TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
2026, Vol. 2680(2) 519–542

DOI: 10.1177/03611981251365899
journals.sagepub.com/home/trr
Sage

Walking in an Urban Environment and a Virtual Reality Replica: Comparisons of Physical Activity Duration and Intensity

Amanda N. Spitzer[1], Matea R. Ramey[1], Yiqing "Skylar" Yu[1], Katrina M. Oselinsky[2], Katie McMahon[1], Brendan Kelley[3,4], David Rojas-Rueda[5], Daniel Dean[5], Sara B. LoTemplio[6], Francisco R. Ortega[3], and Dan J. Graham[1,7]

Abstract

Increasing walking behavior is desirable from public health, environmental, and urban planning perspectives. Virtual reality (VR) has the potential to improve the design of walkable environments. However, the current research was necessary to determine whether walking decisions in VR mirror those in the real world (RW). Participants completed two study sessions: walking in a VR simulation of a historic district (VR session) and walking in the real-life district (RW session). During each session, participants were asked to complete three tasks (e.g., find a restaurant) and stop walking following task completion. Heart rate (HR) data contained a high degree of missingness, so no HR analyses are reported. Nevertheless, walking intensity is addressed through exploratory negative binomial and Poisson regression models predicting duration in light and moderate-to-vigorous physical activity using accelerometry. These models indicated no relationship between physical activity intensity in VR and the RW. Additionally, a paired t-test and mixed-effects model indicated that walking duration was significantly longer in VR than the RW. However, exploratory analyses suggested order effects: those who walked first in the RW walked similar durations in both settings, but those that walked first in VR walked for about 5 min longer in VR (17.8 min) than in the RW (13.0 min). In conclusion, walking intensity in VR may not mimic walking intensity in the RW, however, depending on the order of condition presentation, walking decisions in VR may resemble RW decisions. Possible explanations for the observed order effects include history effects, VR navigation and skill transfer, and participant motivation.

Keywords

virtual reality, walking behavior, built environment, walkability, environmental factors

The built environment is a modifiable factor that boosts community-level walking (*1–3*). Increasing community-level walking behavior in urban environments is beneficial from health, environmental, communal, and economic perspectives. Firstly, increased walking activity is associated with decreased risk for cardiovascular disease (*4*), cancer recurrence (*5*), and cognitive decline in elderly adults (*6*). Concerning wellbeing, walking can increase social connection to people in the neighborhood (*7*) and lift mood among people struggling with their mental health. Though some research has found that mood differentially improves when individuals walk in nature compared with urban environments (*8*, *9*), other research has found that both urban and outdoor nature walks

[1]Department of Psychology, Colorado State University, Fort Collins, CO
[2]Division of Endocrinology, Metabolism & Diabetes, University of Colorado, Anschutz Medical Campus, Aurora, CO
[3]Department of Computer Science, Colorado State University, Fort Collins, CO
[4]Nancy Richardson Design Center, College of Health and Human Sciences, Colorado State University, Fort Collins, CO
[5]Department of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO
[6]Human Dimensions of Natural Resources, Colorado State University, Fort Collins, CO
[7]Department of Community and Behavioral Health, Colorado School of Public Health, Aurora, CO

Corresponding Author:
Amanda N. Spitzer, Amanda.Spitzer@colostate.edu

succeed equally in improving mood (*10*, *11*). Secondly, walking for transport is also a protective factor in climate change as it reduces fossil fuel use and demand for building motor vehicles (*12*, *13*). Thirdly, there are various community-level benefits to walking, as advocated by urban designers. Walking is positively associated with the mutually beneficial social connections that contribute to social capital (*14*, *15*). Nonmotorized transport is linked to greater community engagement (*15*), and walking correlates with larger community networks and greater trust in neighbors (*14*). Walking for transport is also discussed in urban design as an issue of equity (*16*), as walking can provide inexpensive access to necessary resources as well as recreation. Lastly, mindful of the economic benefits of walking, city planners produce designs that optimize such opportunities. Having numerous pedestrians in an area is linked to greater retail sales (*17*), thus, increased walking behavior boosts the local economy and may relate to higher employment rates (*18*). In summary, walking can provide numerous improvements to physical health, mood, environmental health, social capital, and local economies.

Alongside these numerous benefits, modifying the environment to boost walkability also has costs: changing environments can be expensive and time-consuming. Planners face challenges in identifying which of several possible environmental modifications will do the most to increase walking. Correlations between self-reported physical activity (PA) and the built environment in neighborhoods suggest a complex interplay between factors such as open space, traffic intensity, parking, water, infrastructure for walking and cycling, and intangibles like perceived "activity friendliness" (*19*). This interplay complicates the process of isolating any single intervention priority. Beyond the physical landscape, legal and social barriers may further restrict the effectiveness of any attempted intervention (*20*). Additional complexities arise when considering responses in a given community. To focus on urban vegetation as one well-supported predictor of PA, even granular choices like the spacing of trees may influence walkability (*21*). Ideal land cover types to promote walking may vary across racial or ethnic groups (*22*). These site-specific variations suggest that careful attention to local communities and conditions should be paid when seeking to apply previous research to a new context. Together, this complexity and context-dependence make it important to understand how a given intervention (e.g., a new park or trail, improved sidewalks) would fit into the fabric of a proposed community and setting.

The relatively permanent and cost-intensive nature of urban infrastructure makes observational studies and natural experiments the de facto modes for investigations into built environmental interventions to boost walking. These approaches explore associations rather than establish causality because they retroactively examine changes in behavior. To gain insights similar to those in an experimental setting (*23*), observational approaches require development of causal inference frameworks. These frameworks demand fairly detailed situational contexts to approximate randomized "treatment" and "control" groups (e.g., using otherwise similar populations differing in a feature of interest with associated statistical adjustments). Even with these approximations, there remains the possibility of systematic biases via unobserved or uncontrollable variables (*23*). Although observational and causal inference-based designs serve an important role in research, true experimental study designs examining built environmental factors affecting pedestrian behavior remain an important gap. Experimental designs would enable researchers to prospectively compare alternate proposals for a constant study population. Unfortunately, the cost, time investment, and influence on behavior associated with construction makes this a daunting proposition.

