour changes in the last 3 days. These amounted ID.13 additional minutes of MVPA over the last 3 days If ID.11 = 1 or 2, you did positive MVPA behaviour changes in most of the last 3 days. These amounted ID.13 additional minutes of MVPA over the last 3 days
MVPA goals are not achieved (ID.1 value = 0)	Q -state non-achievement	You did not achieve your MVPA goals in any of the last 3 days
	N-look for positive behaviours to reinforce them	If ID.11 = 3, you did positive MVPA behaviour changes in the last 3 days. These amounted ID.13 additional minutes of MVPA over the last 3 days If ID.11 = 1 or 2, you did positive MVPA behaviour changes most of the last 3 days. These amounted ID.13 additional minutes of MVPA over the last 3 days
	O-find negative behaviours to fix	If ID.12 is $\geq$ 2, during the last 3 days you spent fewer minutes doing MVPA compared with a typical day, losing ID.14 minutes of MVPA

Note: The ID and values in the feedback-triggering attribute column are from Table 5.1. The letter placed at the beginning of each general feedback helps to identify them in Table 5.3.

TABLE 5.3. Feedback-trigger rules according to the combination of attributes values

	ID.2=3	ID.2=2 or 1	ID.2=0
ID.1=3	Congratulate for goals (A,J) Reinforce positive changes (B,K)	Congratulate for goals (A,L) Reinforce positive changes (B,K) Reinforce positive trends (M) Reinforce positive behaviours (N)	Congratulate for goals (A) Reinforce positive changes (B) Reinforce positive behaviours (N) Negative behaviour to fix (O)
ID.1=2 or 1	Congratulate for goals (C,J) Reinforce positive changes (B,K) Reinforce positive trend (D) Reinforce positive behaviour (E) Congratulate for habit (F) Promote suggestions (G) Suggest opportunity (H)	Congratulate for goals (C,L) Reinforce positive changes (B,K) Reinforce positive trend (D,M) Reinforce positive behaviour (E,N) Congratulate for habit (F) Promote suggestions (G) Suggest opportunity (H)	Congratulate for goals (C) Reinforce positive changes (B) Reinforce positive trend (D) Reinforce positive behaviour (E,N) Congratulate for habit (F) Promote suggestions (G) Suggest opportunity (H)
ID.1=0	Congratulate for goal (J) Reinforce positive changes (K) Reinforce positive behaviour (E) Negative behaviour to fix (I) Congratulate for habit (F) Promote suggestions (G) Suggest opportunity (H)	Congratulate for goal (L) Reinforce positive changes (K) Reinforce positive trend (M) Reinforce positive behaviour (E,N) Congratulate for habit (F) Negative behaviour to fix (I) Promote suggestions (G) Suggest opportunity (H)	Goal not achieved (P,Q) Reinforce positive behaviour (E,N) Negative behaviour to fix (I,O) Congratulate for habit (F) Promote suggestions (G) Suggest opportunity (H)

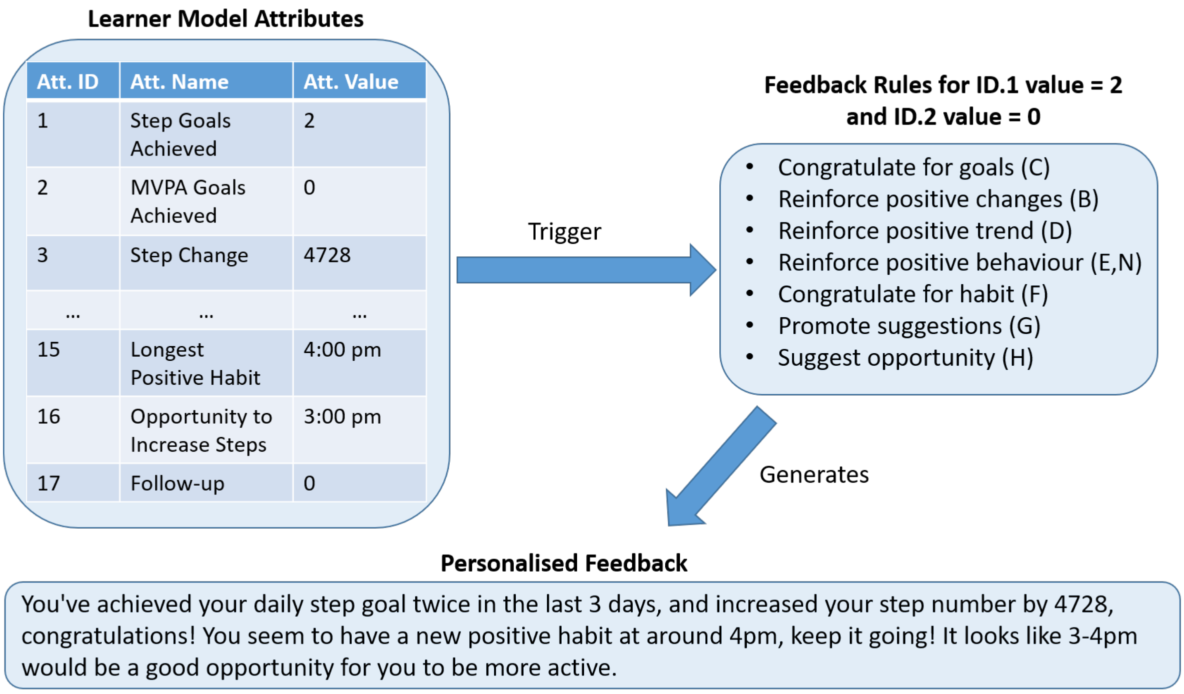

Figure 5.6. Illustration of the resulting personalised feedback using data of a real participant for a particular day

5.4 Validation

We recruited four health sciences, exercise physiology, or physical education experts with experience working with adolescents. We asked them to evaluate via an online survey whether the automated personalised feedback was valuable (across 3 dimensions: helpful, important, and timely) for supporting adolescents who learn how to modify their physical activity behaviour. This online survey, presented real physical activity data from three adolescents (from a previous iEngage education session [111]), and nee. The experts were asked to evaluate and comment on the feedback automatically generated by our algorithms on three days. This survey was approved by The University of Sydney Human Research Ethics Committee (2021/815).

