Kinetic Signatures: A Systematic Investigation of Movement-Based User Identification in Virtual Reality

Jonathan Liebers
University of Duisburg-Essen
Essen, Germany
jonathan.liebers@uni-due.de

Patrick Laskowski
University of Duisburg-Essen
Essen, Germany
patrick.laskowski@stud.uni-due.de

Florian Rademaker
University of Duisburg-Essen
Essen, Germany
florian.rademaker@stud.uni-due.de

Leon Sabel
University of Duisburg-Essen
Essen, Germany
leon.sabel@stud.uni-due.de

Jordan Hoppen
University of Duisburg-Essen
Essen, Germany
jordan.tkocz@stud.uni-due.de

Uwe Gruenefeld
University of Duisburg-Essen
Essen, Germany
uwe.gruenefeld@uni-due.de

Stefan Schneegass
University of Duisburg-Essen
Essen, Germany
stefan.schneegass@uni-due.de

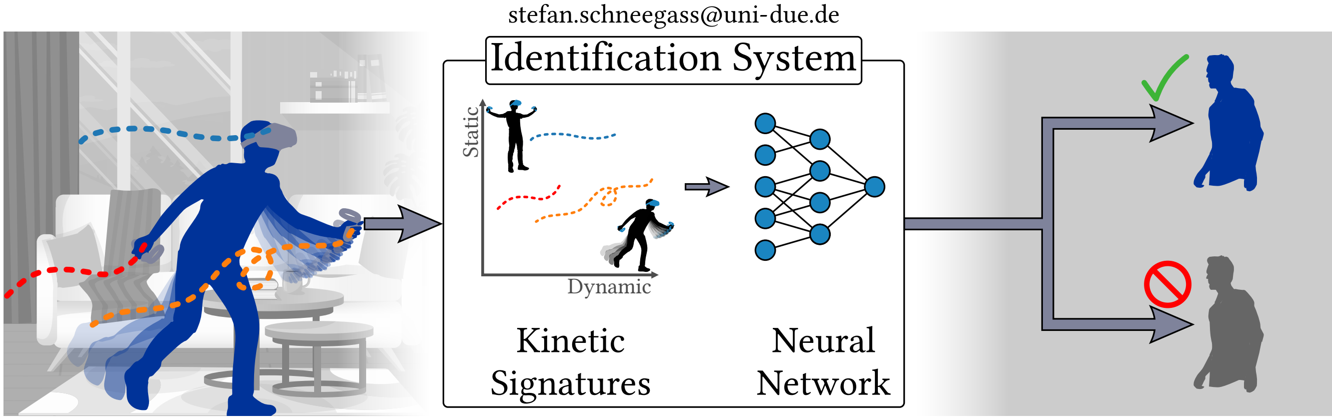


Figure 1: We systematically investigated the identifiability of kinetic signatures in VR. A kinetic signature is a human's spatiotemporal movement data that contains a high degree of individual uniqueness so that it is suitable for implicit user identification in VR. Particularly, we explored movement behavior's static and dynamic components for its biometric uniqueness.

ABSTRACT

Behavioral Biometrics in Virtual Reality (VR) enable implicit user identification by leveraging the motion data of users' heads and hands from their interactions in VR. This spatiotemporal data forms a *Kinetic Signature*, which is a user-dependent behavioral biometric trait. Although kinetic signatures have been widely used in recent research, the factors contributing to their degree of identifiability remain mostly unexplored. Drawing from existing literature, this work systematically examines the influence of static and dynamic components in human motion. We conducted a user study (N = 24) with two sessions to reidentify users across different VR sports and exercises after one week. We found that the identifiability of a kinetic signature depends on its inherent static and dynamic factors, with the best combination allowing for 90.91% identification accuracy after one week had passed. Therefore, this work lays a foundation for designing and refining movement-based identification protocols in immersive environments.

CCS CONCEPTS

• **Human-centered computing** → *Human computer interaction (HCI)*; *Virtual reality*; • **Security and privacy** → *Usability in security and privacy*.

KEYWORDS

identification, virtual reality, kinetic signatures, usable security, task-driven biometrics

ACM Reference Format:

Jonathan Liebers, Patrick Laskowski, Florian Rademaker, Leon Sabel, Jordan Hoppen, Uwe Gruenefeld, and Stefan Schneegass. 2024. Kinetic Signatures: A Systematic Investigation of Movement-Based User Identification in Virtual Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3613904.3642471

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642471

1 INTRODUCTION

Human behavior can show fundamental differences between people. In fact, behavior inherently tends to have an exceptionally high degree of individuality, making it suitable for reliable behavioral biometric user identification, particularly in VR [26, 34, 44, 52]. Widely available off-the-shelf head-mounted displays (HMDs) can elicit this behavior through their sensors using only the motion data of the head and the controllers. This data is so individual that it forms user-specific *kinetic signatures*, which are defined as *traces of users' individual spatiotemporal body behavior, making it suitable for behavioral biometric user identification*. Similar to the idea of so-called "signature moves," kinetic signatures can tell who a person is. As the VR HMD captures this data anyway, it can be used for behavioral biometric identification at no additional cost, benefitting the user [23]. Thereby, the VR device can automatically recognize their users, removing the burden of current authentication methods, which are often cumbersome and time-consuming [5, 11, 12, 59].

Research has shown many examples of applied behavioral biometric identification systems using kinetic signatures in VR [1, 24, 26, 28, 32, 34, 35, 49, 52, 55]. Each previous study utilized its own kinetic signatures in VR, ranging from throwing virtual balls [1, 24, 35–38] to grabbing, pointing, walking, and typing motions [52], to sports such as archery or bowling [26]. However, these works were not conducted systematically concerning the nature of kinetic signatures; instead, they focused on individual applications and interactions. Thus, a central piece of knowledge is missing: a connection between all those individually employed kinetic signatures. For example, why does a bowling ball throw provide better identifiability than an archery shot [26]? Why is walking more beneficial for behavioral biometric identification than pointing or grabbing [52]? So far, these questions remain unanswered, and transferability and generalizability of kinetic signatures are lacking.

Other disciplines, such as kinesiology, sports science, and movement science, provide dimensions for characterizing and classifying human movements and sports [17, 25, 39]. One popular classification method categorizes sports based on static and dynamic dimensions according to the required muscular activities [39, 40]. Static refers to the muscular activity required to hold joints in place, while dynamic refers to muscle activity altering the positions. The benefit of this classification is that domain experts have already categorized multiple sports along these dimensions.

In our work, we applied this classification to behavioral biometrics, providing the first systematic investigation of a kinetic signature's identifiability. Through a user study with two sessions one week apart (N=24), we found that static and dynamic components significantly influence a signature's identifiability. We tested various combinations of static and dynamic movement components in VR, ranging from exercises to sports such as climbing and boxing. Essentially, we found that movements with fewer static components enhance identifiability, followed by highly dynamic components. The best combination resulted in up to 90.91% identification accuracy at a median recall rate of 95.45%. Thus, our results can aid in the design of identification systems and conclude with a discussion on how to ideally employ kinetic signatures, allowing experts to further improve their user identification rates (cf., Figure 1).

> **Contribution Statement.** The contribution of our work is threefold. First, we contribute a systematic empirical investigation (N=24) that presents findings on the effects of static and dynamic components within human motion and how they influence a kinetic signature's grade of behavioral biometric identifiability across two study sessions. Second, we provide a discussion of how experts can design motions that lead to better identification, allowing implicit identification in VR. Finally, we publish our elicited dataset and apparatus to allow further experimentation by the scientific community.

2 FOUNDATIONS AND RELATED WORK

Our work is situated at the crossroads of identification in VR through behavioral biometrics, particularly kinetic signatures, and the foundations of human movement. Additionally, we briefly cover the foundations of authentication and the implications on a system's usability, user experience, and user's privacy. We also provide an brief overview of existing data sets in this field.

2.1 Present and Future Authentication in Virtual Reality

Computers, and VR HMDs as such, regularly need to establish trust in their users' identity. This is imperative to ensure that sensitive information is presented solely to the legitimate user, thus guaranteeing information security. However, it is equally essential to recognize the user's identity to enable the personalization or individualization of applications [7]. For instance, this allows an application to adapt to the user by, e.g., loading their settings or user profile [6]. Stephenson et al. provide a comprehensive Systematization of Knowledge (SoK) of user authentication in Augmented Reality (AR) and VR, evaluating the state-of-the-art of authentication mechanisms by systematizing research efforts and practical deployments by comparing different forms of authentication in as well (e.g., knowledge-based, physical biometrics, and behavioral biometrics, among others) [63].

The best-known authentication process is knowledge-based, where the user presents a knowledge-component to a computer, such as a Personal Identification Number (PIN) or password. However, passwords have a plethora of shortcomings, such as overloading users by requiring them to store dozens of them in their minds [59], being guessable [10, 48], and being difficult to enter [11]. Additionally, bystanders of an immersed VR user can spy on their hand movements while they enter a password [12]. Thus, calls emerged to reshape authentication in the future [4, 5, 16]. Proposed alternatives are possession-based (e.g., security tokens) or preferably biometrics [5, 16]. Tokens can be impractical as they need to be carried around and can be lost, transferred, or stolen; biometrics, however, stay with the user at all times and cannot be changed, which is both an advantage and a disadvantage [22]. As HMDs are worn close to the body, they elicit lots of information on the user's behavior through their sensors, such as their head and controller movements, allowing for identification of the user, as shown in many conducted research projects [1, 24, 26, 28, 32, 34, 35, 49, 52, 55, 70]. These user's movements form a kinetic signature, i.e., the way users move is specific, individual, and suitable for identification. Research showed that such behavior can be hard to imitate [41] and that humans also can perform such identification tasks (e.g., a person being able to identify their friends by their gait) [13]. Cuttings et al. particularly showed that this is possible for people without any familiarity cues, i.e., a person recognizing another person based on their gait pattern without having access to another source of familiarity [13]. Researchers seeking identification for the purpose of authentication, on the other hand, generally utilize algorithms such as deep learning or machine learning [22, 64]; to the best of our knowledge, no exploration of kinetic signatures in VR was performed so far by employing human recognition for the identification task.

2.2 Foundations of Authentication

"Authentication" itself is an umbrella term, as it can be realized through either *verification mode* or *identification mode* [22]. In verification mode, users must provide (1) a claim of identity and (2) proof of their claim [22]. The identity claim must correspond to an identity stored in the computing system (e.g., by providing a username), and the proof acts as a legitimation (e.g., a password corresponding to the username). In behavioral biometrics, this proof is a sample of a user's behavior, such as their head and hand movements (a kinetic signature). However, from a user's perspective, identification mode is more straightforward and more favorable than verification; here, the computer only requires the proof and automatically deducts the corresponding account from the proof, not requiring the user's provided claim of identity [22] (e.g., a user is recognized only by their kinetic signature, not having to provide a username). If this behavioral sample can be elicited from users without requiring their explicit cooperation, the identification can happen without the users realizing it. This enables identification through an implicit interaction, which is highly favorable, as the burden associated with authentication is fully removed from the user, and users might not even realize that they undergo an authentication process [23, 61, 62]. This benefits user experience as users do not have to deal with annoying authentication systems anymore, and at the same time, it increases the usability of the underlying actual system that users primarily seek to use. Of course, the assumption applies that the implicit authentication system functions reliably.

2.3 Foundations of Human Movement and Classification of Sports

One way to view human movement stems from the goal of accomplishing activities, particularly sports. This is possible by activating various muscles inside the body with support from the skeletal and articular systems [3, 19].

Generally speaking, muscles in the human body can activate and thus exert a force on a joint; the joint moves and a movement is the result. Figures 2 and 3(a) provide an overview of the interplay between muscles and joints [9]. To move a joint, muscles can perform a *concentric* (cf., Figure 2(a)) or *eccentric* contraction (cf., Figure 2(b)) [8, 50]. The interplay of concentric and eccentric contractions defines *isotonic* movements. Muscles can also contract and generate forces but not move any joints [8]. This is an *isometric* contraction and occurs, e.g., while carrying heavy objects with an arm that is not fully extended, where muscles are strained through holding the arm in place (cf., Figure 3(a)) [8].

2.3.1 *Classification of Sports.*

While these are the basic foundations of human joint movements, activities, such as sports, are increasingly complex. One popular classification was created by Mitchell et al. [39, 40]. Their classification classifies sports by the degree of static and dynamic components in the exercises on two axes. Static components are related to *isometric* movements, and examples of highly static sports are, e.g., climbing, weight lifting, or windsurfing. Here, the muscles are heavily stressed without creating many joint movements. On the other hand, dynamic components in sports are related to *isotonic* activities created by *concentric* and *eccentric* muscle contractions and associated with other classes of sports, e.g., badminton, tennis, or long-distance running. In their 3×3 classification, there also exist sports that are very static and dynamic at the same time (e.g., boxing, with a high degree of *isotonic* and *isometric* activity) or only a little static and little dynamic in combination (e.g., bowling or billard with low degrees of both) [39, 40]. Additionally, an intermediate level exists per dimension.

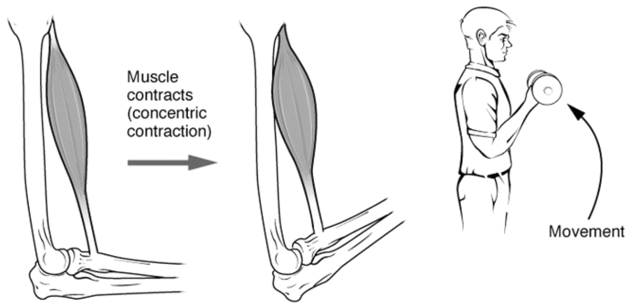
2.3.2 *The Classification's Static and Dynamic Dimensions.*

Mitchell et al. regard the axes "dynamic" and "static" as two independent dimensions that both have different effects on the human body during the exercise of sports [39, 40]. Both dimensions are connected to a different cardiovascular response. Dynamic exercises that are performed with a large mass of muscles cause a marked increase in oxygen consumption and also a substantial increase in cardiac output, heart rate, and stroke volume, among other factors. In contrast, they also state that static exercise causes only a small increase in oxygen consumption, cardiac output, heart rate, and no change in stroke volume [39, 40].