Virtual reality (VR) has been promoted and employed as a method of addressing these issues of cost, time, and causality (*24*, *25*). This application is afforded by the immersive nature of VR, which allows users to experience presence: a state of being where their cognitive processing is so engrossed with the virtual environment that they feel as if they are in that environment (*26*). Presence is achieved through multiple aspects of immersion, however, a key element to consider for physical studies such as walking is that of embodiment. Embodiment is the notion that thoughts, feelings, and behaviors are grounded in physical interactions through the corporeal form (*27*). In VR, the body is used to interact with the digital world without layers of abstraction presented by human input devices (i.e., mouse and keyboard), enabling VR training and exploration of PAs such as jumping (*28*) or climbing (*29*). Thus, VR can produce a convincing illusion that allows for the transferability of results from the virtual environment to the physical world.

There are several benefits to using VR environments as experimental tools to examine built environmental factors of behavior. First is the ability to easily modify and adapt the environment for testing different interventions, allowing for a superb level of control. The tools used for building VR environments (*30*, *31*) allow experimental designers to test differing environments and adjust them mid-experiment (*32*). This process was utilized by Chung and Sparks in which participants experienced different virtual rooms to determine how simulation configuration affects memory (*33*). In sum, the digital control afforded by VR development tools enables deeper exploration of perceptive and cognitive concepts (*34*).

With this technology, researchers can immerse participants in virtual environments, leading to realistic behaviors that they perform in the RW (*35*). In other words, by manipulating the elements of interest within a virtual environment, VR can be used to experimentally test environmental factors affecting walking. This approach can aid municipal governments in their goal to make cost-effective and timely changes to the RW built environment that successfully promote walking behavior.

However, if VR is to be used in this manner, researchers must understand how walking behavior in VR compares with RW walking. In the absence of this knowledge, the value of using VR in walkability research is unclear. As yet, research has not definitively shown that walking decisions in VR are similar to RW decisions. Previous research comparing walking in VR with RW walking has provided mixed results with some studies supporting comparability (*36 –38*) and others noting discrepancies (*39 –44*). Although there is a wide base of literature concerning walking in VR employing various hardware, software, VR navigation methods, and RW locomotion techniques, most are outside the scope of the present study. Therefore, we have focused on those studies that use an immersive head-mounted display (HMD) and overground (i.e., room-scale) walking. These criteria were selected because HMD-delivered VR is currently regarded as the most immersive for walking (*45*); additionally, gait-related mechanics and adaptions in VR using overground walking more similarly match those made in the RW than alternative VR locomotion techniques. Specifically, overground walking in VR results in larger adjustments in gait kinematics compared with treadmill walking, illustrating a greater kinematic response to the demands of the environment (*46*). Further, previous research has shown overground walking in VR to be a feasible method of assessing environmental impacts on walkability (*25*). In accordance, the set-up of the current study employs both methods.

Assessing distance has been found to be both equivalent (*37*, *38*) and divergent (*41*, *43*) in VR when compared with the RW depending on the method of assessment (e.g., verbal estimation, perceptual matching, and visually directed actions). A review summarizing the divergent results in VR distance estimation can be found in Feldstein (*38*). Differences in gait characteristics have also been observed. Canessa et al. reported that participants walked the same distance while navigating a defined walkway in VR as they did in the RW, yet they walked with a greater cadence in VR (*47*). Another study found that participants walked across crosswalks in VR at the average RW pedestrian crossing speed (*36*). Conversely, other studies have observed reduced walking speed (*39*, *40*, *42*, *44*) and slower cadence in VR than in the RW (*44*).

Importantly, walkability research not only relies on distance and walking speed, but also on pedestrian judgments of safety and experience. Individuals evaluate the safety of a street similarly in VR as they do in real life, and their intention to cross is equivalent given similar conditions (e.g., speed of cars) (*37*). Concerning experiential judgments, participants in a study by Bhagavathula et al. perceived traffic as moving faster in VR than in the RW (*37*). VR simulations of RW city locations are judged to be comparable to the corresponding RW locations on some experience dimensions that may influence walking behavior, such as pleasantness and unity, but not on others, such as brightness, and spaciousness (*48*). A separate study found similar results with participants rating environments differently in VR on some dimensions (cohesiveness and cleanliness) but identically on other dimensions (beauty, quietness, and familiarity) (*49*). Of note is that this study employed landscape scenes, so these findings may not hold for streetscapes.

Whereas previous research has focused on assessing walking-related perceptions, rather than behaviors (*37*, *48*, *49*) or on walking behaviors in indoor environments (*39*, *47*), the novel contributions of the present study are to empirically compare VR and RW urban-scale walking-related decisions (such as when to continue versus stop walking while engaging in representative daily tasks) and behaviors using overground walking.

Specifically, the present study investigates how walking in a VR model of a RW street compares to walking in the corresponding RW street. This research can help clarify whether experiments utilizing VR to examine built environmental modifications are likely to provide meaningful findings concerning RW behavior. The aim is to understand whether walking duration and intensity differed in the two settings to inform the interpretation of walking behavior data gathered from VR models of RW environments while controlling for individual differences. To do so, we implemented a within-subjects design so that each participant would complete the same tasks in both VR and the RW. We hypothesized that participants would walk for the same duration at the same intensity in VR as in the RW.