5.4.0.1 Study Results

Each expert evaluated 44 automated feedbacks using a 5-point Likert scale. The resultant 176 answers (4 experts * 44 feedbacks) positively evaluated the automated personalised feedbacks as helpful (mean=88%, s.d.= 11%), important (mean=85%, s.d.= 7%), and timely (mean=85%, s.d.= 8%).

Experts were also asked to provide free text comments about the automated feedbacks. The most frequent response about what they liked was that the feedback reinforced positive behaviour changes. This was expressed by one of the participants as follows: "*[I liked that the feedback was] positive and focused on the good behaviours rather than on the negative aspects*". The most frequent response about what they did not like in the feedback was that they were more focused on explaining the students' behaviours rather than motivating them to generate new behaviour changes. For instance, one of the participants wrote: "*[the automated feedback is] missing encouragement messages to do better the next 3 days*". The most frequent response about what was missing was to incorporate the student's opinion in the feedback loop, as commented by one participant: "*using motivational interviewing might be a better solution than giving them a suggested time to increase physical activity, [it would] allow to make their own suggestions that would work with them*".

Experts found that the automated feedback was good support for the students' learning process; however, they highlighted the importance of using an adolescent-friendly language in a live scenario that would need to be tailored depending on the health intervention population demographics. They also suggested students should be incorporated into the feedback loop to increase their motivation and propose more relevant actions. The study suggests that valuable personalised formative feedback can be automatically generated using a combination of knowledge-specific attributes mapped with the health education application, and physical activity behaviour-specific attributes extracted from the activity tracker.

5.5 Conclusion

We explored the problem of adding the physical activity accelerometer time series data into a learner model for a health education system by providing algorithms that (1) detect physical activity behaviour changes relevant for knowledge acquisition by students, (2) build relevant learner model attributes, and (3) provide timely personalised feedback. Window-based algorithms enable to monitor individual's physical activity behaviour changes over time through relevant learning-related PA behaviour attributes mapped to the learning contents. Although the actual feedback was not implemented live in our system, we provide examples of what type of feedback can be generated using these algorithms and real data from iEngage, a digital health education programme that teaches health knowledge and skills to promote physical activity behaviour change in adolescents. Health education experts commented that the personalised feedbacks generated by our algorithms were useful, important, and timely for supporting healthier PA behaviour changes.

Further research on the creation and evaluation of personalised messages would be interesting to study which aspects of the messages are most likely to affect behaviour change and how these elements may be leveraged with more appealing messages to participants. Likewise, the temporality of the messages can be further explored, for instance, by examining the most effective time to provide feedback and the influence of reminders and follow-up messages responding to their behaviour changes.

5.6 Chapter Contributions and Relationship to Thesis Research Questions

This chapter contribute and is related to the second research question:

RQ2 Can data mining techniques help to detect behaviour changes and generate personalised feedback to support learning?

(1) We presented a set of data mining techniques that use participants' physical activity data to detect behaviour changes as the intervention unfolds. The live monitoring of physical activity changes is relevant to any health education intervention because it provides objective, timely and key information on how they learn healthy behaviours. This information is essential to understand whether the participant is learning and allows providing feedback based on their learning.

(2) We described how we model the detected behaviours and integrate them into a learning model. The integration of the physical dimension of students learning into a learning model is a pivotal step for the generation of automated feedback based on their behaviour. The model is valuable because it allows the automated feedback generation and trigger on the basis of the participant's learning state.

(3) We described how to generate timely personalised feedback for each participant based on their physical activity behaviour changes integrated in the learner model. The automated feedback generated is valuable because it demonstrates how physical activity data can be used to support the participants' learning. Experts found that the generated personalised feedback was useful, important, and timely for supporting each participant's learning.

CHAPTER 6

Conclusion

In this thesis, we explored how data mining methods can be applied to physical activity sensor data to assess and personalise health education programmes. To address the two research questions, each chapter focused on one of four objectives. The first objective was to review the current state of the art in data mining methods for detecting physical activity behaviour changes in health education. The second objective was to propose a data mining method to gain some insights into the impact of a health education intervention on physical activity. The third objective was to develop a data mining technique to detect behavioural changes from streaming data. The fourth objective was to produce personalised feedback based on each participant's behaviour changes to support their learning. This final chapter concludes with a review of the work done for this thesis, explaining how the research questions were addressed, discussing key contributions, and providing indications of potential future work.

6.1 Revisiting the Research Questions

The two research questions of this thesis were:

- **RQ1** Can data mining techniques and models help to assess the effectiveness of physical activity interventions using activity sensor data?
- **RQ2** Can data mining techniques help to detect behaviour changes and generate personalised feedback to support learning?

RQ1 was addressed in Chapter 2 and 3. This required the development of a data mining technique to extract behaviours from physical activity data in order to assess the health education impact. RQ2 was addressed in Chapters 2, 4 and 5. This required using data mining to detect and model behavioural changes to support learning. Figure 1.1 provides an overview of this thesis, showing the initial research questions posed in Chapter 1, how the different chapters relate to these questions, and how chapters are linked together.

In the next section, the contributions and future work of each chapter are discussed.

6.2 Contributions and Future Work

Chapter 2 is a systematic review of the literature on the current state of physical activity sensor data mining techniques in health education to detect behaviour changes in which we discussed common challenges and opportunities. This provided the foundation for addressing RQ1 and RQ2 because it summarised the current knowledge on data mining methods that have been developed to detect behavioural changes in health education and that can be used for intervention assessment (RQ1) and for supporting knowledge acquisition (RQ2).

The first contribution is the identification and description of the current data mining methods used for physical activity sensor data collected in the framework of health education in order to detect behavioural changes. This review analysed the current methods by dividing them into different sub-processes: data capture, data pre-processing, data transformation, data mining, and interpretation. This can help health and computer scientists to understand in detail how these methods work and how the detected behaviour changes can be used to optimise health education. The second contribution was the collection of evidence to highlight the current challenges and opportunities. This helped to identify literature gaps and contradictory, or inconclusive findings that justified the present work and will guide future work on the development of the next generation of data mining methods.

This systematic review showed how data mining can be used to extract behaviour changes from sensor data mining to be then analysed by experts with the aim of improving health education. In the future, it could be useful to periodically update this systematic review because studies on this topic started to be published only a decade ago and have been increasing over time.