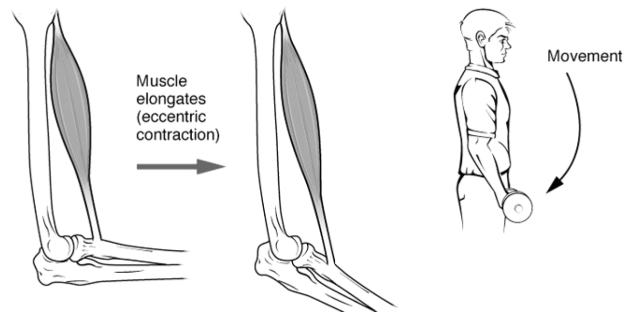
Furthermore, they narrow that their proposed classification of sports "should not be regarded as a rigid classification, but rather a continuum in which some athletes in the same sport could possibly deserve placement in different categories". For example, different players participating in a sport with different roles might be affected differently, e.g., a striker might experience a different effect of "static" and "dynamic" in soccer, compared to a goalkeeper. Given their definition of the classification as a continuum, it could also be possible that slight differences for each sport per class exist, as they might have different effects on specific physiological effects. Mitchell et al. also state that an athlete's physiological response to the sport could also be influenced by factors such as emotional stress or the specific training regimen; thus, the defining dimensions of the classification can also be subject to influence by further factors [39].

2.3.3 *Application to Kinetic Signatures.*

In our work, we apply the classification of Mitchell et al. to activities performed in VR for the purpose of user identification and use their classification to classify the activities conducted in user studies of existing related work. Particularly, we focus on the present levels of "dynamic" (isotonic muscle contractions, cf., Figure 2) and "static" (isometric muscle contractions, cf., Figure 3(a)) movement components. To do so, we form abbreviations. "Static" can be expressed as *Low Static (LS)*, *Medium Static (MS)*, or *High Static (HS)*. "Dynamic" can be either



(a) Concentric Contraction.



(b) Eccentric Contraction.

Figure 2: Figure of isotonic muscle contractions in the human body. (a) Concentric movements happen when the muscle length is reduced due to contraction. When the muscle elongates, its contraction is eccentric, shown in (b) [8, 9].

Low Dynamic (LD), *Medium Dynamic (MD)*, or *High Dynamic (HD)*. Therefore, a *Low Static (LS)* sport that is of *High Dynamic (HD)* nature such as tennis (cf., Figure 3(b) [39, 40]) is abbreviated as *LSHD*. In the following, Figure 4 also provides a comprehensive overview.

2.4 Identification through Kinetic Signatures in Virtual Reality

Researchers proposed many ways to authenticate users in VR. Examples range from 3D pattern entries [72], PIN entry on virtual objects [12, 32], using eye-tracking data [20, 29, 30, 34, 70], finger-tracking traces [27], or patterns in head-movement [31, 43, 68]. Yet, one popular cluster of approaches exists that is based on kinetic signatures and so-called "task-driven biometrics", where the user's head and controller movement patterns are exploited for identification [24].

2.4.1 Task-driven Biometrics in VR, their Classification, and Stability. Pfeuffer et al. conducted an exploration of behavioral biometrics in VR, where they compared a pointing, grabbing, walking, and typing task in a user study ($N = 19$) with two sessions [52]. Given the previously introduced classification of Mitchell et al. and their definition of "static" and "dynamic" through the required degree of isotonic and isometric muscle movements (cf., Figure 3(b)), their pointing and grabbing task can be classified as *Low Static (LS)* and *Low Dynamic (LD)* (*LSLD*), and walking as *Low Static (LS)* and *High Dynamic (HD)* (*LSHD*) [39, 40]. Their typing task is hard to classify since it is executed within a very small area with the hands and fingers; thus, it does not mimic full-body kinetic signatures well, and typing itself is a strong biometric trait [56]. Related to grabbing, Olade et al. also conducted an experiment where users had to pick and place objects in VR [49]. They found an accuracy of up to 98.6%.

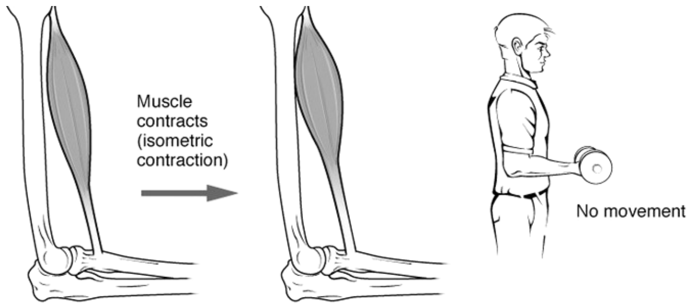
A cluster of works exists on throwing virtual balls in VR. For example, Ajit et al., Kupin et al., and R. Miller et al. extensively studied the throwing-movement in VR [1, 24, 35–38]. Rack et al. also used their dataset to evaluate their distance-based machine learning algorithms [55]. Throwing in VR can be classified as *High Static (HS)* and *Low Dynamic (LD)* (*HSLD*), as it is included as "field events (throwing)" in the classification of Mitchell et al. (cf., Figure 3(b)) [39].

Liebers et al. employed kinetic signatures stemming from well-known sports, such as bowling or archery, thus comparing two sports in addition to virtual modifications of the body [26]. According to the original 3×3 classification of Mitchell et al., archery is classified as *Medium Static (MS)* and *Low Dynamic (LD)* (*MSLD*) and bowling is classified as a *Low Static (LS)* and *Low Dynamic (LD)* (*LSLD*) sport [39]. Throughout their user study ($N = 16$), they found accuracy ratings of 54% for archery (*MSLD*) and 59% for bowling (*LSLD*).

Furthermore, Abdrabou et al. explored eleven contextual factors and activities that influence kinetic signatures in reality, such as beneficial body parts, activities like walking, sitting (down) and tasks during sitting, and walking while carrying items using a motion-capture system in reality [2]. Their user study did not take place in VR. Their tasks can be classified into three groups: (1) chatting, sitting down, watching news, watching horror movies, playing games, and watching animated movies, all falling into *LSLD* since they were mostly executed while sitting. (2) Standing up and walking fall into *LSHD*, as standing up and walking requires isotonic movements of the legs (cf., Figure 3(b)). (3) Walking with glasses, walking with filled glasses, and walking with glasses on a plate fall into *HSHD* due to the required isometric contraction for carrying items in the arms, in addition to the isotonic movements of the legs during walking. They list accuracies between 6.74% to 13.88% for their first group (*LSLD*), 28.82% to 40.75% for the second group (*LSHD*), and 53.01% to 59.21% for the third group (*HSHD*) in their Figure 5 [2].

As many behavioral biometric identification systems employ machine-learning algorithms, Schell et al. investigated different data encodings and model architectures for ideal user identification [60]. Furthermore, R. Miller et al. showed that identification is possible across VR systems, but there is an indication that accuracy is reduced between systems [53].

Most previously presented works rely on distinct kinetic signatures (e.g., one ball throw or one archery shot). However, recent works also focus on non-specific movements, where users' behavior allows for a high degree of freedom in their movements. Examples



(a) Isometric Contraction.

Increasing Static M. Component ↑

High Static and Low Dynamic Sports, e.g., Bobsledding, Luge, Throwing field events, Gymnastics, Martial arts, Sailing, **Sport climbing**, Water skiing, Weight lifting, Windsurfing.	*High Static and High Dynamic Sports*, e.g., **Boxing**, Canoeing, Kayaking, Cycling, Decathlon, Rowing, Speed-skating, Triathlon.
Low Static and Low Dynamic Sports, e.g., Billards, **Bowling**, Cricket, Curling, Golf, Riflery.	*Low Static and High Dynamic Sports*, e.g., Badminton, Cross-country skiing, Field hockey, Orienteering, Race walking, Squash, Long distance running, Soccer, **Tennis.**

Increasing Dynamic Movement (M.) Component →

(b) Simplified Classification of Sports.

Figure 3: Human muscles can also contract but keep the joints still, which is an isometric contraction shown in (a) [8, 9]. (b) Our simplified 2×2 classification matrix of sports using dynamic and static movement components, based on the extrema of the 3×3 classification proposed by Mitchell et al. [39, 40]. We selected the sports marked in bold for our experiment [39].

range from playing games like "Half-Life: Alyx" to "Beat Saber" in VR [28, 44, 55].

One important aspect of kinetic signatures is their stability over time, i.e., how passing time influences the signature. Their identifiability often decreases the more time passes [28, 38]. Therefore, many studies employ multi-session designs, reinviting participants after some time passed to obtain a repeated measure of a user's kinetic signature [2, 26–28, 52]. For example, M. R. Miller et al. conducted a large-scale user study ($N = 232$) for identification in VR, also exploring identifiability over time and Liebers et al. conducted a user study in the field over eight weeks ($N = 15$) [28, 33].

2.4.2 Data Sets for Research on User Identification in VR. Additionally, researchers proposed multiple data sets for conducting evaluations in this area. One of the largest data sets in the area was contributed by Nair et al. ($N > 50.000$), including motion captures from the popular "Beat Saber" and "Tilt Brush" applications [45]. Furthermore, Rack et al. also made a data set available by having their participants play "Who is Alyx?" ($N = 71$) [54]. Also, Wen et al. presented a dataset offering approximately 12-hour gameplay videos from ten real-world games in 10 diverse genres for motion sickness research ($N = 25$) [69]. R. Miller et al. also provide a dataset of VR motions across three VR systems with ten motions per session and $N = 41$ participants, where the participants threw virtual balls [37, 38]. Liebers et al. also published their data sets on Kinetic Signatures using body normalizations with archery and bowling ($N = 16$), and also on using head-rotation data and eye-tracking in VR ($N = 12$) [26, 29].

Compared to the previous data sets, our data set offers the kinetic signatures of $N = 24$ users that were manually annotated. Its kinetic samples are organized into two dimensions (static and dynamic components in human movements) and split by modality: the execution of a sports and self-exercise activity in VR. While we offer data elicited from fewer participants compared to the previous data sets, we present a fully balanced data set across two sessions that is composed of eight systematically chosen activities that participants experienced using a Meta Quest 2 device.

2.4.3 Implications on User's Privacy. While identification methods can be beneficial to the user, they can also pose a threat to their privacy, e.g., if the identification is performed without the user's consent or if it is not in their personal interest. For example, Tricomi et al. provide a framework for user profiling in augmented and virtual reality, stating that it becomes increasingly difficult to hide the personal identity when using a HMD [67]. Moore et al. show that encoding user tracking data in VR can be obfuscated by encoding positional data as velocity data to preserve the user's privacy in VR [42]. Also, Garrido et al. provide a SoK on data privacy in VR, systematizing knowledge on the landscape of VR privacy threats and countermeasures by proposing a taxonomy of data attributes, protections, and adversaries [18]. Notably, Kinetic signatures also can be used to infer personal attributes of VR users [46]. Finally, Nair et al. provide a prototype for increasing user privacy in VR [47]. All in all, it is an ethical imperative that the identification of people using kinetic signatures should be conducted in conformation with their consent.

2.5 Summary and Research Gap

While kinetic signatures are already widely employed in identification systems research, most previously presented works are case studies of mostly singular activities [1, 24, 32, 36]. Rarely, more than one activity is explored, and often it remains unclear why one activity performs better or worse than the other for identification [2, 26, 52]. To the best of our knowledge, a systematic exploration of the kinetic signature's identifiability based on the actual human movement itself does not exist yet. Thereby, it remains unclear *why* a kinetic signature's identifiability for a grabbing or pointing task differs from another signature that was created from walking [52] or *why* throwing a bowling ball differs from launching an arrow in archery [26]. While other researchers already explored personal identifiability over time [28, 33, 38], across different VR systems [53], the obfuscation of tracking data [42, 47], data encodings and machine learning architectures [60], or differences between AR and VR [27, 67], we focus on the actual human movements (kinetic signatures) and their inherent varying degree of identifiability. Our presented research aims to understand what contributes to the inherent identifiability of a kinetic signature in VR by their execution characteristics founded in the human movement sciences. Therefore, we created a simplified version of the classification proposed by Mitchell et al. that consists of their extrema, reducing their 3×3 matrix into a 2×2 matrix (cf., Figure 3(b)). We use this simplified matrix to systematically explore user identification through kinetic signatures in VR throughout our experiment [39].

3 EXPERIMENT

To explore the dependency between the identifiability of kinetic signatures and the dimensions within our simplified classification of sports and movements (cf., Figure 3(b)) that is based on the work of Mitchell et al. [39, 40], we conducted a controlled lab experiment. Based on the classification's dimensions, our central research question was:

RQ How do the *Static* and *Dynamic* movement components influence a kinetic signature's identifiability?

We furthermore wondered about the VR context that the kinetic signatures are elicited in. While the classification proposed by Mitchell et al. primarily aims at sports [39, 40], we wondered if the same principles of static and dynamic movements would be applicable to a more uniform context. For this purpose, we invited the head of university sports, a professional sports scientist at our institution, and reflected in cooperation with her on how a more uniform context could be created in VR. Following this process, we opted to explore sports in VR, which is central to the classification of Mitchell et al. [39, 40]. Additionally, we decided to investigate our research question in a uniform VR context where our participants would perform self-exercises.

3.1 Hypotheses

During our interdisciplinary cooperation, we discussed how the classification dimensions in Figure 3(b) could impact the identifiability of kinetic signatures. We believe that more movements are generally desirable since this allows the HMD to elicit more user-specific data, suitable for identification. Given the definitions of isometric and isotonic movements and their usage in existing literature, we therefore posit the following hypotheses [8, 39, 50]:

H1 Kinetic signatures with a *High Dynamic (HD)* component lead to a significantly *higher* identifiability compared to kinetic signatures with a *Low Dynamic (LD)* component during: (a) self-exercises, (b) sports.

H2 Kinetic signatures with a *High Static (HS)* component lead to a significantly *lower* identifiability compared to kinetic signatures with a *Low Static (LS)* component during: (a) self-exercises, (b) sports.

Hypothesis 1: Dynamic Kinetic Signatures. We informed our hypothesis with the work of Pfeuffer et al., who explored behavioral biometrics in VR by employing, among others, a pointing (*Low Static (LS)* and *Low Dynamic (LD)* – *LSLD*), grabbing (*LSLD*), and walking activity (*LS* and *High Dynamic (HD)* – *LSHD*) throughout a user study ($N = 19$, cf., Section 2.4.1) [52]. In their Table 1, they reported for the three tracked objects of an HMD (head, dominant hand, and non-dominant hand) and all present features an accuracy of 33.34% for pointing (*LSLD*), 27.03% for grabbing (*LSLD*), and 39.31% for walking (*LSHD*) [52]. **H1** is derived from their results, as highly dynamic walking performed better than low-dynamic pointing or grabbing. This applied to their groups Individual/All, Move+Stabilise/All, Distance/All, and Target/Distance/All (cf., their Table 1) [52]). Their results indicate that an increase in dynamic components benefits identification (**H1**).