Materials and Methods

Participants

An a priori sample of 60 participants was selected. We attempted to recruit and schedule as many participants as possible within the constraints of our 3-week data collection plan (see the section covering the procedure). During the summer of 2022, 56 participants were recruited using Colorado State University's a staff email

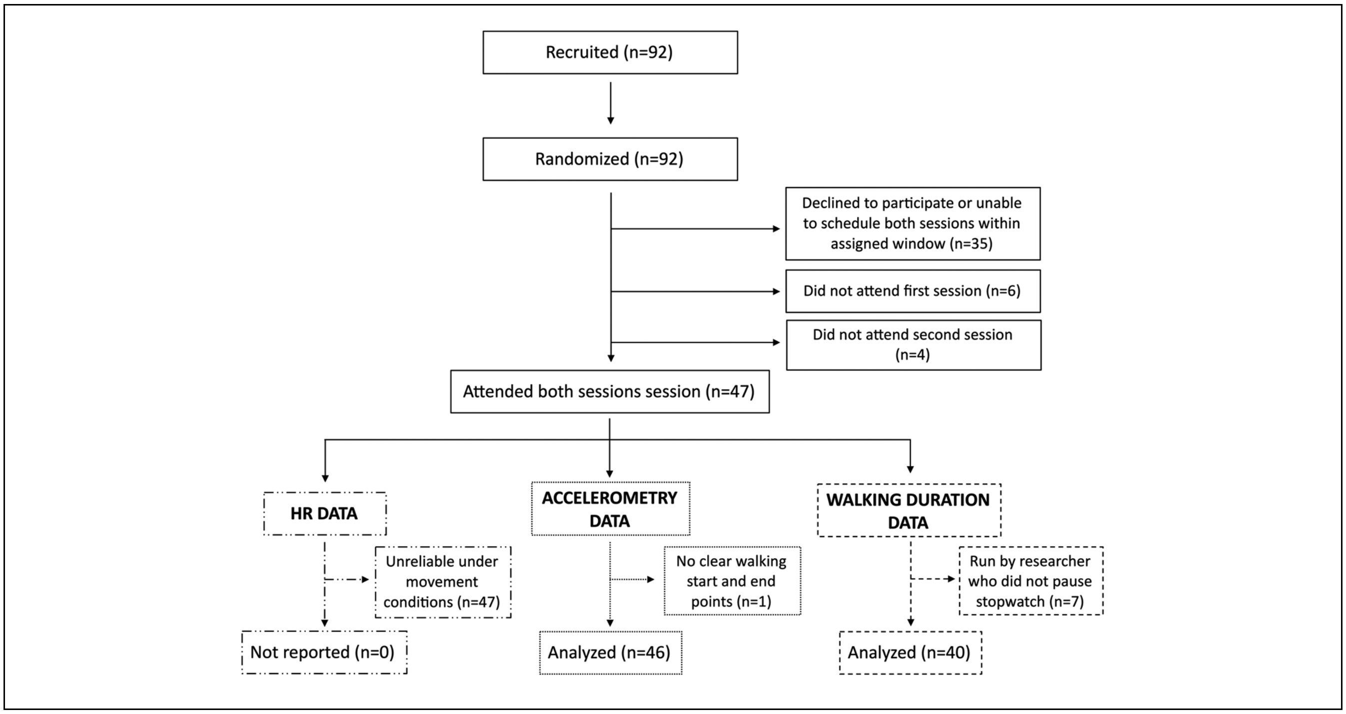

Figure 1. Exclusion flow diagram.
Note: HR = heart rate.

list associated with a large public university in the Western United States to complete two in-person study sessions: a VR session, in which participants walked in a VR simulation, and a RW session, in which participants walked in the real-life counterpart of the simulated environment. Eligibility criteria included being at least 18 years old, possessing the capability of walking for up to 30 min on two separate days, and not being prone to dizziness or other problems that would preclude participants from using VR technology. Several participants did not attend either their first (6) or second session (4), so they were excluded from analyses. This resulted in 47 participants completing both sessions. See Figure 1 for participant attrition information. The participants who attended at least one session were primarily university employees (80.4%), but some were local residents unaffiliated with the university (13.7%); two were retired university employees, and one was a university student. The mean age was 41.7 years (standard deviation [SD] = 14.3) Most of the participants were assigned female at birth (62.7%), and 37.3% were assigned male. The mean height was 5 ft 7 in. (SD = 3.89 in.). Most participants were right-handed (40) though 11 were left-handed, and 1 indicated that they used both hands equally. Participants were compensated with $30 cash. After completion of the first session, they received $10, and the remaining $20 was given to them after completion of the second session.

Apparatus/Measures

Virtual Reality Simulation. This study employed a virtual environment simulation depicting a 0.5-mi portion of the main downtown corridor in a mid-sized city in the Western United States. An image capture of the simulation is depicted in Figure 2a, and the RW environment is illustrated in Figure 2b. The research team had previously developed the environment using SketchUp for three-dimensional modeling, Blender for rendering, and Unity Game Engine for translation into a VR environment. The VR environment was structured with high fidelity to the real environment in relation to the building façades and shapes (i.e., the buildings were the same dimensions in the VR environment as in the real one, and the businesses included in the model were those that existed in the real environment at the time it was constructed, some of which no longer exist). Buildings were textured using photographs of the RW environment, making their appearances identifiable (see Figure 3). The simulation featured greenery, moving cars, and directional sound effects when passing cars and restaurants. Some elements in the VR environment did not have the same high level of fidelity to the real environment, and these included the traffic (there were fewer and less varied vehicles in the VR environment, there were no cyclists or pedestrians in the VR environment) and the range of sounds in the VR environment, which was restricted to generic background auto traffic, bird sounds, and a din

Figure 2. (*a*) Virtual reality (VR) simulation and (*b*) real-world (RW) environment comparison.
Source: RW image from Colorado State University Architectural Virtual Library (*52*).

Figure 3. Detailed photographs of (*a*) a restaurant and (*b*) the FedEx store within the virtual reality simulation.

of conversation when passing by businesses where people congregate (e.g., restaurants, bars) in the VR environment. The standard Unity skybox was applied to simulate weather in VR. This skybox simulated a clear sunny day. The region in which the data were collected has overwhelmingly sunny days (*50*). In fact, we recorded only four RW sessions in which it rained. (Research assistants carried umbrellas for this possibility.) Therefore, the weather in VR sessions mostly mirrored that in the RW. Rotation in the VR environment was enabled through a snap turn function with which a user can rotate left or right using the controller joystick. This method of viewport snapping or snap turns is recommended to reduce simulation sickness (*51*). For further details about the simulation and its development process, see Oselinsky et al. (*25*).

VR Apparatus. The research team loaded the VR simulation onto an Oculus Meta Quest 2 HMD (Model KW49CM). The Meta Quest 2 provides for an untethered VR experience (i.e., does not require a separate high-end PC to run the VR simulation) allowing for greater freedom of user mobility. The HMD head straps, interpupillary distance, and glasses spacer were fit to the participant for clarity and comfort. While in the VR environment, participants held both hand controllers.