Moreover, this literature review could be complemented or extended. For instance, it would be interesting to review other sub-processes or topics related to data mining methods outside this thesis scope, such as ethics, privacy, and security, because they can influence the design of methods.

Chapter 3 presented a novel data mining method to analyse changes in physical activity behaviour in order to understand at a finer level the impact of a health education intervention. This chapter addressed RQ1 because it focused on how data mining techniques and models can help to assess the effectiveness of physical activity interventions using activity sensor data.

The first contribution of this chapter was the method to extract daily behaviour changes from physical activity data to meaningfully assess the health intervention impact relative to the recommended daily goals. Specifically, the method builds clusters of daily intensity, duration and frequency of physical activity bouts extracted from physical activity trackers. Then, these clusters are analysed to identify changes in the participants' behaviour relative to the daily recommended guidelines. The presented method is valuable because it characterises physical activity behaviour changes that are latent in the current analyses. Physical activity intensity, duration and frequency are standard metrics used in health recommendations and guidelines. The flexibility of our method allows analysing behaviour changes using multiple or different criteria. For instance, children's behaviour changes can be analysed using the World Health Organisation physical activity recommendations for children or local physical activity guidelines for children. Similarly, behavioural changes in adults and older adults can be analysed relative to the specific guidelines for each age range and gender.

The second contribution is the use of this method to focus the analysis of physical activity behaviour changes in function of the part of the day in which they occurred. Clusters of physical activity intensity, duration and frequency were created and analysed to assess changes in the participants' behaviour at different times of the day. This is valuable for interventions because it allows understanding specifically at what time of the day the behaviour changes occur. The analysis of behaviour changes by day and at different times of the day might help to refine health guidelines by identifying times of the day more suitable for physical activity.

The presented technique is flexible and can integrate a larger number of physical activity behaviour characteristics and use other time aggregation types. In the future would be interesting to explore what other analysis types can be generated with the proposed method. For instance, using data aggregated at the week level could allow investigating behaviour changes that occur during weekdays or during weekends. Our proposed method could also be extended by exploring the integration of additional information on the participants' behaviours that can be extracted from physical activity sensors. This would require to use, adapt or create other data mining techniques, and the resulting clusters would need new analysis and interpretation approaches. For instance, it would be interesting to integrate other human behaviours that may influence physical activity behaviours, such as sleep. For this, sleep detection algorithms will be required to extract sleep features, such as time spent sleeping, when participants go to bed, and when they wake up, and also sleeping behaviours changes (e.g. bedtime changes). By combining physical activity and sleep features, the new clusters would provide information on the intervention effects on the participants' lives, moving towards a lifestyle assessment of health education interventions. Moreover, it would be interesting to integrate the participants' physical activity routines, such as daily or weekly routines, using sequential-pattern mining algorithms. By combining physical activity behaviour changes and physical activity routines, the new clusters could help to study how behaviour changes affect routines, for instance, the types of physical activity routine changes, when they occur, and their time length and frequency. Similarly, it would be interesting to integrate in the analysis the physical activity type performed by participants. Physical activity recognition techniques could be used to detect different activities, such as sitting, walking or jogging. Supervised data mining techniques would need to be trained to classify different physical activity types. By combining information about physical activity behaviour changes and the physical activity type performed, the new clusters could be used to analyse what activity types changed during the intervention and how they are related to physical activity behaviour changes. Finally, although the k-means algorithm is used in our proposed method because it produced cohesive clusters and is simple to use and interpret, future research can explore the employment of new clustering methods to examine for additional physical activity behaviour patterns and enrich our suggested analysis.

Chapter 4 presented U-BEHAVED, a novel unsupervised data mining algorithm that detects significant changes in physical activity behaviours and habits as soon as they appear. This topic is related to RQ2 because the algorithm detects behavioural changes individually for each participant using data from their physical activity trackers. In Chapter 5, U-BEHAVED is part of the set of algorithms used to detect each participant's behaviour changes. These data are integrated into a user model to generate timely personalised feedback to support their learning.

The first contribution of the algorithm was the early detection of recent behavioural changes from streaming data. This information is valuable for any health intervention because it allows knowing in real time when changes are generated and whether they are healthy. Early detection gives the possibility to generate timely and tailored responses to each participant's needs, such as adjusting their programme content or adapting their physical activity goals. The second contribution of the algorithm was the early detection of behavioural changes sustained over time (habits). This information is valuable for health education because it says whether the behaviours learned by participants are maintained over time. This would allow reinforcing healthy habits once acquired and recovering healthy habits that had been acquired but were no longer performed.

The U-BEHAVED algorithm presents four main advantages. First, it is applicable to any health intervention that uses physical activity trackers because the participants' number of steps per hour is its input, a common output of most activity trackers. Second, this is an unsupervised algorithm and therefore, it does not need to be trained on each participant's physical activity behaviour to detect behaviour changes. The algorithm learns from each participant's behaviour to detect changes. Third, it can be customised to detect behaviour changes at any programme content delivery pace because it has a configurable rolling window that can be set to any appropriate length. Fourth, it can detect behaviour changes close to real-time when data from the participants' physical activity trackers are continuously streamed and cached because the algorithm rolling window would slide to the last time the data are refreshed.

In the future, behavioural changes detected during the intervention might be used as a measure of the participants' learning, thus providing relevant information on the intervention impact, especially when analysed at different levels (individuals or cohort), during or after the intervention. The present work could be extended by exploring the creation of semi-supervised data mining techniques to detect behaviour changes. For instance, upon detecting a behaviour change, our unsupervised algorithm could ask the participant whether the behaviour change actually occurred (or not) in order to label the detection as correct (or incorrect). These labels could serve to train a supervised algorithm (e.g. neural networks) to detect behaviour changes. Although we confirmed that the algorithm can accurately detect behaviour changes, it would be interesting to compare the detection accuracy of the unsupervised (current) and semi-supervised (future) data mining methods.

Also in future work, our algorithm could be modified to handle missing data. The current U-BEHAVED algorithm needs as input a continuous time series of physical activity data to detect behavioural changes, and cannot use incomplete or partial records. However, to improve it robustness, it could be modified to detect when participants stop wearing the tracker, thus skipping these events. As U-BEHAVED bases its detection on past behaviours, it would be interesting to check whether this modification would decrease or increase the detection accuracy, because it may have difficulties in identifying which behaviour is habitual and which is a significant change.