Hypothesis 2: Static Kinetic Signatures. We informed our hypothesis with the work of Liebers et al., who compared two sports activities in VR, namely archery (*Medium Static (MS)* and *Low Dynamic (LD)* – *MSLD*, cf., Section 2.4.1) and bowling (*LSLD*) [26]. They obtained an accuracy of 54% for archery (*MSLD*) and 59% for bowling (*LSLD*), using their better Recurrent Neural Network (RNN) model (cf., their Table 2) [26]. The same order applies for their worse Multilayer Perceptron (MLP) model at respective accuracy values of 38% and 49% for archery and bowling [26]. Their results suggest that a decreased static component increases the identification rate (**H2**).

3.2 Study Design

We chose two independent variables to explore throughout our hypotheses and research question. The first independent variable was *Static Component* with two levels: *Low Static (LS)* and *High Static (HS)*. Additionally, we chose *Dynamic Component* with again two levels: *Low Dynamic (LD)* and *High Dynamic (HD)*. The levels of the variables correspond to our simplified classification in Figure 3(b), and $2 \times 2 = 4$ total combinations exist. As a dependent variable, we chose recall rate, which is a per-participant-defined performance metric that states how often a user is correctly identified (TP) divided by the sum of correct identifications (TP) and misclassification (FN): recall rate $= \frac{TP}{TP+FN}$ We implemented every combination as an activity in our apparatus, both in a sport (SP) setting and in a self-exercise (EX) setting in VR.

To refer to an individual activity, we concatenate the expressions to form an abbreviation. For example, *LSLDSP* refers to a low-static (LS), low-dynamic (LD) sports (SP), whereas *HSHDEX* refers to a high-static (HS), high-dynamic (HD) exercise (EX). Figure 4 provides a comprehensive overview by visualizing the activities, variables, and abbreviations.

3.3 Apparatus

Our apparatus used in the user study consists of an implemented multi-purpose logger and commercially available applications for VR. All VR applications are available on Steam[1] and the total acquisition cost for five different games was approximately 90 USD. All applications used the SteamVR runtime environment[2] and an OpenVR API[3] to be compatible with our logger.

3.3.1 Multi-purpose Logger. We implemented a logger that is capable of capturing the spatial data that is streamed from the HMD. It captures positional, rotational data (Quaternions), button presses, and trigger presses of an attached HMD and its controllers if the executed application supports SteamVR and OpenVR. It saves the data as tab-separated values (TSV) files. The logger is based on the

[1]Valve's Steam is a digital distribution platform for software. https://steampowered.com, last retrieved on March 8, 2024.
[2]SteamVR on steampowered.com. https://store.steampowered.com/app/250820/SteamVR, last retrieved March 8, 2024.
[3]OpenVR SDK on GitHub.com. https://github.com/ValveSoftware/openvr, last retrieved March 8, 2024.

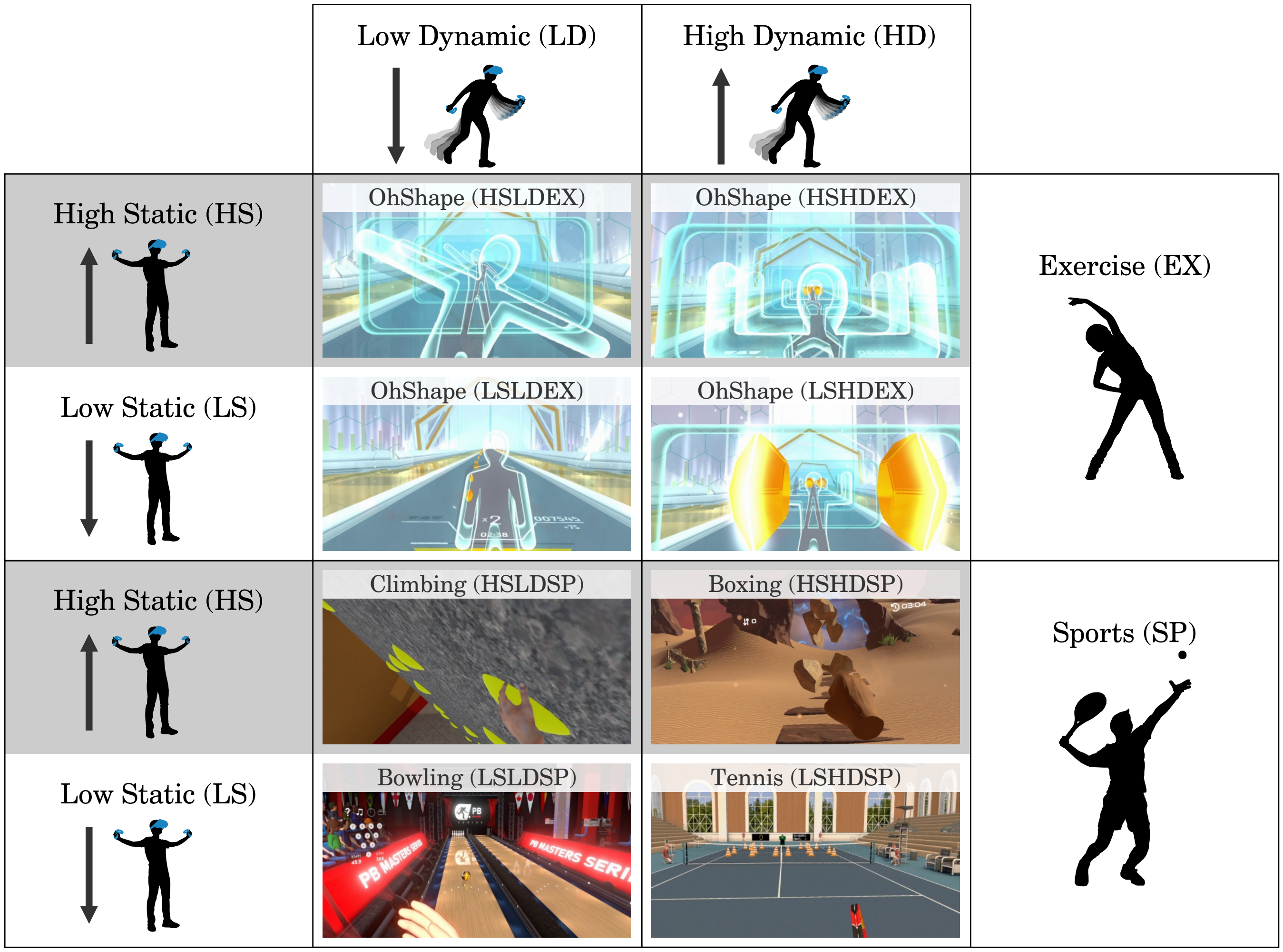


Figure 4: Overview of our independent variables, all implemented combinations, and assigned abbreviations. Each cell in the middle refers to an implemented activity. Abbreviations, concatenated from the variables, act as identifiers for each activity. The first independent variable is *Static Component*, which can take a level of either *Low Static (LS)* or *High Static (HS)*, and it is located to the very left. The second independent variable, *Dynamic Component*, can be realized through *Low Dynamic (LD)* or *High Dynamic (HD)* and resides at the top. At last, we investigate these variables either in the context of *Exercise (EX)* or *Sport (SP)*, depicted at the right.

"triad_openvr" API[4] and receives HMD data in real time. Our logger is manually operated. It features a graphical user interface, and the operator can set markers through keyboard presses to tag specific parts of the captured movement data in real time. The logger additionally draws a timestamp on the monitor, and we used a desktop recorder software to capture the contents of the monitor and of the rendered VR scene. Additionally, we placed a camera in the room where the study took place. All cameras had a synchronized timestamp in view so that any event was comprehensible afterward by inspecting the recordings. We release the source code of the logger[5].

3.3.2 Activities. The second part of our apparatus consists of the five applications we acquired through Steam. We verified that they implemented the necessary APIs for our logger.

Exercises. For *Exercise (EX)*, we chose the game OhShape[6]. OhShape is a rhythm game in which the player is confronted with approaching walls and coins. The walls have cutouts in the form of a human doing different poses, and the coins need to be collected with the

[4]Triad OpenVR Python Wrapper on Github.com https://github.com/TriadSemi/triad_openvr, last retrieved on March 8, 2024.

[5]The source code of our multi-purpose logger can be obtained from http://kinetic-signatures.hcigroup.de.

[6]OhShape on steampowered.com. https://store.steampowered.com/app/1098100/OhShape, last retrieved on March 8, 2024.

controllers. The player has to adjust their shape to fit through the cutouts and catch any approaching coin. The game features a level editor[7] with which we designed four custom levels in the game by creating the walls and coins. Each exercise was three minutes long and contained different types of static and dynamic self-exercise movements. Participants could try each activity for one minute to familiarize themselves with it. We always recentered participants' positions after placing them in the middle of the playspace.

Since we invited the head of university sports at our institution to collaborate, we designed the levels according to her feedback in an iterative process, creating self-exercises (*EX*) from static and dynamic movement components. For *LSLDEX* we created a mixture of short squats to perform in addition to coins that are easy to grab over time. In *HSLDEX* our level consisted of various T-poses together with inclinations to the left and right direction of the upper body. *HSHDEX* was an extension of the previous level, as it combined T-poses and inclinations of the upper body to the left and right direction with a butterfly movement of the arms. At last, for *LSHDEX*, our level required butterfly movements for the arms together with sidestep movements. The auxiliary material of this paper contains a Video Figure that shows a recording of our created levels and Figure 5 provides illustrations.

Sports. To select the *Sport (SP)* games, we used the classification of Mitchell et al. (cf., Figure 3(b)) and picked one game per cell [39, 40]. Each game was played for three minutes. Again, participants could try each game for one minute to become familiar. For all games, we placed participants in the middle of the playspace and recentered their positions before they started playing. We implemented *LSLDSP* by choosing "Premium Bowling"[8], which aims to provide a realistic bowling experience in VR. We selected the local single-player mode and the tournament track map. Then, they started playing bowling. To enable the segmentation of the data stream into individual kinetic signatures, we logged when participants commenced the throwing movements of the bowling ball and when it hit the pins. For *HSLDSP*, we selected climbing and used "Indoor Rock Climbing VR"[9] as a simulator. Here, participants found themselves in a climbing gym and were asked to climb up a large wall. They could climb in any style they liked, and we logged when they first touched the wall, then each grip on the wall, and if they fell or climbed down.

Next, we selected boxing as the activity for *HSHDSP*. We chose "PowerBeatsVR"[10] to be a suitable rhythm-based boxing game. Participants stood in front of a portal and had to punch virtual rocks approaching them with their hands. The level was set to "Beast but no least" on "beginner" difficulty. Here, we logged when the rocks to hit appeared in front of them and when virtual, laser-ray-like guidance lines appeared that participants would follow, imitating a static blocking move. At last, for *LSHDSP* our participants played "First Person Tennis"[11]. We set the game to the "training & forehand" mode and selected the handedness according to participants' true handedness. They then found themselves on a tennis court, could move around, and had to hit a ball that a ball machine launched at them. We logged when the ball was launched and when participants hit the ball back with their virtual racket. Again, we provide a Video Figure in the auxiliary material of this paper that shows recordings within and outside VR of the games. Figure 5 provides screenshots.

[7]OhShapeEditor on Github.com. https://github.com/OddersLab/OhShapeEditor, last retrieved on March 8, 2024.

[8]Premium Bowling on steampowered.com. https://store.steampowered.com/app/898580/Premium_Bowling, last retrieved on March 8, 2024.

[9]Indoor Rock Climbing VR on steampowered.com. https://store.steampowered.com/app/767330/Indoor_Rock_Climbing_VR, last retrieved on March 8, 2024.

[10]PowerBeatsVR on steampowered.com. https://store.steampowered.com/app/810500/PowerBeatsVR_VR_Fitness, last retrieved on March 8, 2024.

3.4 Power Analysis

To determine our required sample size, we conducted an a priori power analysis for a repeated-measures within-factors analysis of variance (ANOVA), given $\alpha = .05$, $\beta = .95$, and effect size $f = .4$. Following our experimental study design, we set the number of groups to eight and the number of measurements to two. G*Power 3.1.9.7 suggested a total sample size of 24 ($\lambda = 15.36, F = 4.50$).

3.5 Participants

We recruited 24 volunteers (7 female, 17 male) via University mailing list and social media (mean age = 26.83, SD = 3.20, 21 were right-handed). We asked participants for their previous VR experience ("I used VR often before participating in this study.") on a Likert item from 1 (completely disagree) to 7 (completely agree) that they then rated at an average of 3.56 (SD=2.30). Also, we asked them to reflect on their previous VR usage ("I have lots of experience with VR.") on a Likert item from 1 (never) to 7 (daily), and we obtained an average of 3.52 (SD=2.32). Hence, our participants were of similar age and had a strongly varying VR experience, with a majority being inexperienced.