Walking Duration. Walking duration was measured by stopwatch. The stopwatch began when the participant started walking and stopped when they ended the walking session. We did not record how many times the participants stopped walking. When RW constraints prohibited walking, the research assistant paused the stopwatch until the participant was free to walk. In other words, whenever participants stopped walking, the research assistant stopped the time on the stopwatch and resumed the stopwatch once participants resumed walking. Participants stopped primarily in response to stoplights in the RW and in response to rotating the simulation when approaching a room boundary in the VR condition. Participants were asked to walk to complete three tasks, which are described within the section focusing on procedure. Although two of the three tasks were completed as soon as a participant found a particular location, the remaining task (finding a restaurant to take a friend visiting the area) could involve finding several locations or stopping as soon as the first restaurant was reached. Therefore, walking duration reflects not only the required tasks but also the decision to continue walking.

Empatica e4 Wristband. The Empatica e4 wristband was used to measure HR and activity intensity via accelerometry. Although it is possible to collect similar accelerometry data from the Meta Quest 2 HMD, the computational cost to collect that data each frame may have introduced stutter or frame drops, which, in turn, would increase the risk of adverse VR symptoms (i.e., cybersickness). The e4 is a lightweight (25 g) research-grade device worn on the nondominant wrist like a watch and is equipped with four sensors measuring movement and physiological responses. It records HR data at a frequency of 1 Hz (*53*), and its reliability concerning HR measurement has been established among adults in various common research paradigms (*54*). It measures accelerometry data in three dimensions on a -2- to 2-g scale with a frequency of 32 Hz (*55*). This wrist-worn device was chosen over a dual-system of an HR chest-strap and hip-worn accelerometer because the Empatica e4 is a single device and less invasive. Participants were asked to don the HR and accelerometry device in a public space for the RW condition. For participants' privacy, the Empatica e4 wristband was selected as it does not require placement on the torso under clothes.

Postexperience Survey. At the conclusion of each study session, participants completed a short survey. The survey following the RW session included 17 items. It inquired about the number of previous visits to the historic district, physical discomfort, judgments and emotions about the experience, distractions, recommendations, and demographic information. The survey did not inquire about the participant's prior knowledge of the buildings in the RW. For most items about physical discomfort, participants rated the degree to which they experienced various dimensions of discomfort, such as eyestrain, nausea, and dizziness, on a 5-point Likert scale (1 = none at all, 5 = a great deal). Other questions relating to emotions and judgments presented statements like "The environment felt safe" or "The environment was attractive" and were answered on a 5-point Likert scale from strongly disagree (1) to strongly agree (5). Lastly, two free-response questions asked, "Was anything in the environment distracting?" and "What would make you want to spend more time walking around this environment?"

The survey presented at the end of the VR session included all items from the RW survey and eight additional items specific to the VR experience for a total of 25 items. These additional items inquired about familiarity with VR and the adjustment to, navigation of, immersion in, and realism of this particular VR experience. For items concerning navigation, immersion, and realism, participants rated their agreement with statements such as "The environment seemed realistic and believable" and "I felt immersed in the VR environment" on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). One VR-specific free-response question was added: "What changes to the environment would you recommend to make it feel more real/immersive?" The survey, walking duration, accelerometry, and heart interbeat interval data (HR raw data) will be made available on OSF. This study was previously presented at the 44th Annual Meeting & Scientific Sessions of the Society of Behavioral Medicine (*56*).

Procedure

Before recruitment, the host university's institutional review board approved the study (#3526). After expressing their interest in participation by responding to the recruitment email, possible participants were assigned a participant ID and randomized to complete either the RW session first order (RW->VR group) or the VR session first order (VR->RW group) to counterbalance for order effects. Each participant was subjected to both VR and RW conditions. Following this randomization, participants were prompted to sign up for two study sessions that corresponded to their randomized order. Most manufacturers discourage the use of their HMD devices outdoors owing to potential lens damage from the sun (*57*, *58*). Further, HMDs rely on sensors for tracking movement. The researchers tested the HMD in outdoor environments and observed that bright outdoor lighting (i.e., sunlight) can interfere with these sensors, leading to inaccurate tracking and a disruptive experience. Therefore, we needed to rent a large (60 × 80 ft) indoor space to complete the VR sessions. Because of financial constraints, this space was only rented for a single week. Consequently, data collection lasted 3 weeks. The RW->VR group completed their RW session during the first week of data collection. The indoor space was rented during the second week, during which all participants completed their VR session. The VR->RW group performed the RW session in the third and final week of data collection. At the onset of both study sessions and before completing any research tasks, participants gave informed consent.

RW Session

For the RW session, participants arrived at a landmark in the historic district. After obtaining informed consent, a research assistant placed the Empatica e4 wristband on the participant's nondominant wrist. The research assistant instructed the participant to walk around and attempt to find and use a crosswalk, to find a restaurant where they would want to bring a friend who is visiting the city, and to find a specific shipping courier store. These tasks were chosen because they illustrate behaviors that are typically performed in the RW setting (the historic district). Specifically, the use of crosswalks is necessary to navigate the historic district. Finding a restaurant to bring a visiting friend requires a choice to be made on behalf of the participant (i.e., they could choose the first restaurant they saw, or they could choose to walk until they were satisfied with their decision). Lastly, finding the specific shipping courier store represents the task of finding a specific location—an endeavor that can occur frequently in the RW. This specific shipping courier store was selected because it is a well-known store (FedEx) but difficult to locate. The participant was unable to use their phone, or any other type of device that utilized a global positioning system to complete any of the tasks. This was done to ensure that the attempted tasks in the RW and VR session were the same. The participant was reminded that they could choose to stop walking and take a break if they felt uncomfortable at any point, and if they did not wish to continue, they could quit the study. The temperature and humidity were measured using a digital hygrometer at the starting location. After marking the temperature, humidity, and time, the research assistant informed the participant that they could begin their tasks.