Lastly, as explained in section 4.6, U-BEHAVED is flexible enough to accept other aggregation levels of the participants' number of steps as input. Exploring new aggregation levels would allow detecting other behavioural changes. For instance, when data are aggregated to higher levels, such as per day or week, it might detect long-term changes.

Chapter 5 propose a model for supporting knowledge acquisition in physical learning environments based on the participants' physical activity behaviours. This topic is related to RQ2 because it describes a range of data mining techniques to detect behavioural changes from the participants' physical activity data. These findings were modelled into a learner model to create personalised feedback to support the participants' learning. A narrative summative and informative feedback is generated and triggered in response to specific behaviours. The feedback seeks to guide the participants' learning process, helping them to think about their physical activity behaviours and encouraging healthy behaviours.

The first contribution of this work was the creation of a set of algorithms that detect behaviour changes as the intervention unfolds using the participants' physical activity data. The real-time detection of physical activity changes is relevant for any health education intervention because it provides objective, timely, and crucial information on how participants learn healthy behaviours, which can be used to support their learning. The second contribution was to demonstrate how to model the participants' behaviours and how to design a learning model to integrate them. This is valuable for health education because it combines the physical dimension of learning in a learner model. This allows assessing how participants learn and also generating and triggering feedback based on their behaviours. Finally, the third contribution was to show how timely personalised feedback can be generated. This is relevant because it demonstrates how physical activity data can be used to produce feedback with the aim of supporting the participants' learning. Experts validated our personalised feedback by saying that it was helpful, important and timely for supporting the participants' learning.

In the future, new data mining techniques may extract more participants' behaviours from activity tracker data that could be added to the learner model. This may give a broader view of how participants' behaviours change during knowledge acquisition, and may allow assessing these behaviours to further improve personalisation. For instance, the time spent at different physical activity intensities could be extracted and analysed to understand how this changes as participants learn. Moreover, by relating these data to the participants' behavioural changes, more detailed feedback on their physical activity habits could be provided. Also, would be interesting to investigate other ways of delivering personalised feedback. In this chapter, narrative feedbacks were used, but vibrotactile or sound feedback could be tested. The latter may be suitable delivery options for people with special needs, for instance, participants who cannot read. In addition, it would be interesting to explore the possibility of opening the learning model to participants. This could make participants think about their own behaviours and self-regulate their learning, which could boost knowledge acquisition. In addition, the learning model could be expanded or modified to focus on supporting health education teachers. Teachers could monitor the behaviours of the whole cohort and follow each participant's progress. This could give to teachers timely key information on each participant's learning that may be used to provide person-to-person detailed personalised feedback or to modify the programme to fit the cohort's needs. Finally, a user study needs to be conducted to evaluate the impact of our automated feedback on learning. We note that this study was part of the original thesis plan, but unfortunately had to be abandoned due to COVID-19 lockdown restrictions.

6.3 Summary and Final Reflections

This thesis presented data mining methods to extract changes in physical activity behaviour from physical activity tracker data in order to create elaborated intervention evaluations and to generate timely personalised feedback during the intervention.

Extracting physical activity patterns from physical activity trackers is valuable for health education because it provides detailed insights into the participants' physical activity learning dimension, paving the way to thorough assessments and personalisation. This information can also be valuable for areas related to health science, such as exercise physiology and sports science, because by providing details about how the participants' PA occurs and changes, it helps to carry out comprehensive analyses on human physical activity behaviours.

Sensor technology continues to progress and is becoming largely adopted in health education. This thesis can serve as a first guide on how to develop novel data mining methods for extracting other physical activity patterns from sensor data. Furthermore, it can be used as a reference on how to develop insightful physical activity behavioural features and as a blueprint on how these features can be modelled and analysed to evaluate interventions and support learning. Finally, we hope this thesis will serve as a stepping stone toward the many promising avenues for future research we have discussed.

Bibliography

[1] Activinsights Ltd., *GENEActiv Original - Wrist-Worn Actigraphy Device | GENEActiv Accelerometers*, 2017.

[2] A. Adadi and M. Berrada, 'Peeking inside the black-box: A survey on explainable artificial intelligence (xai),' *IEEE access*, vol. 6, pp. 52 138–52 160, 2018.

[3] M. A. Adams, W. D. Johnson and C. Tudor-Locke, 'Steps/day translation of the moderate-to-vigorous physical activity guideline for children and adolescents,' *International Journal of Behavioral Nutrition and Physical Activity*, vol. 10, no. 1, p. 49, 2013.

[4] A. Aguilera, C. A. Figueroa, R. Hernandez-Ramos *et al.*, 'mHealth app using machine learning to increase physical activity in diabetes and depression: clinical trial protocol for the DIAMANTE Study,' *BMJ Open*, vol. 10, no. 8, e034723, Aug. 2020.

[5] M. C. Angelides, L. A. C. Wilson and P. L. B. Echeverría, 'Wearable data analysis, visualisation and recommendations on the go using android middleware,' English, *Multimedia Tools and Applications*, vol. 77, no. 20, pp. 26 397–26 448, Oct. 2018.

[6] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou and Y. Amirat, 'Physical human activity recognition using wearable sensors,' *Sensors*, vol. 15, no. 12, pp. 31 314–31 338, 2015.

[7] R. C. Bailey, J. Olson, S. L. Pepper, J. Porszasz, T. J. Barstow and D. M. Cooper, 'The level and tempo of children's physical activities: An observational study,' *Medicine and Science in Sports and Exercise*, vol. 27, no. 7, pp. 1033–1041, 1995.

[8] R. S. Baker and K. Yacef, 'The State of Educational Data Mining in 2009 : A Review and Future Visions,' *Journal of Educational Data Mining*, vol. 1, no. 1, pp. 3–17, 2009.

[9] L. Bao and S. S. Intille, 'Activity recognition from user-annotated acceleration data,' in *International Conference on Pervasive Computing*, Springer, 2004, pp. 1–17.

[10] G. Baquet, G. Stratton, E. Van Praagh and S. Berthoin, 'Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: A methodological issue,' *Preventive Medicine*, vol. 44, no. 2, pp. 143–147, 2007.