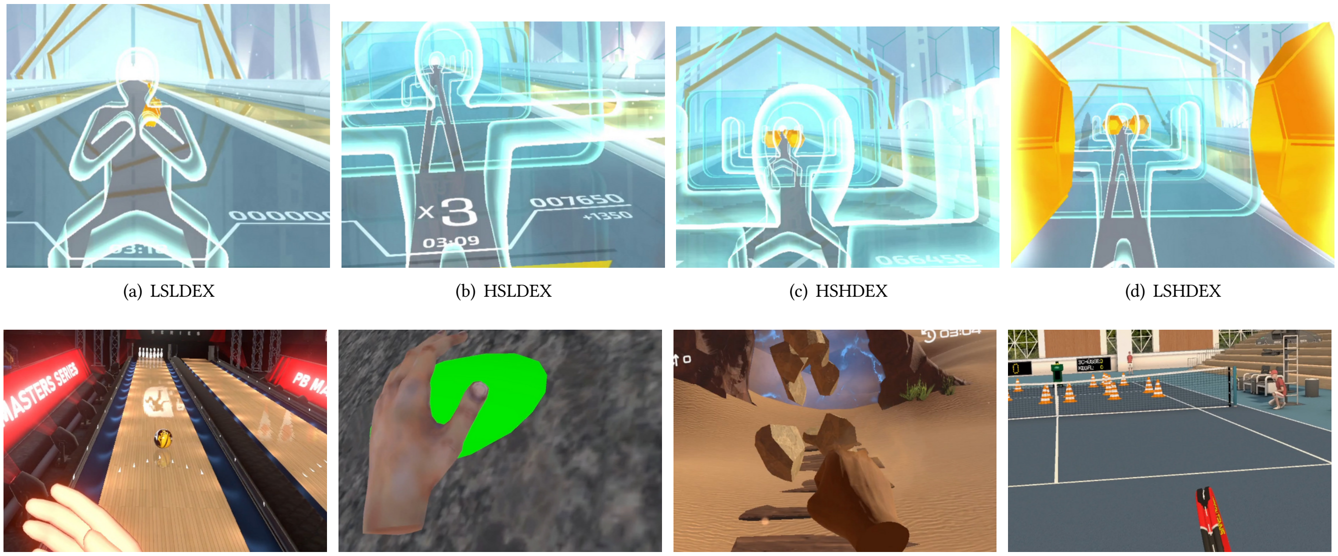
3.6 Procedure

First, we welcomed participants and explained the procedure to them upon their arrival. Next, we ensured that all questions they had about the study's progression were fully addressed. Then, we obtained participants' written and informed consent. Furthermore, we assured participants that they could, at any point in time, revoke their consent to participate in the study without any detriments. We then provided them a short explanation of the VR system if they were unfamiliar and assisted them in putting on the HMD by helping them adjust the straps and inter-pupillary distance. We then tested two blocks. The first block consisted of the *Exercise (EX)* activities, and the second block was the *Sport (SP)* activities. Each block consisted of four conditions that implemented all combinations of our independent variables' levels (cf., Figure 4). Within each block, we randomized the order of activities using a Latin Square. Upon starting the activity, participants could try every activity for one minute to familiarize themselves. Then, we tested each condition for a total of three minutes. In total, each session took approx. 60 minutes per participant. The session was then repeated one week later using the same procedure, meaning that participants played the exact same levels in each condition for the same duration, again.

3.6.1 *Study Setting.*

We conducted our user study in a seminar room with a large play area of $8m \times 6m$. The room was air-conditioned

[11]First Person Tennis on steampowered.com. https://store.steampowered.com/app/454140/First_Person_Tennis_The_Real_Tennis_Simulator, last retrieved on March 8, 2024.



(a) LSLDEX (b) HSLDEX (c) HSHDEX (d) LSHDEX

(e) LSLDSP (f) HSLDSP (g) HSHDSP (h) LSHDSP

Figure 5: Screenshots of our apparatus for *Exercise (EX)* in (a) to (d) and *Sport (SP)* in (e) to (h). Following Figures 3(b) and 4, *LSLDSP* was implemented through bowling in (e), *HSLDSP* with climbing in (f), *HSHDSP* with boxing in (g) and *LSHDSP* with tennis in (h). For *Exercise (EX)*, all levels were implemented in “OhShape”, shown in (a) to (d). A Video Figure is available in the supplementary material, showcasing each activity within VR and in reality.

to provide an equal temperature-wise experience for all participants. As an HMD, we chose the widely available “Meta Quest 2” that does not require a wired connection to a computer. We used a host computer with an Intel i7-9700 CPU, 16GB RAM, and a Nvidia RTX 2060 GPU to execute the SteamVR runtime and our multi-purpose logger and streamed the applications to the HMD using Oculus Air Link and a WiFi connection. At all times during the study, two experimenters were available. One experimenter took care of the participants and externally observed their movements. Additionally, another experimenter took care of the host computer and operated the logger, being able to see the participants in reality and a rendering of the participants’ view in VR on the monitor simultaneously.

3.6.2 Questionnaires. After each activity in *Sport (SP)*, we asked two Likert items, L1 and L2, on a scale from 1 (completely disagree) to 7 (completely agree). L1 presented the statement “I found the movement very exerting”. For L2, the statement was “I found the movement in its execution very natural”. Additionally, we asked the NASA-Task Load Index (TLX) questionnaire [57, 58] to understand participants’ workload and to validate our apparatus using participants’ subjective feedback.

3.7 Ethics

Before conducting the study, we obtained the approval of our institution’s ethics committee. Furthermore, we obtained the informed consent of our participants, and we assured them that they could cancel their participation in the study at any time without any detriments. To further protect participants’ privacy, we created two sets of randomized IDs, where we assigned one ID to their data and the other ID to their metadata. After we finished the study, we deleted the mapping so that no backtracking is possible; hence, we conducted a full pseudonymization of their data. Additionally, we would like to emphasize that the present work presents an implicit identification procedure. Insofar as the contents of this research are used for the implementation of such a system, it is necessary that established ethical guidelines are followed. For example, users should know how their data is used, that an implicit identification system processes their data, and what the intended purpose of the data processing is. Likewise, their informed consent should be obtained.

4 ANALYSIS

We created a deep-learning-based identification system similar to previous work that identified our participants [29, 32]. We trained it exclusively with the data obtained from the first session of our study and tested it with data from the second session, which took place one week later. Using these results, we then applied quantitative statistics to explore the effects of dynamic and static movement components using a factorial analysis.

4.1 Data Set

Our data set consists of the output of our logger that we elicited during the study, which was initially one continuous data stream. It consists of our participants’ positional and rotational coordinates obtained from the HMD, timestamps, activity markers, identifiers of each activity, and participant identifiers. During the elicitation process, we observed participants both in reality and within VR since we utilized Oculus Air Link, which also rendered the VR view on the host computer’s monitor. During this observation in the study in the *SP* activities, we manually annotated participants' data by placing numerical markers into the stream. We then split the data by the markers into individual samples containing exactly one kinetic signature. For the *Sport (SP)* activities, one kinetic sample corresponds to one bowling throw, a single climbing move, one boxing punch, or one hit in tennis in the exercises. For the *Exercise (EX)* activities, where the levels were deterministic and linear, we segmented the data into individual kinetic signatures by applying a timestamp-based splitting algorithm. Here, a kinetic sample is one consistent motion that participants performed at a time, e.g., one performed squat or one executed butterfly arm movement.

Since certain metrics (e.g., accuracy) are sensitive to imbalanced data sets, we decided to keep our dataset strictly balanced per activity. Therefore, we first matched the number of samples between participants, meaning, that in each activity all participants contributed the same number of samples and no participant is over- or underrepresented. Next, we matched the number of samples that compose the training data set with the testing data set; therefore, every participant contributed the same number of samples per activity to the training and testing data set, so that both sets have the same cardinality per activity. Since the number of samples provided could vary during the elicitation in our user study, we downsampled participants' contributed kinetic signature samples per activity. To do so, we removed samples that occurred later in the activity stream until a match between all participants' contributions was acquired by choosing the minimum number provided by any participant for each activity. By doing so, every participant contributed 87 kinetic signature samples (hereafter referred to as "samples") for *LSLDEX*, 54 samples for *HSLDEX*, 98 samples for *HSHDEX*, and 98 samples for *LSHDEX*. Similarly, for *Sport (SP)*, participants contributed 12 throws of a bowling ball (*LSLDSP*), 45 climbing movements (*HSLDSP*), 247 boxing punches (*HSHDSP*), and 33 tennis hits (*LSHDSP*) to our data set. These numbers apply to both the training and testing data set (e.g., there is 12 throws of a bowling ball in the training data set and in the testing data set for every participant). Per activity, the number of contributed samples per participant for the training and testing data sets is equal, and no participant is over- or underrepresented.

In total, we elicited 14.1 GB of data from our user study. We publish our data set online to enable replication[12].

4.2 Data Split

Since the user study encompassed two sessions, we split the data by session and activity to perform a hold-out validation. Hence, the data set was partitioned into eight parts for training, containing exclusively the data from the first session, and eight other exclusive parts for testing, containing the data from the second session. We never mixed data across sessions. The data from the first session was then used to train a deep-learning model per activity, which was then tested with the data from the second session to determine the evaluation metrics within said activity. We refrained from performing a dedicated hyperparameter tuning; instead, we used the default parameters of the models that were established by previous literature [21, 32]. We chose a hold-out validation style to show that we can re-identify participants exclusively with their kinetic signatures from the second session after a one-week interval.

4.3 Preprocessing

We applied the same preprocessing to all of our data on the level of the kinetic signature samples, i.e., individually per executed movement after we split the data by the markers. For each kinetic sample, we initially removed any information not originating from the controllers or the HMD (i.e., we stripped timestamps, user identifiers, and all other metadata). Since we sampled participants' position and rotation, we remained with 21 columns of data, representing the positions (i.e., "pos.x", "pos.y", and "pos.z") per object and its rotation (i.e., "quat.x", "quat.y", "quat.z", and "quat.w").

In the next step, we subtracted all position values per sample from the very first value in the same column. Consequently, we obtained relative positions in relation to the beginning of each kinetic sample. This procedure stripped parts of the per-participant bias from the data, such as absolute differences in body height denoted by "pos.y". Also, if a participant changed their initial position, for instance, by taking a step back, their initial position was always reset to zero. Although spatiotemporal human behavior always depends on body physiology, we were aware that this processing could not remove all components of a user's physiology. However, previous studies showed that applying further transformations to normalize physiology led to higher identification rates [26]. With this in mind, we concluded our specific feature transformation to emphasize the component of human behavior in each kinetic sample. Figure 6 provides plots of a sample per activity from our participants after applying preprocessing that we also used for visual inspection.

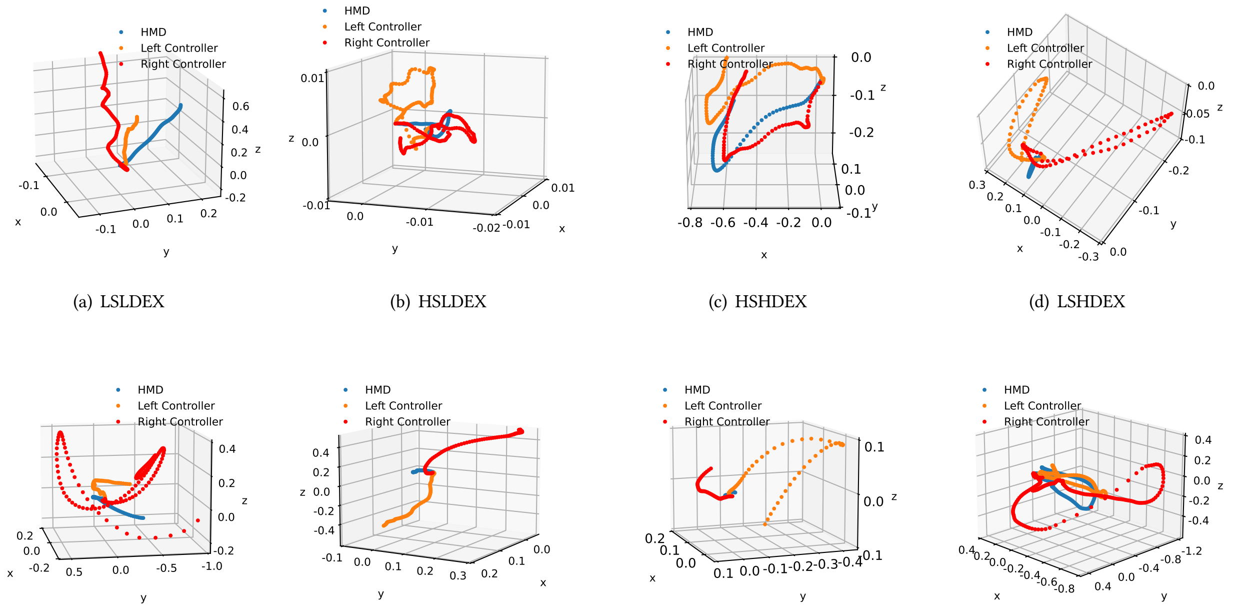
In the final step, we applied a min-max normalization to the data per kinetic sample since neural networks benefit from data being normalized into an interval of [0, 1] per column. We also applied padding and zero-based pre-padding to the mean length of the kinetic samples for each activity to unify their shape. Additionally, since we did not restrict the sampling rate of our logger, we downsampled the data to 60 Hz, as we saw that our logger logged faster than the HMD could stream new data, leading to interpolated duplicates. Each model was trained on the sliced kinetic samples obtained from the first session of the study. We then predicted the slices with the same parameters from the second session to elicit the evaluation metrics.

4.4 Deep Learning Model Architectures, Training, and Selection

We chose eleven different deep-learning model architectures for time-series classification from previous work to create an identification system. Since our elicited kinetic signatures are composed of spatiotemporal data, that is, human behavior sampled from an HMD over time, models for time-series classification are particularly suitable to create an identification system. We selected these models since they are founded in the review on time-series classification by Fawaz et al., and we also utilized their implementation with default hyperparameters[13] [21]. Some hyperparameters were

[12]Our data set can be obtained from http://kinetic-signatures.hcigroup.de.

[13]Implementation of Fawaz et al: "dl-4-tsc" on Github.com [21]. https://github.com/hfawaz/dl-4-tsc/tree/master/classifiers, last retrieved on March 8, 2024. Note: their CNN model can be parametrized by "same" and "valid" padding. We subsequently



(a) LSLDEX (b) HSLDEX (c) HSHDEX (d) LSHDEX

(e) LSLDSP (f) HSLDSP (g) HSHDSP (h) LSHDSP

Figure 6: Scatterplots of exemplary kinetic signatures elicited in our experiment from participants’ worn HMD, right and left controllers after applying the preprocessing. The preprocessing sets all position coordinates to zero at the beginning of each sample. Each scatterplot belongs to one activity.

dynamically set by their implementation and we did not alter this behavior. The proposed model architectures of Fawaz et al. were previously applied in several projects on behavioral biometrics by different researchers [29, 32, 70]. Initially, we trained all models proposed by Fawaz et al. since we did not know which of the models would converge and how their performances would behave [21]. We trained all models per activity with the data from the first session after it was preprocessed, i.e., we split the dataset by activity to avoid mixing kinetic signatures between activities. Then, we tested the models with participants’ data from the second session, again per activity, so that kinetic signatures would not be mixed and determined the per-participant performance recall rates of the models.