As the participant completed their tasks, the research assistant walked behind them so that they did not influence the participant's chosen pace and direction of walking. When participants completed each task, they notified the research assistant. Participants were not required to walk for the 25-min duration. Participants could choose to stop walking at any point in the session. If the participant completed all three tasks, the walking segment of the session was concluded. When the stopwatch showed 20 min, the research assistant asked the participant how they felt and whether they wanted to continue walking. If they chose to stop walking, the walking session ended. If the participant chose to continue walking, they were given five additional minutes before the research assistant ended the walking segment.

Following the session's walking segment, the research assistant removed the Empatica e4 wristband from the participant's wrist and directed them to a street bench or chair where they could rest and complete the postexperience survey. Following the survey, the research assistant thanked the participant, provided compensation, and debriefed them.

VR Session. VR study sessions were held in an indoor room measuring roughly 60 × 80 ft (18.3 × 24.4 m). On arrival, the participant gave informed consent. Next, a research assistant placed the Empatica e4 wristband on the participant's nondominant wrist and walked the research assistant to the center of the room. The participant was then trained on the use of hand controllers. The research assistant showed the participant the two buttons (the trigger button and thumbstick) and clarified that these were the only buttons that the participants would touch. The participant was also educated on the proper fit of the HMD. If the participant wore glasses, a glasses spacer was utilized in the headset. The research assistant fit the HMD to the participant using the head straps and verified the clarity of the image with them. The starting position of the VR simulation was the same landmark where the RW session began. Then, the researcher gave both hand controllers to the participant and explained how to use the trigger button to jump to a new location. The research assistant emphasized that walking should be the primary method for navigating the environment, and that this button was only to be used if they became stuck in the virtual environment (i.e., inside a building). Next, the research assistant explained how to use the thumbstick to rotate the environment. This was to be used when the participant reached a boundary in the RW but wanted to continue walking in the same direction in the VR environment. The participant was reminded that these were the only buttons the participants would touch. The participant practiced rotating the VR environment around them until they were comfortable. The participant was then reminded that they were able to remove the headset if they felt uncomfortable at any point and were told twice that if they did not wish to continue, they could quit the VR experience. Additionally, after 20 min they were asked if they would like to continue and were eventually stopped at 25 min.

The research assistant instructed the participant to walk around the VR environment and attempt to complete the same three tasks as in the RW session (1. find and use a crosswalk; 2. find a restaurant where they would want to dine with a friend from out of town; 3. find the shipping courier store). The VR environment was designed with identifiable details from the RW environment; see Figure 3 for an image capture of various businesses (i.e., a restaurant) in the virtual environment compared with the same businesses in the RW. The research assistant explained that they would walk beside the participant throughout the entire experience, and when the participant approached an RW boundary (i.e., a wall), the research assistant would tell them to stop, to rotate their body so that no obstacle is in their path, and then instruct the participant to use the hand controllers to rotate the VR environment around them so that they can walk in their desired direction. The temperature and humidity were measured using a digital hygrometer. After marking the temperature, humidity, and time, the research assistant informed the participant that they could begin their tasks.

When participants completed each task, they notified the research assistant. Participants were not required to walk for the 25-min duration. Participants could choose to stop walking at any point in the session. If the participant completed all three tasks, the walking segment of the session was concluded. When the stopwatch showed 20 min, the research assistant asked the participant how they felt and whether they would like to continue walking. If they chose to stop walking, the walking session ended. If the participant chose to continue walking, they were given five additional minutes before the research assistant ended the walking segment.

Following the walking portion of the session, the participant removed the equipment with the help of the research assistant and subsequently completed the post-experience survey while seated. After the survey, the research assistant thanked the participant, provided compensation, and debriefed them.

Analyses

RStudio Statistical Software (v. 2022.07.2 + 576 & v.2023.12.1 + 402) (*59*) and R Statistical Software (v.4.1.1 and v.4.3.3) (*60*) were employed to perform transformations and analyses. To determine the similarity of VR and RW walking, paired *t*-tests and mixed-effects models controlling for weather, completion of all tasks, and order of conditions were performed. Alpha was set at 0.05 for all analyses.

Data Cleaning

Concerning walking duration, some participants completed all tasks or walked for the maximum of 25 min, and then asked whether they could continue walking. They were permitted to do so, but the time at which they originally completed the tasks or the maximum of 25 min was used in relevant analyses, respectively. This was done to treat them identically to those participants who did not ask to continue walking. Lastly, two of the eight researchers who ran VR sessions did not pause the stopwatch when participants paused walking. However, all other researchers did. To amend this discrepancy, the VR sessions run by those two researchers (total of 7) were removed from walking duration analyses. Figure 1 depicts all exclusions through the stages of the study.

Calculation of Activity Index

We conservatively identified accelerometry data between the start and conclusion of the walking session within each participant's larger accelerometry dataset using the R package "adept" created by Karas et al. (*61*). This package identifies periods of walking from accelerometry data by comparing the vector magnitude of the data to a template of the vector magnitude from an accelerometer worn in the same location while walking. Select templates are included in the package "adeptdata" (*62*) whereas researchers are able to create additional templates. Both approaches were used in the current study together with an additional approach intended to identify resting and nonresting periods using the "adept" package described by Karas et al. (*63*). Converging data from these three approaches identified the walking session start and end points within participants' accelerometry datasets.

Following this identification, the activity index (AI) was calculated. AI was proposed by Bai et al. and is more sensitive to sedentary behavior and low-intensity PA than activity counts, a common accelerometry metric (*64*). AI has been shown to differentiate PA intensities with higher accuracy than activity counts (*64*). Unlike ActiGraph and Actical activity counts, which use proprietary algorithms, AI can be calculated by the researcher from raw accelerometry data, using an open-source algorithm; consequently, it can be easily computed from any accelerometry device that reports raw data, like the Empatica e4 wristband. The AI metric reflects a ratio of the magnitude of the accelerometry data from the study period to the magnitude of the data when the accelerometer is still and not worn. To obtain data from the still accelerometer (i.e., the comparison data), each of the three Empatica e4 devices used in the current study was left on a still desk (as suggested in Bai et al. [*64*]) for approximately 15 min. During each study session, the researcher noted which Empatica e4 wristband was used so that the comparison data for that same device could be used in the calculation of AI.