[11] T. Baranowski and C. de Moor, 'How many days was that? intra-individual variability and physical activity assessment,' *Research Quarterly for Exercise and Sport*, vol. 71, no. sup2, pp. 74–78, 2000.

[12] T. Batool, Y. Vanrompay, A. Neven, D. Janssens and G. Wets, 'CTASS: an intelligent framework for personalized travel behaviour advice to cardiac patients,' *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–13, May 2018.

[13] I. Ben-Gal, 'Outlier detection,' in *Data mining and knowledge discovery handbook*, Springer, 2005, pp. 131–146.

[14] K. Borodulin, K. R. Evenson, F. Wen, A. H. Herring and A. Benson, 'Physical activity patterns during pregnancy,' *Medicine and science in sports and exercise*, vol. 40, no. 11, p. 1901, 2008.

[15] J. E. Brinton, M. D. Keating, A. M. Ortiz, K. R. Evenson and R. D. Furberg, 'Establishing linkages between distributed survey responses and consumer wearable device datasets: A pilot protocol,' *JMIR research protocols*, vol. 6, no. 4, e6513, 2017.

[16] F. C. Bull, S. S. Al-Ansari, S. Biddle *et al.*, 'World health organization 2020 guidelines on physical activity and sedentary behaviour,' *British journal of sports medicine*, vol. 54, no. 24, pp. 1451–1462, 2020.

[17] C. Caillaud, S. Ledger, C. Diaz, G. Clerc, O. Galy and K. Yacef, 'Iengage: A digital health education program designed to enhance physical activity in young adolescents,' *PREPRINT (Version 2) available at Research Square [https://doi.org/10.21203/rs.3.rs-1456896/v2]*, August 2022.

[18] H. K. Chen, F. H. Chen and S. F. Lin, 'An ai-based exercise prescription recommendation system,' *Applied Sciences (Switzerland)*, vol. 11, no. 6, pp. 1–11, 2021.

[19] K. L. Cheung, D. Durusu, X. Sui and H. de Vries, 'How recommender systems could support and enhance computer-tailored digital health programs: A scoping review,' *Digital Health*, vol. 5, pp. 1–19, 2019.

[20] P. Chillón, J. Panter, K. Corder, A. Jones and E. Van Sluijs, 'A longitudinal study of the distance that young people walk to school,' *Health Place*, vol. 31, pp. 133–137, 2015.

[21] D. Chousiadas, A. Menychtas, P. Tsanakas and I. Maglogiannis, 'Advancing Quantified-Self Applications Utilizing Visual Data Analytics and the Internet of Things,' in ser. IFIP Advances in Information and Communication Technology, L. Iliadis, I. Maglogiannis and H. Papadopoulos, Eds., vol. 436, Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, pp. 263–274.

[22] A. R. Cooper, A. Goodman, A. S. Page *et al.*, 'Objectively measured physical activity and sedentary time in youth: The International children's accelerometry database (ICAD),' *International Journal of Behavioral Nutrition and Physical Activity*, vol. 12, no. 1, pp. 1–10, 2015.

[23] P. N. Dawadi, D. J. Cook and M. Schmitter-Edgecombe, 'Modeling patterns of activities using activity curves,' *Pervasive and mobile computing*, vol. 28, pp. 51–68, 2016.

[24] L. De Moura Carvalho, V. Furtado, J. E. de Vasconcelos Filho and C. M. G. F. Lamboglia, 'Using Machine Learning for Evaluating the Quality of Exercises in a Mobile Exergame for Tackling Obesity in Children,' in *Proceedings of SAI Intelligent Systems Conference*, Springer, 2016, pp. 373–390.

[25] Y. Demetriou and O. Höner, 'Physical activity interventions in the school setting: A systematic review,' *Psychology of Sport and Exercise*, vol. 13, no. 2, pp. 186–196, 2012.

[26] A. Dias Pereira dos Santos, K. Yacef and R. Martinez-Maldonado, 'Forró trainer: Automated feedback for partner dance learning,' in *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*, ser. UMAP '17, Bratislava, Slovakia: ACM, 2017, pp. 103–104.

[27] A. Dias Pereira dos Santos, K. Yacef and R. Martinez-Maldonado, 'Let's dance: How to build a user model for dance students using wearable technology,' in *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, ser. UMAP '17, Bratislava, Slovakia: Association for Computing Machinery, 2017, pp. 183–191.

[28] C. Diaz, 'A clustering approach for modelling and analysing changes in physical activity behaviours from accelerometers,' *Open Science Framework*, 2022, Available at https://osf.io/f5uw6/.

[29] C. Diaz, 'Unsupervised early detection of physical activity behaviour changes from wearable accelerometer data,' *Open Science Framework*, 2023, Available at https://osf.io/q9yw2/.

[30] C. Diaz, C. Caillaud and K. Yacef, 'Mining sensor data to assess changes in physical activity behaviors in health interventions: Systematic review,' *JMIR Medical Informatics*, 2022.

[31] C. Diaz, C. Caillaud and K. Yacef, 'Unsupervised early detection of physical activity behaviour changes from wearable accelerometer data,' *Sensors*, vol. 22, no. 21, 2022.

[32] C. Diaz and L. Ha, 'Physical activity accelerometer data preprocessing,' *Open Science Framework*, 2022, Available at https://osf.io/ng3tk/.

[33] C. Diaz and K. Yacef, 'Detecting behaviour changes in accelerometer data,' in *Knowledge Discovery in Healthcare Data (KDH@IJCAI-ECAI Stockholm)*, vol. 2148, 2018, pp. 21–26.

[34] C. Díaz, O. Galy, C. Caillaud and K. Yacef, 'A clustering approach for modeling and analyzing changes in physical activity behaviors from accelerometers,' *IEEE Access*, vol. 8, pp. 224 123–224 134, 2020.

[35] C. Díaz, O. Galy, C. Caillaud and K. Yacef, 'Modelling physical activity behaviour changes for personalised feedback in a health education application,' *Proceedings of the 30th International Conference on Computers in Education*, pp. 369–374, 2022.

[36] T. B. Dijkhuis, F. J. Blaauw, M. W. van Ittersum, H. Velthuijsen and M. Aiello, 'Personalized physical activity coaching: A machine learning approach,' *Sensors (Switzerland)*, vol. 18, no. 2, pp. 1–15, 2018.