4.5 Quantitative Factorial Analysis

We round our analysis off with a quantitative analysis of our models’ performance rates to determine the effect of our two independent variables on the identification rates. Here, we chose the “recall rate”, which is defined as the number of true positives divided by the sum of true positives and false negatives per participant (i.e., $\frac{TP}{TP+FN}$). Thus, recall rate is the quotient of the true identity assignments divided by the sum of true and false identity assignments. It essentially tells how well a person can be identified by an identification system and was also used in previous work [27].

treated both settings as individual models, denoted as CNN (V) for “valid” padding and CNN (S) for “same” padding.

To understand the influences of our independent variables *Static Component* and *Dynamic Component* onto recall rate, we conducted two separate two-way repeated measures analysis of variance (RM-ANOVA). We chose to split our data set by modality for the analysis (*Exercise (EX)* vs. *Sport (SP)*) since it is not reflected in our hypotheses to compare across modality, i.e., we did not seek to compare the recall rates obtained in *Exercise (EX)* against recall rates in *Sport (SP)*. We proceeded this way, as we assumed that the change of modality has too many factors to be considered. Factors stem from the implementation, such as being within only a single application (OhShape) vs. being in multiple applications (the different sports games), as every application considerably changes the lightning, movement, visual effects, sound effects, and many more details. In addition, the human factor also played a major role. While our participants, on one hand, performed self-exercises, which is a very well-defined task, the contextual rules of the task fundamentally changed in *Sport (SP)*, depending on the respective game. Due to the design of OhShape, their movements within this application were strongly guided, as the walls and coins required certain poses to be performed and thus told the participants how to behave. For the sports games, participants’ degrees of freedom were way higher than OhShape, as they could freely decide how they act in each sport. For example, we did not enforce which way participants climbed, in which direction they moved during tennis, or how they interacted with their racket. Thereby, we believe that the four different *Exercise (EX)* can hardly be compared to four different *Sport (SP)*, where each *Sport (SP)* has its own set of rules defining the task. All in all, both the contextual requirements of the task for the human to perform and the implementation stayed constant for the different activities in *Exercise (EX)* but differed per *Sport (SP)*. Hence, we chose to split the analysis.

5 RESULTS

First, we selected models to be used within our analysis. The first part of the analysis then explored *Exercise (EX)* and the second part explored *Sport (SP)*.

5.1 Model Selection

To obtain our independent variable *recall rate*, we first trained our selected models with the respective data. Here, we always trained all model architectures per activity and assessed their performance. The resulting accuracy ratings are listed in Table 1. We averaged their results per modality and found that Inception works best for *Exercise (EX)* at a mean accuracy of 71.74% and FCN worked best for *SP* at a mean accuracy of 74.93%. While there is little difference between these two model architectures within *EX* and *SP* respectively, the overall mean performance degrades between *Exercise (EX)* and *Sport (SP)*, as *SP* performs approximately three to four percent better than *EX* (cf., Table 1). We then obtained the respective recall rate from each activity's models. Figure 10 in Appendix A shows each model's confusion matrix, providing insights into its biometric identification performance and Figure 7 depicts a histogram of the obtained recall rates per modality.

5.2 Kinetic Signature's Identifiability in Exercises

To explore the influence of the two independent variables, *Static Component* and *Dynamic Component* onto recall rate, we conducted a two-way factorial RM-ANOVA. *Static Component* and *Dynamic Component* had two respective levels, *Low Static (LS)* and *High Static (HS)* for the former, and *Low Dynamic (LD)* and *High Dynamic (HD)* for the latter.

5.2.1 Assumptions. At first, we checked the assumptions of RM-ANOVA. We conducted Levene's test for the homogeneity of variances across groups (homoscedasticity), which did not indicate that our groups have significant differences in variance (F(3, 92) = .4196, p = .7394). The sphericity assumption was necessarily met due to the present levels of our design. At last, we checked if our data samples were normally distributed using the Shapiro-Wilk test, which indicated that they were not normally distributed (W = .9044, p < .0001). This was confirmed in a histogram as shown in Figure 7(a). Therefore, we applied the aligned rank transform (ART) procedure before running the RM-ANOVA [71].

5.2.2 Main Effects. In the analysis of *Exercise (EX)*, we found a statistically significant main effect for *Dynamic Component* (F(1, 69) = 4.8421, p = .0311, η_p^2 = .0656). *High Dynamic (HD)* overall yields a higher recall rate (Med = .86, IQR = .29) compared to *Low Dynamic (LD)* (Med = .72, IQR = .31), supporting **H1(a)**. We did not find a significant main effect for *Static Component* (F(1, 69) = 1.9517, p = .1669), not supporting **H2(a)**. Additionally, the interaction effect between *Static Component* and *Dynamic Component* was not significant (F(1, 69) = 1.9694, p = .1650). Figure 7(b) provides a boxplot of the significant main effect in *Exercise (EX)*.

5.3 Kinetic Signature's Identifiability in Sports

We continued with the analysis of *Static Component* and *Dynamic Component* in *Sport (SP)*, where we again conducted a RM-ANOVA analogously to the analysis of both variables in *Exercise (EX)*.

5.3.1 Assumptions. Again, we conducted Levene's test for the homogeneity of variances across groups (homoscedasticity). As before, it did not indicate that our groups have significant differences in variance (F(3, 92) = 1.5328, p = .2113). Sphericity, again, was necessarily met. We then continued checking if our data would be normal distributed using a Shapiro-Wilk test, which again indicated that the recall rate values follow a non-normal distribution (W = .8788, p < .0001, cf., Figure 7(a)). Therefore, we once more applied the ART procedure [71].

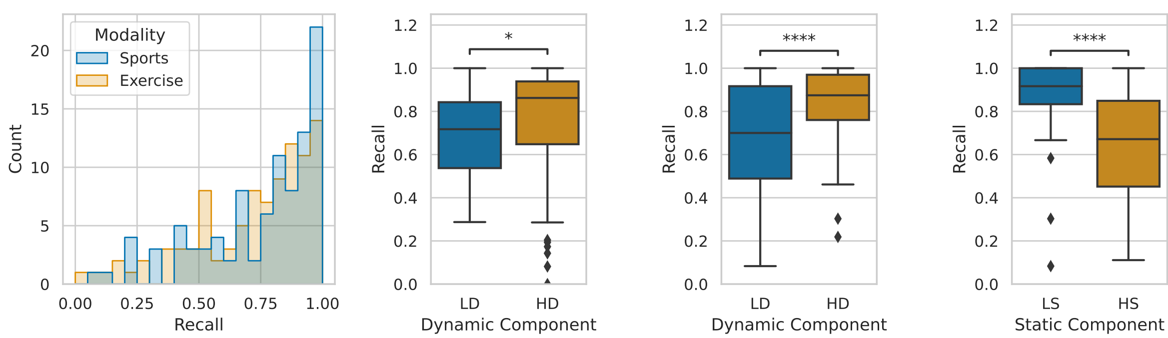
5.3.2 Main Effects. In *Sport (SP)*, we found two significant main effects, one for *Static Component* and one for *Dynamic Component*, respectively (cf., Figures 7(c) and 7(d)). We found a significant main effect for *Static Component* (F(1, 69) = 66.9425, p < .0001, η_p^2 = .4924). *High Static (HS)* (Med = .67, IQR = .40) lead to lower recall rate values than *Low Static (LS)* (Med = .92, IQR = .17), supporting **H2(b)**. Furthermore, we found a significant main effect for *Dynamic Component*, where *High Dynamic (HD)* (Med = .87, IQR = .21) lead to higher values than *Low Dynamic (LD)* (Med = .70, IQR = .43), F(1, 69) = 24.4615, p < .0001, η_p^2 = .2617, supporting **H1(b)**. However, the interaction between *Static Component* and *Dynamic Component* also had a significant effect (F(1, 69) = 13.0590, p = .0006, η_p^2 = .1591). Thus, it is important to explore the interaction between both.

5.3.3 Interaction Effect between Static and Dynamic Component. We found that five out of six possible pair-wise contrast tests were significant after p-value correction using Holm's method for the interaction effect between *Static Component* and *Dynamic Component*. First, *HSHD* (Med = .80, IQR = .24) yielded higher recall rates than *HSLD* (Med = .51, IQR = .29), t(69) = 4.1123, p = .0004, supporting **H1(b)**. Also, *HSHD* (Med = .80, IQR = .24) resulted in lower values than *LSHD* (Med = .95, IQR = .15), t(69) = -3.7998, p = .0009, supporting **H2(b)**. *HSHD* (Med = .80, IQR = .24) was lower than *LSLD* (Med = .92, IQR = .19), t(69) = -2.7978, p = .0133, supporting **H2(b)**, but opposing **H1(b)**. On the other hand, *HSLD* (Med = .51, IQR = .29) was lower than *LSHD* (Med = .95, IQR = .15), t(69) = -7.9121, p < .0001, supporting **H1(b)** and **H2(b)**. Furthermore, *HSLD* (Med = .51, IQR = .29) was lower than *LSLD* (Med = .92, IQR = .19), t(69) = -6.9100, p < .0001, supporting **H2(b)**. At last, the comparison of *LSHD* (Med = .95, IQR = .15) with *LSLD* (Med = .92, IQR = .19) was not significant (t(69) = 1.0020, p = .3198), thus **H1(b)** remained unsupported for this contrast.

To summarize this interaction effect, it becomes apparent that **H1(b)** and **H2(b)** were mostly supported, with two exceptions. First, as we could not find a significant difference between *LSLD* and *LSHD* (cf., fourth significance indicators from the top in Figure 8(a)), so **H1(b)** remained unsupported for this case. Second, *HSHD* in comparison with *LSLD* supported **H2(b)**, as *HS* lead to lower identifiability than *LS*, yet it opposed **H1(b)** in this contrast,

Table 1: Accuracy ratings in percent of all trained models per activity, sorted by the last column. $\overline{\mathrm{ACC}_{\mathrm{EX}}}$ is the mean of the four trained models for *Exercise (EX)* per architecture and $\overline{\mathrm{ACC}_{\mathrm{SP}}}$ for all *Sport (SP)* activities and its four trained models per architecture (cf., Figure 4). The best model for *EX* is Inception. For *SP* it is FCN. The highest reached mean accuracies per *EX* and *SP* are respectively marked with an asterisk (*).

Model	LSLDEX	HSLDEX	HSHDEX	LSHDEX	LSLDSP	HSLDSP	HSHDSP	LSHDSP	$\overline{\mathrm{ACC}_{\mathrm{EX}}}$	$\overline{\mathrm{ACC}_{\mathrm{SP}}}$
FCN	72.08	69.21	68.75	74.57	84.72	50.65	74.71	89.65	71.15	*74.93
Inception	69.68	69.83	69.30	78.15	82.29	50.46	73.90	90.91	*71.74	74.39
ResNet	69.64	67.82	67.35	73.21	81.25	45.65	73.97	89.77	69.51	72.66
Encoder	46.17	49.23	61.78	65.86	80.56	34.44	56.48	73.48	55.76	61.24
TWIESN	42.53	55.63	54.63	61.14	68.06	36.48	48.75	81.94	53.48	58.81
MCDCNN	56.80	55.94	56.04	55.14	63.89	37.41	50.96	74.49	55.98	56.69
MLP	46.17	48.07	51.79	49.15	63.19	27.87	50.46	77.02	48.79	54.64
CNN (S)	32.33	38.89	42.18	54.63	70.14	24.35	44.96	66.04	42.01	51.37
CNN (V)	32.38	32.72	41.96	48.94	59.72	24.07	41.28	67.30	39.00	48.09
MCNN	4.17	4.17	4.17	4.17	4.17	4.17	4.17	4.17	4.17	4.17
t-leNet	4.17	4.17	4.17	4.17	4.17	4.17	4.17	4.17	4.17	4.17



(a) (b) Dynamic Main Effect in *EX*. (c) Dynamic Main Effect in *SP*. (d) Static Main Effect in *SP*.

Figure 7: Histogram of the distribution of recall rate values split by modality in (a). Then, (b) shows the significant main effect within *EX*. For *SP*, (c) and (d) show the two significant main effects for *Dynamic Component* and *Static Component*, respectively. Annotated asterisks denote significance levels: **** = p < .0001, *** = p < .001, ** = p < .01, * = p < .05, ns = not significant.

since *HD* led to lower values than *LD*. This is also depicted in Figure 8(a), where the comparison of the third and fourth boxplot (enumerated from left to right) turned out significant. Figure 8(b) provides an interaction plot. This can mean that **H2(b)** dominates **H1(b)** in this comparison, as both factors were altered. We saw in the main effects that η_p^2 was .4924 for *Static Component* related to **H2(b)** and .2617 for *Dynamic Component's* main effect, related to **H1(b)**. This indicates that **H2(b)** outweighs **H1(b)** in this pairwise contrast test, as it is impossible for both to remain supported in this specific comparison.

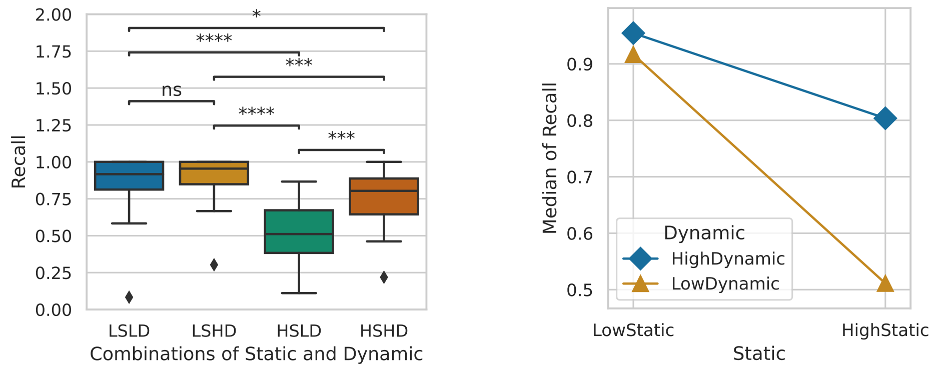
5.4 Questionnaires

We conducted the raw NASA-Task Load Index (RTLX) questionnaire to obtain our participants' feedback and validate our apparatus based on their responses. Besides the RTLX, we asked them two individual Likert Items, L1 and L2.

5.4.1 NASA TLX. We asked participants for their task load after every activity, using the RTLX [57, 58]. We did so to understand whether we could find differences in task load between our eight different activities, i.e., if the order of activities sorted by task-load in *Exercise (EX)* differs from the activities in *Sport (SP)*. First, we calculated participants' median scores for each activity in *Exercise (EX)* and *Sport (SP)*. Next, we ranked the activities per modality.