Using only the data between the determined walking session start and end points, we calculated AI over a 1-s epoch as presented by Bai et al. using the "ActivityIndex" R package (*64*). To do so, all accelerometry data were first transformed into units of *g* as required by the package. Next, we calculated the mean AI for each session to address the varying session durations. To the best of our knowledge, there are no established thresholds to differentiate intensities of PA for AI; however, mean AI has been previously employed to analyze wrist-mounted accelerometry data (*65*). Mean AI can be interpreted relatively with higher mean AIs indicating higher intensity PA (*64*).

Treatment of Missing Data

For all participants, the e4 wristband reported triaxial accelerometer data. When identifying accelerometry data between the start and conclusion of each walking session, there was no clear threshold for one of the sessions. This session was omitted from the AI analyses. Additionally, two participants in the RW condition showed an abnormally long measurement session (52.7 and 118.4 min) with a large (˜17 and ˜47 min) period of inactivity near the middle/end of the measurement session, which suggests a measurement error. The timestamp of the measurement onset aligned with the scheduled time of the participant session, indicating that the device was most likely not turned off correctly at the end of the session. Since participants donned the device within seconds or minutes of beginning the walking session, and the researchers constrained the walking session to be at most 25 min of active walking, device measurement sessions should not be longer than 35 min even when including time spent stopped at stoplights and other barriers. For these two abnormal sessions, we chose to remove any data occurring after 35 min following the onset of measurement. For both sessions, this point was within the long period of inactivity indicating that it was appropriately separating the participant's walking session from the data occurring afterward. Following removal of these data, the accelerometry data between the start and end of the walking session were determined using the identical method as with other sessions.

The HR data provided by the Empatica e4 wristbands were characterized by a high degree of missingness. This led the researchers to examine the validity of Empatica e4 devices under light PA conditions (e.g., walking). Empatica notes that these devices can produce missing values when the wrist is in motion (*66*), thereby producing biased data under our study's conditions. For this reason, we do not report analyses on HR data.

Results

Fifty-one participants completed their first session. Four participants did not return after their first session (three completed VR, one completed RW), resulting in data for both sessions from 47 participants. The exclusions detailed above produced a final sample size of 46 participants for mean AI analyses and 40 participants for walking duration analyses.

Postexperience Survey

Just under half of the 47 participants had experienced VR previously (22), primarily via video games (16) or educational/entertainment outlets (12) (see Table 1). Roughly 80% of the 47 participants had visited the historic district in real life over 20 times, and 60% had done so over 50 times, indicating most participants had familiarity with the RW location. Participants with experience in VR seemed to walk for longer in VR than those without experience. Additionally, walking duration in VR and the RW may diverge with increasing age. Owing to the limited sample size, inferential statistics are not reported.

Figure 4 visualizes postexperience survey responses concerning perceptions of both environments (Figure 4a), experience of adverse symptoms (Figure 4b), and experiences of the VR environment (Figure 4c). The majority of these participants felt little to no discomfort although this percentage was lower in VR (76% in VR versus 96% in the RW). For a comparison of symptomology, see Figure 4b. Participants felt less comfortable wearing the VR equipment when compared with only the Empatica e4 wristband in the RW, however, 76% still at least agreed that the VR equipment was comfortable (Figure 4a). This trend continued for judgments of safety (VR: 92%, RW: 95%), relaxation (VR: 70%, RW: 100%), and attractiveness (VR: 60%, RW: 98%). Nearly all participants in the RW agreed that the tasks were easy (92%), but only 68% did so in VR. However, more participants completed the RW in their second session (28) than did so in their first (19), so these judgments may be influenced by learning effects.

In relation to experience in VR environments (Figure 4c), 87% of participants adjusted to the VR environment in less than 5 min, whereas two participants never adjusted. Only about a quarter of participants disagreed with the statement that the environment was realistic and believable (26%), with about a third responding neutrally (34%), and 40% agreeing. Lastly, the majority of participants felt immersed in the VR environment (79%).

To better understand which street elements were influential in walking behavior, two raters independently open-coded participants' postsurvey responses. After

Table 1. Participant Characteristics and LPA, MVPA, and Walking Duration across Conditions

Participant characteristic	*n*	LPA duration (min) VR (RW)	MVPA duration (min) VR (RW)	Walking duration (s) VR (RW)
Experience with VR				
No	24	4.92 (.83)	1.79 (12.21)	975.91 (751.68)
Yes	22	5.73 (1.24)	3.86 (13.57)	1,208.94 (760.22)
Number of previous visits to Old Town				
0	0	na	na	na
1–5	3	7.67 (0.00)	4.00 (14.00)	1,130.33 (816.33)
6–20	6	7.00 (1.50)	0.67 (19.83)	1,319.00 (1121.40)
21–50	9	5.56 (0.67)	4.89 (16.44)	987.71 (858.00)
⩾51	29	5.00 (1.11)	2.66 (10.29)	1,053.24 (646.36)
Age				
⩽30	13	6.08 (1.00)	2.85 (16.46)	1,183.00 (905.5)
31–45	21	4.90 (1.10)	2.86 (11.5)	942.22 (724.39)
46–60	4	0.50 (0.75)	0.25 (15.75)	1,113.00 (698.00)
⩾61	9	8.44 (0.89)	4.33 (10.00)	1,270.50 (587.33)

Note: VR = virtual reality; RW = real world; LPA = light physical activity; MVPA = moderate-to-vigorous physical activity; na = not applicable.
Values for *n* may not sum to number of participants as participants were permitted to skip questions. LPA and MVPA durations were calculated using MIMS (described below).