[37] C. Dobbins and R. Rawassizadeh, 'Clustering of physical activities for quantified self and mhealth applications,' in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 1423–1428.

[38] J. Dollman, K. Norton and L. Norton, 'Evidence for secular trends in children's physical activity behaviour,' *British Journal of Sports Medicine*, vol. 39, no. 12, pp. 892–897, 2005.

[39] V. Echeverria, R. Martinez-Maldonado, T. Power, C. Hayes and S. B. Shum, 'Where is the nurse? towards automatically visualising meaningful team movement in healthcare education,' in *Artificial Intelligence in Education*, C. Penstein Rosé, R. Martínez-Maldonado, H. U. Hoppe *et al.*, Eds., Cham: Springer International Publishing, 2018, pp. 74–78.

[40] D. W. Esliger, A. V. Rowlands, T. L. Hurst, M. Catt, P. Murray and R. G. Eston, 'Validation of the GENEA accelerometer,' *Medicine and Science in Sports and Exercise*, vol. 43, no. 6, pp. 1085–1093, 2011.

[41] M. Ester, H.-P. Kriegel, J. Sander and X. Xu, 'A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,' in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, 1996.

[42] Z. Fang and M. J. Langford, *Package ' GENEAread '*, 2013.

[43] U. Fayyad, G. Piatetsky-Shapiro and P. Smyth, 'The kdd process for extracting useful knowledge from volumes of data,' *Communications of the ACM*, vol. 39, no. 11, pp. 27–34, 1996.

[44] Fitbit, *Google*, URL: https://www.fitbit.com/, [Accesed: 2-October-2022].

[45] E. M. Forman, S. G. Kerrigan, M. L. Butryn *et al.*, 'Can the artificial intelligence technique of reinforcement learning use continuously-monitored digital data to optimize treatment for weight loss?' English, *Journal of Behavioral Medicine*, pp. 1–15, Aug. 2018.

[46] Y. Fukuoka, M. Zhou, E. Vittinghoff, W. Haskell, K. Goldberg and A. Aswani, 'Objectively measured baseline physical activity patterns in women in the mped trial: Cluster analysis,' *Journal of Medical Internet Research*, vol. 20, no. 2, pp. 1–12, 2018.

[47] O. Galy, K. Yacef and C. Caillaud, 'Improving pacific adolescents' physical activity toward international recommendations: Exploratory study of a digital education app coupled with activity trackers,' *JMIR mHealth and uHealth*, vol. 7, no. 12, pp. 1–11, 2019.

[48] F. Gasparetti, L. M. Aiello and D. Quercia, 'Personalized weight loss strategies by mining activity tracker data,' *User Modeling and User-Adapted Interaction*, vol. 30, no. 3, pp. 447–476, Sep. 2020.

[49] C. Gjestvang, T. Stensrud, B. H. Hansen, E. Kolle and L. A. Haakstad, 'Are fitness club members likely to meet the current physical activity recommendations?' *Translational Sports Medicine*, vol. 3, no. 2, pp. 75–83, 2020.

[50] B. Gupta, M. Tiwari and S. Singh Lamba, 'Visibility improvement and mass segmentation of mammogram images using quantile separated histogram equalisation with local contrast enhancement,' *CAAI Transactions on Intelligence Technology*, vol. 4, no. 2, pp. 73–79, 2019.

[51] L. Ha, C. E. Wakefield, D. Mizrahi *et al.*, 'A digital educational intervention with wearable activity trackers to support health behaviors among childhood cancer survivors: Pilot feasibility and acceptability study,' *JMIR Cancer*, vol. 8, no. 3, e38367, Aug. 2022.

[52] W. L. Haskell, I.-M. Lee, R. R. Pate *et al.*, 'Physical activity and public health: Updated recommendation for adults from the american college of sports medicine and the american heart association,' *Medicine & science in sports & exercise*, vol. 39, no. 8, pp. 1423–1434, 2007.

[53] H. Hermens, H. op den Akker, M. Tabak, J. Wijsman and M. Vollenbroek, 'Personalized Coaching Systems to support healthy behavior in people with chronic conditions.,' *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology*, vol. 24, no. 6, pp. 815–26, Dec. 2014.

[54] L. Jaeschke, A. Steinbrecher, S. Jeran, S. Konigorski and T. Pischon, 'Variability and reliability study of overall physical activity and activity intensity levels using 24 h-accelerometry-assessed data,' *BMC public health*, vol. 18, no. 1, pp. 1–10, 2018.

[55] I. Janssen and A. G. LeBlanc, 'Systematic review of the health benefits of physical activity and fitness in school-aged children and youth,' *International Journal of Behavioral Nutrition and Physical Activity*, vol. 7, no. 1, p. 40, 2010.

[56] I. Jolliffe, 'Principal Component Analysis,' in *International Encyclopedia of Statistical Science*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 1094–1096.

[57] E. Jovanov, A. Milenkovic, C. Otto and P. C. De Groen, 'A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation,' *Journal of NeuroEngineering and Rehabilitation*, vol. 2, 2005.

[58] L. A. Kelly, J. J. Reilly, D. M. Jackson, C. Montgomery, S. Grant and J. Y. Paton, 'Tracking physical activity and sedentary behavior in young children.,' *Pediatric exercise science*, vol. 19, no. 1, pp. 51–60, Feb. 2007.

[59] S. Kokoska and D. Zwillinger, *CRC Standard Probability and Statistics Tables and Formulae, Student Edition*. 1999, p. 15.

[60] P. Krebs, J. O. Prochaska and J. S. Rossi, 'Defining what Works in Tailoring: A Meta-Analysis OF Computer Tailored Interventions for Health Behavior Change,' *Prev Med*, vol. 51, no. 3-4, pp. 214–221, 2010.

[61] J. R. Kwapisz, G. M. Weiss and S. A. Moore, 'Activity recognition using cell phone accelerometers,' *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[62] P. H. Lee, Y. Y. Yu, I. McDowell, G. M. Leung and T. H. Lam, 'A cluster analysis of patterns of objectively measured physical activity in Hong Kong,' *Public Health Nutrition*, vol. 16, no. 8, pp. 1436–1444, 2013.