Exercises. We find the following ranking for the four different activities within *Exercise (EX)*, in descending order: (1) *HSHDEX* (Med = 35.83, IQR = 28.33), (2) *LSHDEX* (Med = 31.25, IQR = 24.17), (3) *HSLDEX* (Med = 27.08, IQR = 14.38), and (4) *LSLDEX* (Med = 13.33, IQR = 11.25).

Friedman's test for ordinal data revealed a significant effect of group ($\chi^2(3)$ = 53.67, p = .0001, N = 24). The post-hoc pairwise comparisons using Wilcoxon signed-rank test showed that *LSLDEX* lead to significantly lower scores than *HSLDEX* (W = 3.5, Z = -4.1861, p = .0001, r = .6042). *LSLDEX* also lead to significantly smaller scores than *HSHDEX* (W = 0, Z = -4.2864, p = .0001, r = .6187). Next, *LSLDEX* also yielded significantly lower values than *LSHDEX* (W = 5, Z = -4.1429, p = .0001, r = .5980). Furthermore, *HSLDEX* provided smaller



(a) Pairwise Contrast Tests in *SP*.

(b) Plotted Interaction Effect in *SP*.

Figure 8: The boxplots in (a) show the contrasts for the interaction effect between *Static Component* and *Dynamic Component* in *SP*. Annotated asterisks denote significance levels: **** = p < .0001, *** = p < .001, ** = p < .01, * = p < .05, ns = not significant. The interaction plot in (b) shows the interaction effect between *Static Component* and *Dynamic Component* in *Sport (SP)*.

scores *HSHDEX* (W = 13, Z = -3.8023, p = .0001, r = .5489). *HSLDEX* also yielded smaller scores than *LSHDEX* (W = 53, Z = -2.5857, p = .0483, r = .3732). However, only for the scores in *HSHDEX* vs. *LSHDEX*, we could not find any significant differences (W = 202.5, Z = 1.96210, p = .2960). The p-values of the pairwise comparisons were adjusted using Bonferroni's method.

Sports. We proceeded analogously for *Sport (SP)* and found the following ranking in descending order: (1) *HSHDSP* (Med = 38.75, IQR = 19.38), (2) *LSHDSP* (Med = 36.25, IQR = 20.625), (3) *HSLDSP* (Med = 23.33, IQR = 25.63), and (4) *LSLDSP* (Med = 21.25, IQR = 12.50).

Again, we conducted a Friedman's test and saw that there was a significant effect of group ($\chi^2(3)$ = 29.5277, p = .0001, N = 24). Next, we conducted pairwise comparisons using Wilcoxon's signed-rank test. *LSLDSP* lead to smaller TLX scores than *HSLDSP*, but the comparison was not significant (W = 108.5, Z = -1.1858, p = 1). *LSLDSP* resulted in significantly smaller values than *HSHDSP* (W = 2, Z = -4.1369, p < .0001, r = .5971). For *LSLDSP* we also found that its values were significantly smaller than *LSHDSP* (W = 29.5, Z = -3.4432, p = .0012, r = .4970). When comparing *HSLDSP*, we once more found significantly smaller values compared to *HSHDSP* (W = 37, Z = -3.2287, p = .0038, r = .4660). For *HSLDSP* vs. *LSHDSP*, we could not find significant differences (W = 63, Z = -1.8249, p = .4136). Comparing *HSHDSP* with *LSHDSP* yielded significantly smaller values, again (W = 187.5, Z = 1.5065, p = .8155). The p-values of the pairwise comparisons were adjusted using Bonferroni's method.

Here, we find that our participants' ranked the activities in *Exercise (EX)* and *Sport (SP)* in the same order by task-load. We mostly found significant differences, with only a few exceptions. As the overall ranking is identical for *Exercise (EX)* and *Sport (SP)*, we assumed, that participants mostly experienced identical relative task loads by the corresponding activities, supporting the validity of our apparatus through their subjective feedback.

5.4.2 Individual Likert Items. At last, we also asked participants for their responses to two individual Likert items after every activity. "L1" presented the statement "I found the movement very exerting" on a scale from 1 (completely disagree) to 7 (completely agree). The second item, "L2", requested their rating on the statement "I found the movement in its execution very natural", on the same scale. Figure 9 provides a boxplot with participants' results. Overall, participants' responses did not find the activities exceptionally exerting, with only *HSHDEX* yielding the highest exertion levels at a median of three. Other than that, participants mostly rated their movements as natural, with *LSLDSP* (bowling) being the most natural.

6 DISCUSSION

In this section, we discuss the (1) overall identifiability of kinetic signatures, (2) the influences of static and dynamic movement components, (3) the applicability of our hypotheses outside of VR, (4) the implicit nature of kinetic signatures and how they can ideally be employed in real-world applications, and (5) limitations of our investigation.

6.1 Overall Identifiability of Kinetic Signatures

Overall, kinetic signatures are well-suited for re-identification by deep-learning-based identification systems, even after some time passes. In our experiment, we reinvited participants after one week for their second session and exclusively used the first session for training, simulating identification with unseen data after one week. From Table 1, we can see that the identification accuracy per model varies per condition; however, the best results for *LSHDEX* were obtained by Inception, and for *LSHDSP* it was FCN at respective accuracy ratings of 78.15% and 90.91%. For the latter, the median recall rate was 95.45%.

In identification systems, the base chance of a correct random guess is $\frac{1}{N}$; thus, for our work, it is $\frac{1}{24}$ = 4.167%. Previous works that reinvited participants for a second session used N=16 participants with a base chance of 6.25% reported up to 90% as their highest model's accuracy when applying body normalizations [26], or 63.55% for N=19 [52] at a base chance of 5.26%. Thus, our best results exceed these values, given our decreased probability of a random guess and harder classification problem, as we had +26.32%

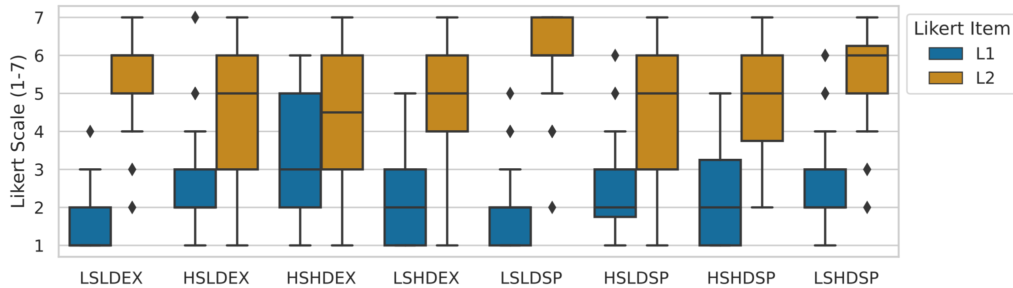


Figure 9: Participants' responses to the two individual Likert items on a scale from 1 (completely disagree) to 7 (completely agree). L1 was "I found the movement very exerting" and L2 was "I found the movement very natural".

participants compared to Pfeuffer et al. and +50% participants compared to Liebers et al. [26, 52]. However, we must acknowledge that we also had activities in our experiment that yielded considerably lower accuracy ratings (e.g., *HSLDSP*: climbing at 50.65%).

Concerning our apparatus, we saw from participants' subjective feedback in the RTLX that they ranked our activities within *Exercise (EX)* identically to the activities *Sport (SP)*, according to the required task-load per activity. Hence, we believe that our apparatus reflected the various levels of *Static Component* and *Dynamic Component* within *Exercise (EX)* and *Sport (SP)* well.

6.2 Influences of Static and Dynamic Movement Components on a Kinetic Signature's Identifiability

Static Components. Starting with hypothesis **H2**, stating that *High Static (HS)* yields lower identifiability in kinetic signatures compared to *Low Static (LS)*, we can acknowledge **H2(b)** for *Sport (SP)*. There, we found a significant main effect supporting **H2(b)** and a significant interaction effect that consistently supported **H2(b)** across all contrast tests in the interaction effect. This also confirms the findings of Liebers et al. [26], who also found significant differences between *LSLD* and *MSLD*. However, for *Exercise (EX)*, we found a trend in the data that this hypothesis might apply, but the differences were not statistically significant (**H2(a)**). Thus, we can only accept **H2(b)**, but cannot accept **H2(a)**.

Dynamic Components. For **H1**, stating that *High Dynamic (HD)* yields higher identifiability compared to *Low Dynamic (LD)*, the situation is more complex. First, **H1(a)** can be acknowledged throughout *Exercise (EX)*, as we found a significant main effect and no significant interaction effect. For *Sport (SP)*, we found a significant main effect, supporting **H1(b)**. Similarly, Pfeuffer et al. found that *HD* performed better than *LD* [52].

Albeit, in the pair-wise contrast test, we did not find significant differences between *LSHDSP* (sports: tennis) vs. *LSLDSP* (sports: bowling). The descriptive statistics reveal a trend in the correct direction, but the differences were insignificant (p = .3198). Both median values were very high (.92 and .95), as they belong to a scale of [0; 1], indicating that our system was highly capable of performing correct identifications in both activities. Thereby, **H1(b)** remains unsupported for this specific comparison but was supported in almost any other pairwise comparisons and, additionally, throughout the main effect in *Sport (SP)*. The last contrast test revealed that *HSHDSP* (sports: boxing) lead to significantly lower identifiability compared to *LSLDSP* (sports: bowling), opposing **H1(b)** for these two specific activities in *SP*.

However, it must be noted that **H1(b)** and **H2(b)** are mutually exclusive for this specific contrast test. If *LSLDSP* (bowling) would have led to significantly lower identifiability than *HSHDSP* (boxing), **H1(b)** would be supported but **H2(b)** would be opposed. Here, the opposite is the case, and we believe that moderation of the effects took place during their interaction, as both variables were altered in this comparison (*LSLD* vs. *HSHD*). Given both main effects in *Sport (SP)*, we find that *Static Component* has an η_p^2 of .4924 and *Dynamic Component* has an η_p^2 of .2617. Thus, the effect denoted by η_p^2 of *Static Component* is almost twice the value of *Dynamic Component*, and it could be possible that the change from *Low Static (LS)* to *High Static (HS)* outweighed the change from *High Dynamic (HD)* to *Low Dynamic (LD)*. Thus, we can accept **H1(a)**. Additionally, we overall can accept **H1(b)**, but it is moderated by **H2(b)**.

Conclusions. We accept **H1**, since we accepted both **H1(a)** and **H1(b)**. Our results indicate that *High Dynamic (HD)* components lead to significantly higher identifiability of kinetic signatures than *Low Dynamic (LD)* components. For **H2**, we could only accept **H2(b)** for *Sport (SP)*. We found no significant evidence for **H2(a)**, yet the data followed a supporting trend. Thus, we cannot universally accept **H2**, as we only have evidence within *Sport (SP)* that *High Static (HS)* leads to significantly lower identifiability than *Low Static (LS)*. Additionally, the interaction between *Static Component* and *Dynamic Component* shows that the former dominates the latter, meaning that *Low Static (LS)* increases the dentifiability of kinetic signature more than *High Dynamic (HD)*.

6.3 Reflecting on Literature Outside of VR

Abdrabou et al. explored a variety of eleven tasks using a motion-capture system in reality (i.e., not in VR, $N = 22$, cf., Section 2.4.1) [2]. They list accuracies between 6.74% to 13.88% for their *LSLD* activities, 28.82% to 40.75% for the *LSHD* activities, and 53.01% to 59.21% for *HSHD* activities in their Figure 5 [2]. Therefore, *High Dynamic (HD)* yielded higher ranked identifiability in their experiment than *Low Dynamic (LD)*, as indicated by **H1**. In their experiment, *High Static (HS)* is better than *Low Static (LS)*, opposing **H2**. However, this opposition happens during the interaction of *HS* and *HD* in *reality*, which is fundamentally different from a setup in VR [14]. For example, their participants carried real loads, likely resulting in a higher degree of isometric muscle contractions. In contrast, our participants did not experience any external weight in VR due to its inherent limitations [14]. Thus, the interaction of *High Static (HS)* and *High Dynamic (HD)* plays a role, in addition to the factor of *reality* vs. *virtual reality*, as it potentially influences the isometric efforts required in *HS*. Therefore, **H1** is confirmed by their results, but **H2** is not, while the general applicability of our hypotheses to a setup in reality (not in VR) can be questioned. Exploring these effects might be a promising research opportunity.

6.4 Implementing Implicit Identification using Kinetic Signatures

We explored implicit identification through kinetic signatures and various activities. A popular misconception is that these activities are meant as drop-in replacements for common authentication methods, such as a password entry in VR being replaced with a throw of a bowling ball. This is not the intended use case. Instead, our findings are meant to enable identification through implicit human-computer interaction, to identify users "through actions they would carry out anyway" [23, 61]. This means that the HMD can identify its user while they perform any suitable activity during an arbitrary point in time of their usage of the device and keep the verified information on its user's identity until the information is required. This relieves users of the burden of today's authentication methods [5, 59]. Additionally, the HMD can continuously re-identify its user. This is benefitted by an implicit identification capable of happening silently and transparently in the background, using implicit interactions [61]. As a consequence, security can be increased by employing a continuous authentication system [65, 66].

To employ our findings in a real-world system, it would be ideal to create an application where users at certain points perform highly dynamic movements with little static movements involved. Usually, users would perform these movements in response to the rendered virtuality; thereby, the VR application knows when such a context is applicable to the user. Once such a context is imposed, the application can scan the user's kinetic signatures. For this reason, triggers might be used, e.g., if an interaction of the user is required in response to the context before the actual context starts (e.g., if the context requires the user to dynamically throw an object, the beginning can be deducted from the user pressing the controller's gripper buttons). Then, the kinetic signatures can be obtained and used for identification through, for example, neural networks. Additionally, majority voting can be applied to multiple elicited kinetic signatures, which usually enhances the identification accuracy. Here, identification does not happen per individual kinetic signature but per the majority voting when identifying multiple signatures, often leading to more stable results [29].

6.5 Limitations

We acknowledge two limitations concerning our experiment. First, we cannot fully exclude any physiological factor in our behavioral biometrics research. Second, our investigation is limited by the lack of external forces and weight in VR.

Physiological Factors. Human behavior always depends on human physiology to some degree, as physiology limits how a movement can be executed. In our work, we did not apply any body normalization in VR; therefore, we rely on our preprocessing to remove most physiological factors [26]. For example, we effectively could remove body height as a primary bias, yet other factors, such as arm length, which correlate with body height, cannot be trivially removed. Therefore, we acknowledge that our behavioral data still reflects participants' physiology to a small extent.