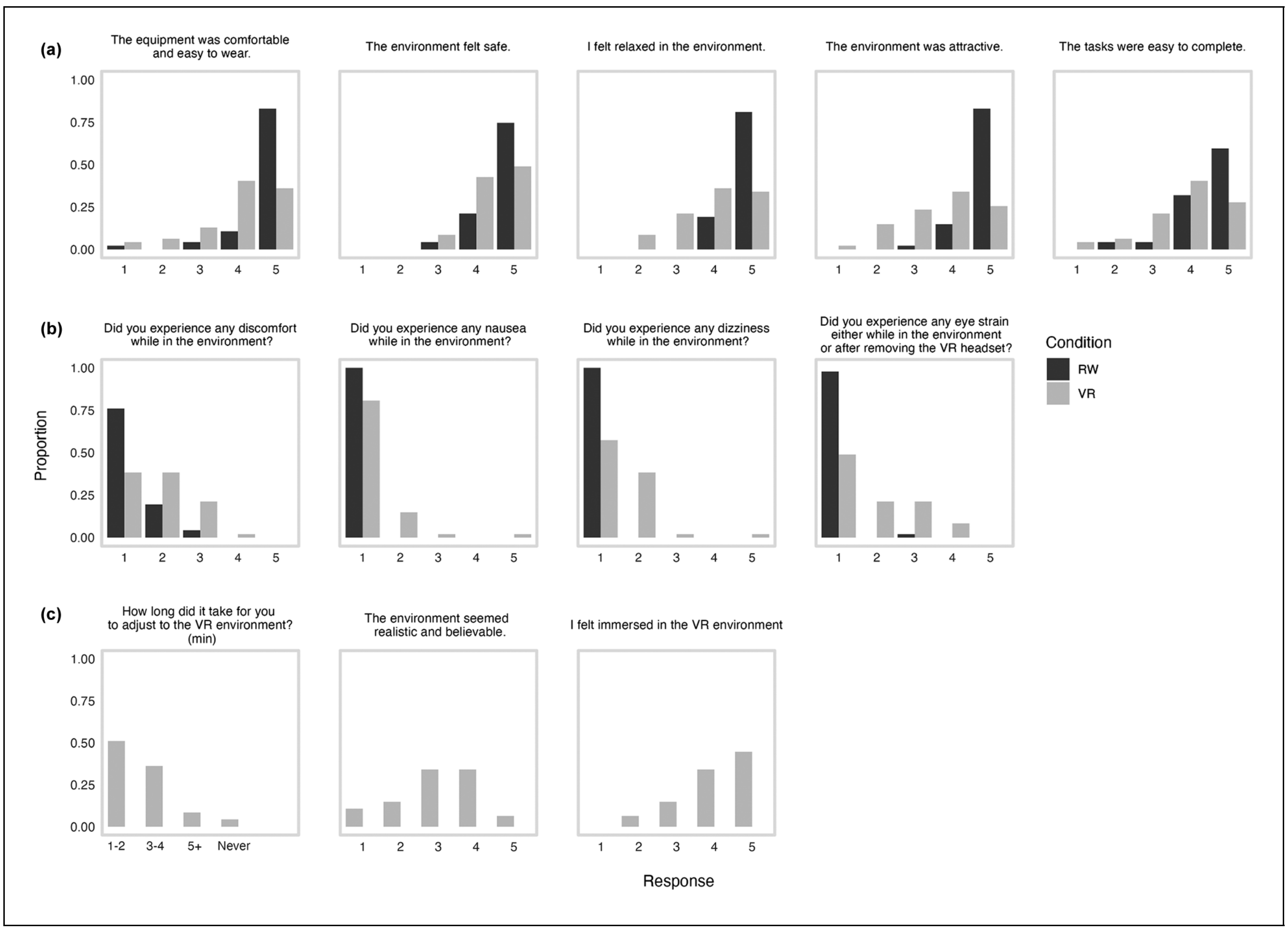

Figure 4. Quantitative survey responses by condition: (*a*) perception statements: 1 = strongly disagree, 5 = strongly agree; (*b*) symptom questions: 1 = none at all, 5 = a great deal; (*c*) VR only statements: 1 = strongly disagree, 5 = strongly agree.
Note: VR = virtual reality.