[63] R. M. Leech, S. A. McNaughton and A. Timperio, 'The clustering of diet, physical activity and sedentary behavior in children and adolescents: A review,' *International Journal of Behavioral Nutrition and Physical Activity*, vol. 11, no. 1, p. 4, 2014.

[64] A. Liberati, D. G. Altman, J. Tetzlaff *et al.*, 'The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions:

explanation and elaboration,' *Journal of clinical epidemiology*, vol. 62, no. 10, pp. 1–34, 2009.

[65] M. L. A. Lustria, S. M. Noar, J. Cortese, S. K. Van Stee, R. L. Glueckauf and J. Lee, 'A meta-analysis of web-delivered tailored health behavior change interventions,' *Journal of health communication*, vol. 18, no. 9, pp. 1039–1069, 2013.

[66] J. Macqueen, 'Some methods for classification and analysis of multivariate observations,' *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, no. 233, pp. 281–297, Nov. 1967.

[67] A. E. Mark and I. Janssen, 'Influence of Bouts of Physical Activity on Overweight in Youth,' *American Journal of Preventive Medicine*, vol. 36, no. 5, pp. 416–421, 2009.

[68] R. Martinez-Maldonado, V. Echeverria, O. C. Santos, A. D. P. D. Santos and K. Yacef, 'Physical learning analytics: A multimodal perspective,' in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, New York, NY, USA: Association for Computing Machinery, 2018, pp. 375–379.

[69] R. Martinez-Maldonado, K. Yacef, A. D. P. D. Santos *et al.*, 'Towards Proximity Tracking and Sensemaking for Supporting Teamwork and Learning,' in *Proceedings - IEEE 17th International Conference on Advanced Learning Technologies, ICALT 2017*, IEEE, 2017, pp. 89–91.

[70] U. Maurer, A. Smailagic, D. P. Siewiorek and M. Deisher, 'Activity recognition and monitoring using multiple sensors on different body positions,' in *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, IEEE, 2006, 4–pp.

[71] D. Merom, C. Tudor- Locke, A. Bauman and C. Rissel, 'Active commuting to school among nsw primary school children: Implications for public health,' *Health Place*, vol. 12, no. 4, pp. 678–687, 2006.

[72] R. Miller, W. Brown and C. Tudor-Locke, 'But what about swimming and cycling? how to "count" non-ambulatory activity when using pedometers to assess physical activity,' *Journal of Physical Activity and Health*, vol. 3, no. 3, pp. 257–266, 2006.

[73] Misfit, *Fossil group*, URL: https://www.misfit.com/, [Accesed: 2-October-2022].

[74] J. S. Mollee, E. F. M. Araújo, A. Manzoor, A. T. van Halteren and M. C. A. Klein, 'Explaining Changes in Physical Activity Through a Computational Model of Social Contagion,' in *Springer Proceedings in Complexity*, Springer International Publishing, 2017, pp. 213–223.

[75] J. S. Mollee and M. C. A. Klein, 'Empirical validation of a computational model of influences on physical activity behavior,' in *Advances in Artificial Intelligence: From Theory to Practice*, S. Benferhat, K. Tabia and M. Ali, Eds., Cham: Springer International Publishing, 2017, pp. 353–363.

[76] E. M. Murtagh, J. L. Mair, E. Aguiar, C. Tudor-Locke and M. H. Murphy, 'Outdoor walking speeds of apparently healthy adults: A systematic review and meta-analysis,' *Sports medicine (Auckland, N.Z.)*, vol. 51, no. 1, pp. 125–141, Jan. 2021.

[77] A. Nazari, A. Dehghan, S. Nejatian, V. Rezaie and H. Parvin, 'A comprehensive study of clustering ensemble weighting based on cluster quality and diversity,' *Pattern Analysis and Applications*, vol. 22, no. 1, pp. 133–145, 2019.

[78] T. S. Olds, K. Ridley, J. Dollman and C. A. Maher, 'The validity of a computerized use of time recall, the multimedia activity recall for children and adolescents,' *Pediatric Exercise Science*, vol. 22, no. 1, pp. 34–43, 2010.

[79] W. H. Organization, *Global Recommendations on Physical Activity for Health WHO, Geneva 2010*. Jan. 2010, vol. 60, pp. 1–58.

[80] W. H. Organization, *WHO guidelines on physical activity and sedentary behaviour: at a glance*. World Health Organization, 2020, iv, 17 p.

[81] M. Ouzzani, H. Hammady, Z. Fedorowicz and A. Elmagarmid, 'Rayyan-a web and mobile app for systematic reviews,' *Systematic Reviews*, vol. 5, no. 1, pp. 1–10, 2016.

[82] H. Parvin and B. Minaei-Bidgoli, 'A clustering ensemble framework based on selection of fuzzy weighted clusters in a locally adaptive clustering algorithm,' *Pattern Analysis and Applications*, vol. 18, no. 1, pp. 87–112, 2014.

[83] L. R. S. Phillips, G. Parfitt and A. V. Rowlands, 'Calibration of the GENEA accelerometer for assessment of physical activity intensity in children,' *Journal of Science and Medicine in Sport*, vol. 16, no. 2, pp. 124–128, 2013.

[84] G. Plasqui, A. G. Bonomi and K. R. Westerterp, 'Daily physical activity assessment with accelerometers: New insights and validation studies,' *Obesity Reviews*, vol. 14, no. 6, pp. 451–462, 2013.

[85] R Core Team, 'R: A language and environment for statistical computing,' 2021.

[86] M. Rabbi, A. Pfammatter, M. Zhang, B. Spring and T. Choudhury, 'Automated Personalized Feedback for Physical Activity and Dietary Behavior Change With Mobile Phones: A Randomized Controlled Trial on Adults,' *JMIR mHealth and uHealth*, vol. 3, no. 2, e42, 2015.

[87] S. Ratel, C. Williams, J. Oliver and N. Armstrong, 'Effects of age and recovery duration on performance during multiple treadmill sprints,' *International journal of sports medicine*, vol. 27, no. 01, pp. 1–8, 2006.

[88] S. Ratel, P. Duché and C. A. Williams, 'Muscle fatigue during high-intensity exercise in children,' *Sports Medicine*, vol. 36, no. 12, pp. 1031–1065, 2006.