Additionally, we had three left-handed participants in our experiment. This is representative of the population sample that we took, assuming that 10.6% of the general population are left-handed [51]. These three participants might have elicited behavior that is easier to identify compared to the right-handed participants' behavior. We believe that this reflects how identification systems work within the general population.

Weight in VR. One central limitation of our work is also the role of weight in VR. While dynamic movements and the associated isotonic muscle contractions can be easily performed in VR, isometric contractions are limited, as most consumer-ready VR cannot provide or apply external weight or forces to their user. However, isometric contractions can be used in VR by using a body's own weight (e.g., as in body-weight self-exercises). The application of external forces or external weights, however, is non-trivial although researchers provided case studies of such systems [14, 15] We did not employ such a solution, as they are not widely available. Therefore, we must acknowledge the limitation that our *High Static (HS)* components are impacted by taking place in VR. We believe that the impact is larger for *Sport (SP)*, as both *HSLDSP* (climbing) and *HSHDSP* (boxing) did not provide the feedback of weight to participants. For *Exercise (EX)*, we believe that the difference is less since participants performed self-exercises that were affected by their own body weight. However, given our participants' subjective feedback using raw NASA-Task Load Index (RTLX), we did not find a different ranking given their overall workload between *Exercise (EX)* and *Sport (SP)*.

7 CONCLUSION

In recent years, research showed that a wide range of distinct behaviors can uniquely identify users in VR. While this provides great starting points by showing that different types of movement and sports can be used for identification, it remained unclear what part of each behavior is responsible for the overall identification performance. In this work, we systematically investigated how different types of kinetic signatures influence user identification in VR. We based our work on an existing taxonomy that classifies sports in two dimensions, namely, static and dynamic [39]. We found that kinetic signatures with low static components are beneficial for user identification, followed by highly dynamic components, indicating that certain movement types elicit more individual behavior than others. Thus, this work contributes to a better understanding of what characterizes behavior to be particularly suited for user identification. Therefore, our findings help remove the burden of today's prevalent authentication methods for users in next-generation identification systems using implicit interactions and kinetic signatures. Moreover, our findings serve as a solid foundation for continued future systematic explorations of behavioral biometric systems.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to Ms. Katrin Hertel of the University Sports at the University of Duisburg-Essen for her extraordinary commitment, expertise and support of our research project. The presented work was funded by the German Research Foundation (DFG) under project numbers 426052422 and 425869382.

REFERENCES

[1] A. Ajit, N. Banerjee, and S. Banerjee. 2019. Combining Pairwise Feature Matches from Device Trajectories for Biometric Authentication in Virtual Reality Environments. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE Computer Society, Los Alamitos, CA, USA, 9–97. https://doi.org/10.1109/AIVR46125.2019.00012

[2] Yasmeen Abdrabou, Lukas Mecke, Radiah Rivu, Sarah Prange, Quy Dat Nguyen, Vanessa Voigt, Florian Alt, and Ken Pfeuffer. 2023. How Unique do we Move? Understanding the Human Body and Context Factors for User Identification. In *Mensch und Computer 2023*, Markus Stolze, Frieder Loch, Matthias Baldauf, Florian Alt, Christina Schneegass, Thomas Kosch, Teresa Hirzle, Shadan Sadeghian, Fiona Draxler, Kenan Bektas, Katrin Lohan, and Pascal Knierim (Eds.). ACM, New York, NY, USA, 127–137. https://doi.org/10.1145/3603555.3603574

[3] Bruce Abernethy, Vaughan Kippers, Laurel Traeger Mackinnon, Robert J. Neal, and Stephanie Hanrahan. 1997. *The Biophysical Foundations of Human Movement*. Human Kinetics, Champaign, Illinois, USA.

[4] Anne Adams and Martina Angela Sasse. 1999. Users Are Not the Enemy. *Communications of the ACM* 42, 12 (1999), 40–46. https://doi.org/10.1145/322796.322806

[5] Florian Alt and Stefan Schneegass. 2022. Beyond Passwords—Challenges and Opportunities of Future Authentication. *IEEE Security & Privacy* 20, 1 (2022), 82–86. https://doi.org/10.1109/MSEC.2021.3127459

[6] Muhammad Asif and John Krogstie. 2012. Taxonomy of personalization in mobile services. In *Proceedings of 10th IADIS International Conference e-Society*. IADIS Press, Berlin, Germany, 343–350.

[7] Mirjam Augstein, Eelco Herder, and Wolfgang Wörndl (Eds.). 2019. *Personalized human-computer interaction*. De Gruyter, Berlin and Boston. https://doi.org/10.1515/9783110552485

[8] J. Gordon Betts, Kelly A. Young, James A. Wise, Eddie Johnson, Brandon Poe, Dean H. Kruse, Oksana Korol, Jody E. Johnson, Mark Womble, and Peter DeSaix et al. 2022. *Anatomy and Physiology 2e*. OpenStax, Chapter 10.4 Nervous System Control of Muscle Tension. Available online at https://openstax.org/books/anatomy-and-physiology-2e/pages/10-4-nervous-system-control-of-muscle-tension. Access for free at http://openstax.org, last retrieved on March 8, 2024.

[9] J. Gordon Betts, Kelly A. Young, James A. Wise, Eddie Johnson, Brandon Poe, Dean H. Kruse, Oksana Korol, Jody E. Johnson, Mark Womble, and Peter DeSaix et al. 2022. *Anatomy and Physiology 2e*. OpenStax, Chapter 10.4 Nervous System Control of Muscle Tension, Figure 10.13. Figure licensed under CC-BY-4.0. Figure available online at https://openstax.org/books/anatomy-and-physiology-2e/pages/10-4-nervous-system-control-of-muscle-tension#fig-ch10_04_01. Access for free at http://openstax.org. Figure was adapted into a horizontal layout by splitting it into Figures 2(a), 2(b), and 3(a) of this work and by changing its color to greyscale. Hyperlinks last retrieved on March 8, 2024.

[10] Joseph Bonneau, Cormac Herley, Paul C. van Oorschot, and Frank Stajano. 2015. Passwords and the evolution of imperfect authentication. *Communications of the ACM* 58, 7 (2015), 78–87. https://doi.org/10.1145/2699390

[11] Sascha Brostoff and M. Angela Sasse. 2003. "Ten strikes and you're out": Increasing the number of login attempts can improve password usability. In *Workshop on Human-Computer Interaction and Security Systems at CHI 2003*. ACM, Fort Lauderdale, Florida, 4 pages. https://discovery.ucl.ac.uk/id/eprint/19826/

[12] Ceenu George, Mohamed Khamis, Emanuel von Zezschwitz, Henri Schmidt, Marinus Burger, Florian Alt, and Heinrich Hu\textsectionmann. 2017. Seamless and Secure VR: Adapting and Evaluating Established Authentication Systems for Virtual Reality. In *Proceedings 2017 Workshop on Usable Security*. Internet Society, San Diego, CA, USA, 12 pages. https://doi.org/10.14722/usec.2017.23028

[13] James E. Cutting and Lynn T. Kozlowski. 1977. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society* 9, 5 (1977), 353–356. https://doi.org/10.3758/BF03337021

[14] Sarah Faltaous, Marvin Prochazka, Jonas Auda, Jonas Keppel, Nick Wittig, Uwe Gruenefeld, and Stefan Schneegass. 2022. Give Weight to VR: Manipulating Users' Perception of Weight in Virtual Reality with Electric Muscle Stimulation. In *Mensch und Computer 2022*, Max Mühlhäuser, Christian Reuter, Bastian Pfleging, Thomas Kosch, Andrii Matviienko, Kathrin Gerling, Sven Mayer, Wilko Heuten, Tanja Döring, Florian Müller, and Martin Schmitz (Eds.). ACM, New York, NY, USA, 533–538. https://doi.org/10.1145/3543758.3547571

[15] Cathy Mengying Fang and Chris Harrison. 2021. Retargeted Self-Haptics for Increased Immersion in VR without Instrumentation. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, Jeffrey Nichols, Ranjitha Kumar, and Michael Nebeling (Eds.). ACM, New York, NY, USA, 1109–1121. https://doi.org/10.1145/3472749.3474810

[16] David Ferbrache. 2016. Passwords are broken – the future shape of biometrics. *Biometric Technology Today* 2016, 3 (2016), 5–7. https://doi.org/10.1016/S0969-4765(16)30049-2

[17] David L. Gallahue, Jackie Goodway, and John C. Ozmun. 2020. *Understanding motor development: Infants, children, adolescents, adults* (eighth edition ed.). Jones & Bartlett Learning, Burlington MA.

[18] Gonzalo Munilla Garrido, Vivek Nair, and Dawn Song. 2023. SoK: Data Privacy in Virtual Reality. https://doi.org/10.48550/arXiv.2301.05940

[19] Marc D. Gellman (Ed.). 2020. *Encyclopedia of Behavioral Medicine*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-39903-0

[20] Ceenu George, Daniel Buschek, Andrea Ngao, and Mohamed Khamis. 2020. GazeRoomLock: Using Gaze and Head-Pose to Improve the Usability and Observation Resistance of 3D Passwords in Virtual Reality. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso de Paolis and Patrick Bourdot (Eds.). Springer International Publishing, Cham, 61–81.

[21] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. 2019. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* 33, 4 (2019), 917–963. https://doi.org/10.1007/s10618-019-00619-1

[22] Anil K. Jain, Arun A. Ross, and Karthik Nandakumar. 2011. Introduction. In *Introduction to Biometrics*, Anil K. Jain, Arun A. Ross, and Karthik Nandakumar (Eds.). Springer US, Boston, MA, 1–49. https://doi.org/10.1007/978-0-387-77326-1_1

[23] Markus Jakobsson, Elaine Shi, Philippe Golle, and Richard Chow. 2009. Implicit Authentication for Mobile Devices. In *Proceedings of the 4th USENIX Conference on Hot Topics in Security (HotSec'09)*. USENIX Association, USA, 9.

[24] Alexander Kupin, Benjamin Moeller, Yijun Jiang, Natasha Kholgade Banerjee, and Sean Banerjee. 2019. Task-Driven Biometric Authentication of Users in Virtual Reality (VR) Environments. In *MultiMedia Modeling*, Ioannis Kompatsiaris, Benoit Huet, Vasileios Mezaris, Cathal Gurrin, Wen-Huang Cheng, and Stefanos Vrochidis (Eds.). Lecture Notes in Computer Science, Vol. 11295. Springer International Publishing, Cham, 55–67. https://doi.org/10.1007/978-3-030-05710-7_5

[25] María La Sanz de Garza and Paolo Emilio Adami. 2020. Definition of Athletes and Classification of Sports. In *Textbook of Sports and Exercise Cardiology*, Axel Pressler and Josef Niebauer (Eds.). Springer International Publishing, Cham, 3–11. https://doi.org/10.1007/978-3-030-35374-2_1

[26] Jonathan Liebers, Mark Abdelaziz, Lukas Mecke, Alia Saad, Jonas Auda, Uwe Gruenefeld, Florian Alt, and Stefan Schneegass. 2021. Understanding User Identification in Virtual Reality Through Behavioral Biometrics and the Effect of Body Normalization. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445528

[27] Jonathan Liebers, Sascha Brockel, Uwe Gruenefeld, and Stefan Schneegass. 2022. Identifying Users by Their Hand Tracking Data in Augmented and Virtual Reality. *International Journal of Human–Computer Interaction* (2022), 28 pages. https://doi.org/10.1080/10447318.2022.2120845

[28] Jonathan Liebers, Christian Burschik, Uwe Gruenefeld, and Stefan Schneegass. 2023. Exploring the Stability of Behavioral Biometrics in Virtual Reality in a Remote Field Study: Towards Implicit and Continuous User Identification through Body Movements. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology* (, Christchurch, New Zealand,) *(VRST '23)*. Association for Computing Machinery, New York, NY, USA, Article 30, 12 pages. https://doi.org/10.1145/3611659.3615696

[29] Jonathan Liebers, Patrick Horn, Christian Burschik, Uwe Gruenefeld, and Stefan Schneegass. 2021. Using Gaze Behavior and Head Orientation for Implicit Identification in Virtual Reality. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology (VRST '21)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3489849.3489880

[30] Jonathan Liebers and Stefan Schneegass. 2020. Gaze-based Authentication in Virtual Reality. In *Symposium on Eye Tracking Research and Applications*, Andreas Bulling, Anke Huckauf, Eakta Jain, Ralph Radach, and Daniel Weiskopf (Eds.). ACM, New York, NY, USA, 1–2. https://doi.org/10.1145/3379157.3391421

[31] M. Sivasamy, V. N. Sastry, and N. P. Gopalan. 2020. VRCAuth: Continuous Authentication of Users in Virtual Reality Environment Using Head-Movement. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, Coimbatore, India, 518–523. https://doi.org/10.1109/ICCES48766.2020.9137914

[32] Florian Mathis, John Williamson, Kami Vaniea, and Mohamed Khamis. 2020. RubikAuth: Fast and Secure Authentication in Virtual Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382827

[33] M. R. Miller, E. Han, C. DeVeaux, E. Jones, R. Chen, and J. N. Bailenson. 2023. A Large-Scale Study of Personal Identifiability of Virtual Reality Motion Over Time. https://doi.org/10.48550/arXiv.2303.01430

[34] M. R. Miller, F. Herrera, H. Jun, J. A. Landay, and J. N. Bailenson. 2020. Personal identifiability of user tracking data during observation of 360-degree VR video. *Scientific Reports* 10, 1 (2020), 17404. https://doi.org/10.1038/s41598-020-74486-y

[35] R. Miller, A. Ajit, N. K. Banerjee, and S. Banerjee. 2019. Realtime Behavior-Based Continual Authentication of Users in Virtual Reality Environments. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE Computer Society, Los Alamitos, CA, USA, 253–2531. https://doi.org/10.1109/AIVR46125.2019.00058

[36] R. Miller, N. K. Banerjee, and S. Banerjee. 2021. Using Siamese Neural Networks to Perform Cross-System Behavioral Authentication in Virtual Reality. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisboa, Portugal, 140–149. https://doi.org/10.1109/VR50410.2021.00035

[37] R. Miller, N. K. Banerjee, and S. Banerjee. 2022. Combining Real-World Constraints on User Behavior with Deep Neural Networks for Virtual Reality (VR) Biometrics. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Christchurch, New Zealand, 409–418. https://doi.org/10.1109/VR51125.2022.00060

[38] R. Miller, N. K. Banerjee, and S. Banerjee. 2022. Temporal Effects in Motion Behavior for Virtual Reality (VR) Biometrics. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Christchurch, New Zealand, 563–572. https://doi.org/10.1109/VR51125.2022.00076

[39] J. H. Mitchell, W. Haskell, P. Snell, and S. P. van Camp. 2005. Task Force 8: classification of sports. *Journal of the American College of Cardiology* 45, 8 (2005), 1364–1367. https://doi.org/10.1016/j.jacc.2005.02.015

[40] J. H. Mitchell, W. L. Haskell, and P. B. Raven. 1994. Classification of sports. *Journal of the American College of Cardiology* 24, 4 (1994), 864–866. https://doi.org/10.1016/0735-1097(94)90841-9

[41] Bendik B. Mjaaland, Patrick Bours, and Danilo Gligoroski. 2011. Walk the Walk: Attacking Gait Biometrics by Imitation. In *Information Security*, Mike Burmester, Gene Tsudik, Spyros Magliveras, and Ivana Ilić (Eds.). Lecture Notes in Computer Science, Vol. 6531. Springer Berlin Heidelberg, Berlin, Heidelberg, 361–380. https://doi.org/10.1007/978-3-642-18178-8_31

[42] A. G. Moore, R. P. McMahan, H. Dong, and N. Ruozzi. 2021. Personal Identifiability and Obfuscation of User Tracking Data From VR Training Sessions. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, Bari, Italy, 221–228. https://doi.org/10.1109/ISMAR52148.2021.00037

[43] Tahrima Mustafa, Richard Matovu, Abdul Serwadda, and Nicholas Muirhead. 2018. Unsure How to Authenticate on Your VR Headset? Come on, Use Your Head!. In *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics (IWSPA '18)*. Association for Computing Machinery, New York, NY, USA, 23–30. https://doi.org/10.1145/3180445.3180450

[44] Vivek Nair, Wenbo Guo, Justus Mattern, Rui Wang, James F. O'Brien, Louis Rosenberg, and Dawn Song. 2023. Unique Identification of 50,000+ Virtual Reality Users from Head & Hand Motion Data. http://arxiv.org/pdf/2302.08927v1

[45] Vivek Nair, Wenbo Guo, Rui Wang, James F. O'Brien, Louis Rosenberg, and Dawn Song. 2023. Berkeley Open Extended Reality Recordings 2023 (BOXRR-23): 4.7 Million Motion Capture Recordings from 105,852 Extended Reality Device Users. https://doi.org/10.48550/arXiv.2310.00430

[46] Vivek Nair, Christian Rack, Wenbo Guo, Rui Wang, Shuixian Li, Brandon Huang, Atticus Cull, James F. O'Brien, Marc Latoschik, Louis Rosenberg, and Dawn Song. 2023. Inferring Private Personal Attributes of Virtual Reality Users from Head and Hand Motion Data. arXiv:2305.19198 [cs.HC]

[47] Vivek C. Nair, Gonzalo Munilla-Garrido, and Dawn Song. 2023. Going Incognito in the Metaverse: Achieving Theoretically Optimal Privacy-Usability Tradeoffs in VR. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, Sean Follmer, Jeff Han, Jürgen Steimle, and Nathalie Henry Riche (Eds.). ACM, New York, NY, USA, 1–16. https://doi.org/10.1145/3586183.3606754

[48] Alex Nosenko, Yuan Cheng, and Haiquan Chen. 2022. Learning Password Modification Patterns with Recurrent Neural Networks. In *Secure Knowledge Management In The Artificial Intelligence Era*, Ram Krishnan, H. Raghav Rao, Sanjay K. Sahay, Sagar Samtani, and Ziming Zhao (Eds.). Springer International Publishing, Cham, 110–129.

[49] Ilesanmi Olade, Charles Fleming, and Hai-Ning Liang. 2020. BioMove: Biometric User Identification from Human Kinesiological Movements for Virtual Reality Systems. *Sensors* 20, 10 (2020), 2944. https://doi.org/10.3390/s20102944

[50] Johnny Padulo, Guillaume Laffaye, Karim Chamari, and Alberto Concu. 2013. Concentric and Eccentric: Muscle Contraction or Exercise? *Sports Health: A Multidisciplinary Approach* 5, 4 (July 2013), 306–306. https://doi.org/10.1177/1941738113491386

[51] Marietta Papadatou-Pastou, Eleni Ntolka, Judith Schmitz, Maryanne Martin, Marcus R. Munafò, Sebastian Ocklenburg, and Silvia Paracchini. 2020. Human handedness: A meta-analysis. *Psychological bulletin* 146, 6 (2020), 481–524. https://doi.org/10.1037/bul0000229

[52] Ken Pfeuffer, Matthias J. Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 2019. Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300340

[53] R. Miller, N. K. Banerjee, and S. Banerjee. 2020. Within-System and Cross-System Behavior-Based Biometric Authentication in Virtual Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Atlanta, Georgia, USA, 311–316. https://doi.org/10.1109/VRW50115.2020.00070

[54] Christian Rack, Tamara Fernando, Murat Yalcin, Andreas Hotho, and Marc Erich Latoschik. 2023. Who is Alyx? A new behavioral biometric dataset for user identification in XR. *Frontiers in Virtual Reality* 4 (2023), 11 pages. https://doi.org/10.3389/frvir.2023.1272234

[55] Christian Rack, Konstantin Kobs, Tamara Fernando, Andreas Hotho, and Marc Erich Latoschik. 2023. Extensible Motion-based Identification of XR Users using Non-Specific Motion Data. arXiv:2302.07517 [cs.HC]

[56] Soumen Roy, Jitesh Pradhan, Abhinav Kumar, Dibya Ranjan Das Adhikary, Utpal Roy, Devadatta Sinha, and Rajat Kumar Pal. 2022. A Systematic Literature Review on Latest Keystroke Dynamics Based Models. *IEEE Access* 10 (2022), 92192–92236. https://doi.org/10.1109/ACCESS.2022.3197756

[57] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[58] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Elsevier, Amsterdam, The Netherlands, 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9

[59] M. Angela Sasse, Michelle Steves, Kat Krol, and Dana Chisnell. 2014. The Great Authentication Fatigue – And How to Overcome It. In *Cross-cultural design*, P. L. P. Rau (Ed.). Lecture notes in computer science Information systems and application, incl. Internet/web and HCI, Vol. 8528. Springer, Cham, 228–239. https://doi.org/10.1007/978-3-319-07308-8_23

[60] Christian Schell, Andreas Hotho, and Marc Erich Latoschik. 2022. Comparison of Data Encodings and Machine Learning Architectures for User Identification on Arbitrary Motion Sequences. In *2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, CA, USA, 11–19. https://doi.org/10.1109/AIVR56993.2022.00010

[61] Albrecht Schmidt. 2000. Implicit human computer interaction through context. *Personal Technologies* 4, 2-3 (2000), 191–199. https://doi.org/10.1007/BF01324126

[62] E. Shi, Y. Niu, M. Jakobsson, and R. Chow. 2011. Implicit Authentication through Learning User Behavior. In *Information Security*, M. Burmester, G. Tsudik, S. Magliveras, and I. Ilić (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 99–113. https://doi.org/10.1007/978-3-642-18178-8_9

[63] Sophie Stephenson, Bijeeta Pal, Stephen Fan, Earlence Fernandes, Yuhang Zhao, and Rahul Chatterjee. 2022. SoK: Authentication in Augmented and Virtual Reality. In *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 267–284. https://doi.org/10.1109/SP46214.2022.9833742

[64] Kalaivani Sundararajan and Damon L. Woodard. 2019. Deep Learning for Biometrics. *ACM Computing Surveys* 51, 3 (2019), 1–34. https://doi.org/10.1145/3190618

[65] Issa Traore and Ahmed Awad E. Ahmed (Eds.). 2012. *Continuous Authentication Using Biometrics: Data, Models, and Metrics*. IGI Global, Hershey, PA, USA.

[66] Issa Traoré and Ahmed Awad E. Ahmed. 2012. Introduction to Continuous Authentication. In *Continuous Authentication Using Biometrics: Data, Models, and Metrics*, Issa Traore and Ahmed Awad E. Ahmed (Eds.). IGI Global, Hershey, PA, USA, 1–22. https://doi.org/10.4018/978-1-61350-129-0.ch001

[67] Pier Paolo Tricomi, Federica Nenna, Luca Pajola, Mauro Conti, and Luciano Gamberini. 2023. You Can't Hide Behind Your Headset: User Profiling in Augmented and Virtual Reality. *IEEE Access* 11 (2023), 9859–9875. https://doi.org/10.1109/ACCESS.2023.3240071

[68] Xue Wang and Yang Zhang. 2021. Nod to Auth: Fluent AR/VR Authentication with User Head-Neck Modeling. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411763.3451769

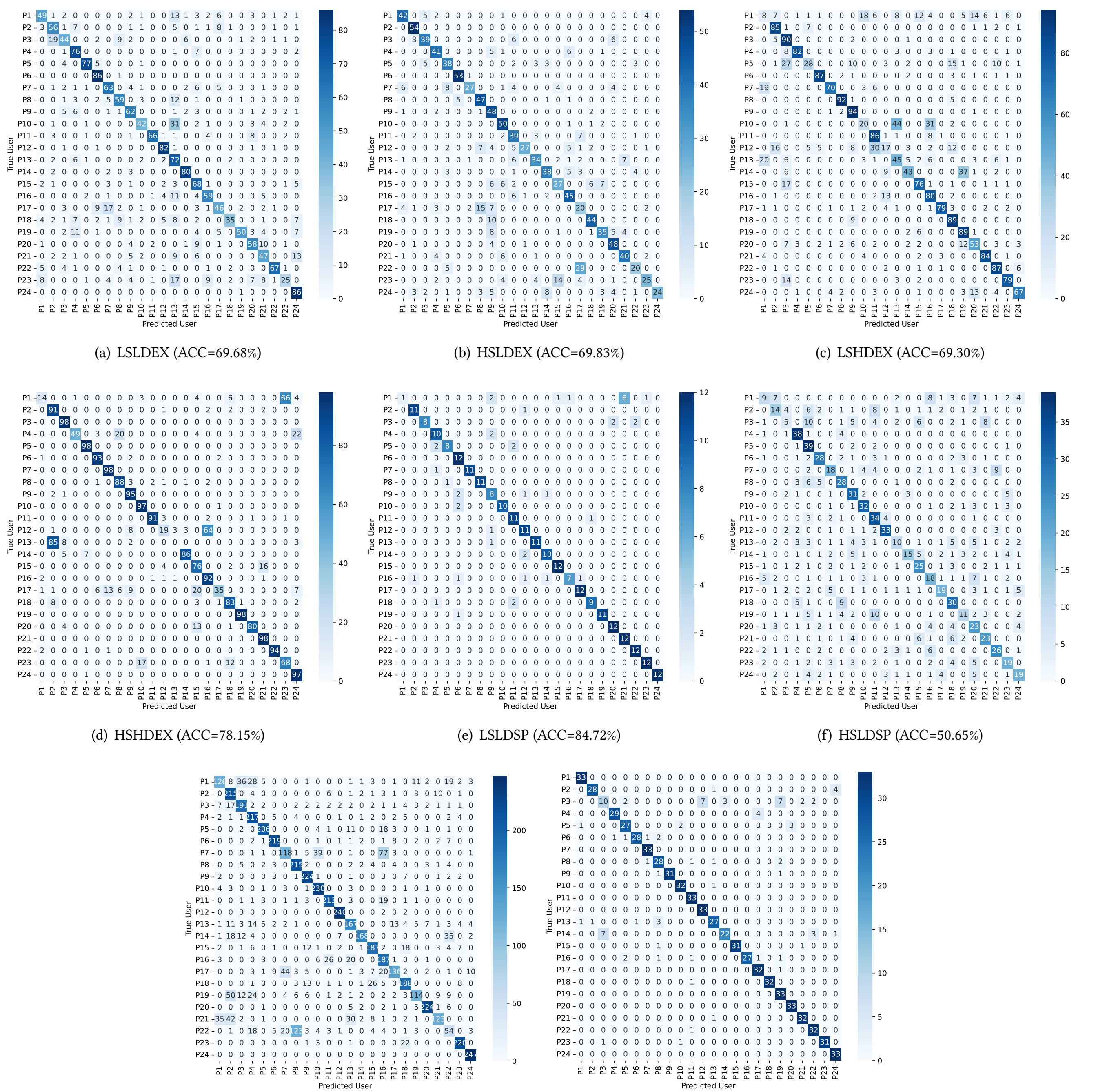
[69] Elliott Wen, Chitralekha Gupta, Prasanth Sasikumar, Mark Billinghurst, James Wilmott, Emily Skow, Arindam Dey, and Suranga Nanayakkara. 2023. VR.net: A Real-world Dataset for Virtual Reality Motion Sickness Research. https://doi.org/10.48550/arXiv.2306.03381

[70] Mariusz Wierzbowski, Grzegorz Pochwatko, Paulina Borkiewicz, Daniel Cnotkowski, Michal Pabis-Orzeszyna, and Pawel Kobylinski. 2022. Behavioural Biometrics in Virtual Reality: To What Extent Can We Identify a Person Based Solely on How They Watch 360-Degree Videos?. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, Singapore, 417–422. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00090

[71] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[72] Z. Yu, H. Liang, C. Fleming, and K. L. Man. 2016. An exploration of usable authentication mechanisms for virtual reality systems. In *2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*. IEEE, Jeju, South Korea, 458–460. https://doi.org/10.1109/APCCAS.2016.7804002

A CONFUSION MATRICES



(a) LSLDEX (ACC=69.68%)

(b) HSLDEX (ACC=69.83%)

(c) LSHDEX (ACC=69.30%)

(d) HSHDEX (ACC=78.15%)

(e) LSLDSP (ACC=84.72%)

(f) HSLDSP (ACC=50.65%)

(g) HSHDSP (ACC=74.71%)

(h) LSHDSP (ACC=89.65%)

Figure 10: Confusion Matrices per activity. As models, we selected Inception for *Exercise (EX)*, as shown in (a) to (d). For *Sport (SP)* we selected FCN and its confusion matrices are shown in (e) to (h). Table 1 lists all accuracies.