[89] N. Ravi, N. Dandekar, P. Mysore and M. L. Littman, 'Activity recognition from accelerometer data,' in *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence - Volume 3*, ser. IAAI'05, Pittsburgh, Pennsylvania: AAAI Press, 2005, pp. 1541–1546.

[90] K. Ridley, S. Zabeen and B. K. Lunnay, 'Children's physical activity levels during organised sports practices,' *Journal of Science and Medicine in Sport*, vol. 21, no. 9, pp. 930–934, 2018.

[91] W. Robertson, S. Stewart-Brown, E. Wilcock, M. Oldfield and M. Thorogood, 'Utility of accelerometers to measure physical activity in children attending an obesity treatment intervention,' *Journal of obesity*, vol. 2011, 2010.

[92] L. Rokach and O. Maimon, 'Chapter 15 Minkowski : Distance Measures for Numeric Attributes,' *Data Mining and Knowledge Discovery Handbook*, pp. 322–352, 2005.

[93] A. V. Rowlands, 'Moving forward with accelerometer-assessed physical activity: Two strategies to ensure meaningful, interpretable, and comparable measures,' *Pediatric Exercise Science*, vol. 30, no. 4, pp. 450–456, 2018.

[94] S.-J. Salvy and J. C. Bowker, 'Peers and obesity during childhood and adolescence: A review of the empirical research on peers, eating, and physical activity,' *Journal of obesity & weight loss therapy*, vol. 4, no. 1, 2014.

[95] O. C. Santos, 'Training the body: The potential of aied to support personalized motor skills learning,' *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 730–755, 2016.

[96] O. C. Santos, 'Education still needs Artificial Intelligence to support personalized motor skill learning: Aikido as a case study,' *CEUR Workshop Proceedings*, vol. 1432, pp. 72–81, 2015.

[97] C. A. Schaefer, C. R. Nigg, J. O. Hill, L. A. Brink and R. C. Browning, 'Establishing and evaluating wrist cutpoints for the GENEActiv accelerometer in youth,' *Medicine and Science in Sports and Exercise*, vol. 46, no. 4, pp. 826–833, Apr. 2014.

[98] H. Schäfer, J. Bachner, S. Pretscher, G. Groh and Y. Demetriou, 'Study on Motivating Physical Activity in Children with Personalized Gamified Feedback,' *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization - UMAP '18*, pp. 221–226, 2018.

[99] T. Semplonius and T. Willoughby, 'Long-term links between physical activity and sleep quality,' *Med. Sci. Sports Exerc*, vol. 50, pp. 2418–2424, 2018.

[100] R. J. Shephard, 'Is active commuting the answer to population health?' *Sports medicine*, vol. 38, no. 9, pp. 751–758, 2008.

[101] G. Siemens, 'Learning Analytics: The Emergence of a Discipline,' *American Behavioral Scientist*, vol. 57, no. 10, pp. 1380–1400, 2013.

[102] G. Sprint, D. J. Cook and M. Schmitter-Edgecombe, 'Unsupervised detection and analysis of changes in everyday physical activity data,' *Journal of Biomedical Informatics*, vol. 63, pp. 54–65, 2016.

[103] V. A. Temple and J. W. Walkley, 'Physical activity of adults with intellectual disability,' *Journal of Intellectual and Developmental Disability*, vol. 28, no. 4, pp. 342–353, 2003.

[104] V. Thambawita, S. A. Hicks, H. Borgli *et al.*, 'Pmdata: A sports logging dataset,' in *Proceedings of the 11th ACM Multimedia Systems Conference*, ser. MMSys '20, Istanbul, Turkey: Association for Computing Machinery, 2020, pp. 231–236.

[105] C. Tudor-Locke and D. R. Bassett, 'How many steps/day are enough?' *Sports medicine*, vol. 34, no. 1, pp. 1–8, 2004.

[106] C. Tudor-Locke, C. L. Craig, W. J. Brown *et al.*, 'How many steps/day are enough? for adults,' *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, no. 1, pp. 1–17, 2011.

[107] S. J. Verstraete, G. M. Cardon, D. L. De Clercq and I. M. De Bourdeaudhuij, 'Effectiveness of a two-year health-related physical education intervention in elementary schools,' *Journal of Teaching in Physical Education*, vol. 26, no. 1, pp. 20–34, 2007.

[108] S. L. Williams and D. P. French, 'What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour—and are they the same?' *Health Education Research*, vol. 26, no. 2, pp. 308–322, Feb. 2011.

[109] M. Worsley, 'Multimodal learning analytics as a tool for bridging learning theory and complex learning behaviors,' *3rd Multimodal Learning Analytics Workshop and Grand Challenges, MLA 2014*, pp. 1–4, 2014.

[110] X. Wu, V. Kumar, J. R. Quinlan *et al.*, 'Top 10 algorithms in data mining,' *Knowledge and information systems*, vol. 14, no. 1, pp. 1–37, 2008.

[111] K. Yacef, C. Caillaud and O. Galy, 'Supporting learning activities with wearable devices to develop life-long skills in a health education app,' in *Artificial Intelligence in Education*, Cham: Springer International Publishing, 2018, pp. 394–398.

[112] C.-C. Yang and Y.-L. Hsu, 'A review of accelerometry-based wearable motion detectors for physical activity monitoring,' *Sensors*, vol. 10, no. 8, pp. 7772–7788, 2010.

[113] H. Zheng, C. Nugent, P. McCullagh *et al.*, 'Smart self management: Assistive technology to support people with chronic disease,' *Journal of Telemedicine and Telecare*, vol. 16, no. 4, pp. 224–227, 2010.

[114] M. Zhou, Y. Fukuoka, Y. Mintz *et al.*, 'Evaluating machine learning–based automated personalized daily step goals delivered through a mobile phone app: Randomized controlled trial,' *JMIR mHealth and uHealth*, vol. 6, no. 1, 2018.

[115] Z. Zhou, W. Dai, J. Eggert *et al.*, 'A real-time system for in-home activity monitoring of elders,' *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009*, pp. 6115–6118, 2009.

[116] J. Zhu, D. H. Dallal, R. C. Gray *et al.*, 'Personalization Paradox in Behavior Change Apps,' *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–21, Apr. 2021.