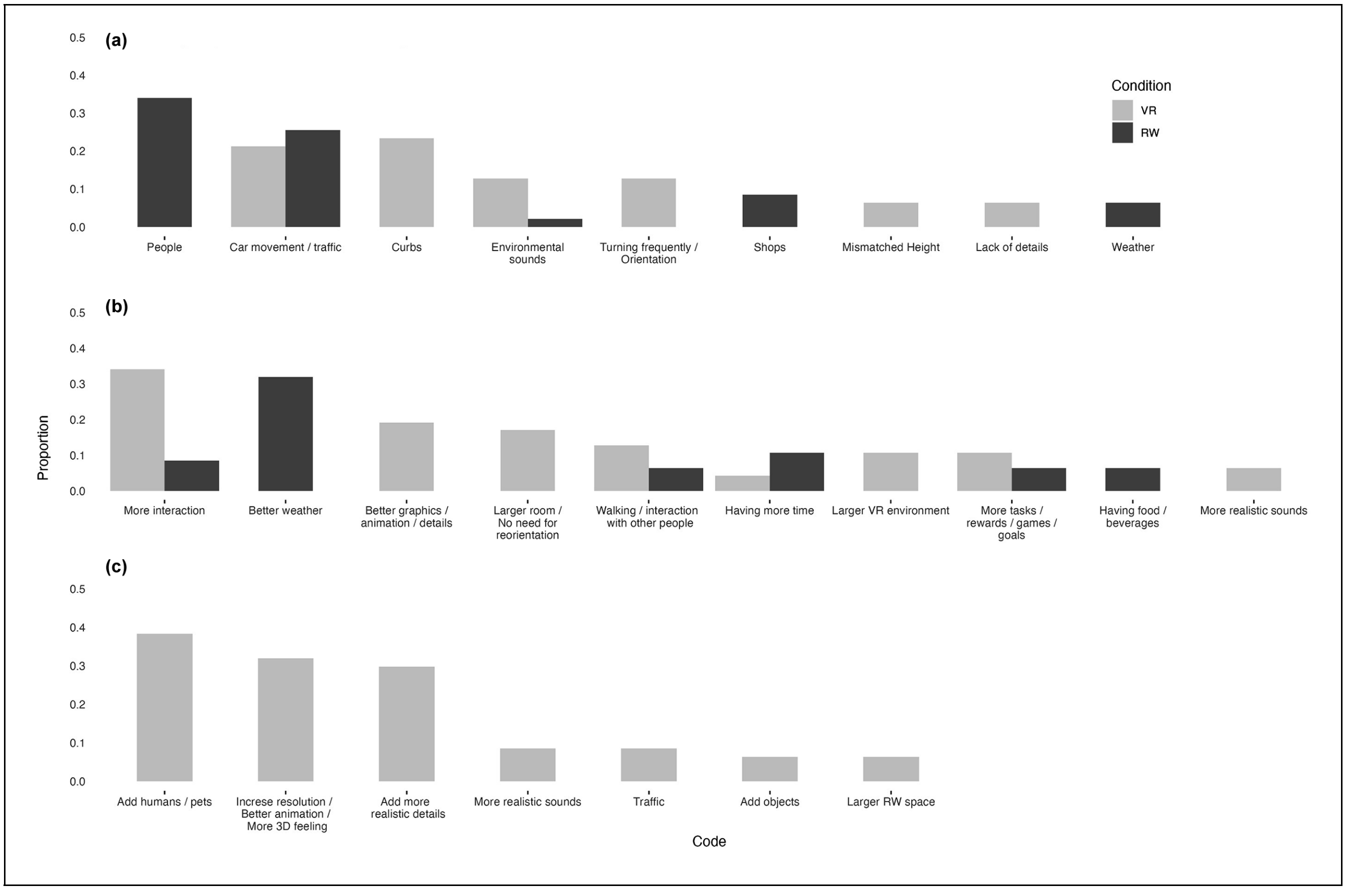

Figure 5. Survey free responses by condition: (*a*) Was anything in the environment distracting? (*b*) What would make you want to spend more time walking around this environment? (*c*) VR only: What changes to the environment would you recommend to make it feel more real/immersive?
Note: VR = virtual reality.

coding, the raters discussed and resolved disagreements. When a resolution was unable to be reached, a third rater was consulted. Codes mentioned by at least three different participants are visualized within Figure 5. Participants commonly identified car movement/traffic (21%) and adjustment to walking with street curbs (23%) as distractions in VR. Three participants found a mismatch between their actual height and their height in VR to be distracting. Participants also frequently found car movement/traffic to be distracting in the RW (26%), as well as the presence of other people (34%). They cited better weather (32%), more interactions (9%), and having more time (11%) as elements that would encourage them to walk more in the RW. In VR, participants discussed wanting more interaction in the virtual environment (34%), including with other people (13%), better graphics and animation (19%), and having a larger walking space in the RW (17%) in response to this question. Lastly, they recommended adding humans and pets (38%), increasing the resolution and improving the animation of the simulation (32%), and adding more realistic details (30%) to make the VR environment more realistic.

Activity Index

The AI difference data did not conform to a normal distribution according to a Shapiro test for normality (p = 0.001). Mean AI data for the RW condition were bimodal whereas the VR mean AI data were unimodal (see Figure 6). The mean of mean AIs in the VR condition was 27.1 (SD = 7.33), and the mean in the RW condition was 32.5 (SD = 26.6). However, further exploration of the mean AI distributions uncovered a concentration of mean AI values in the RW clustered near zero and a second concentration clustered near 50 (see Figure 7). In contrast, mean AI values in VR presented much less spread and centered around 30. This pattern prompted the researchers to investigate the role of arm-swinging in the mean AI values derived from wrist-mounted e4 devices. The e4 wristband is designed to be wrist-worn with a built-in wrist strap (*67*) and is validated for wrist

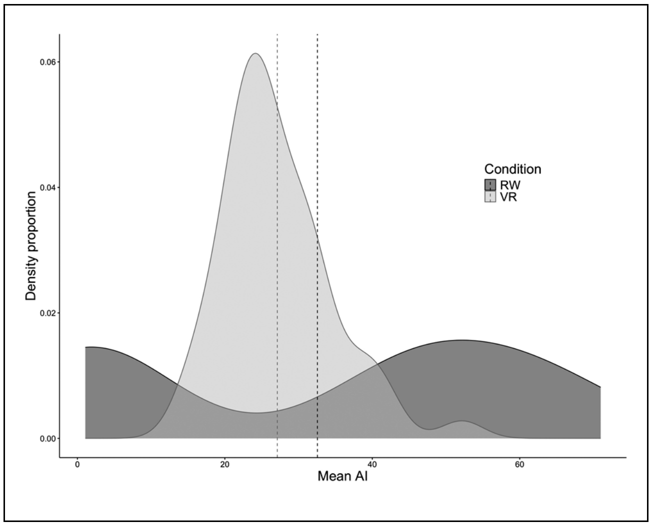

Figure 6. Mean AI distribution by condition.
Note: AI = activity index.

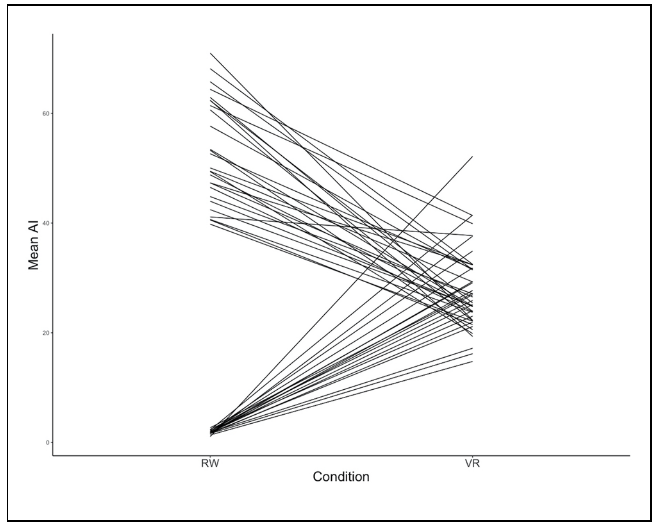

Figure 7. Mean AI by condition per participant.
Note: AI = activity index. Each line represents one participant.

use (*55*). However, participants were free to swing their arms in the RW condition but were constrained to holding the VR controllers in front of them while in the VR condition. Pilot data from one of the researchers (ANS) walking at a standardized speed on a treadmill for 50 s while swinging arms or holding VR controllers demonstrated a discrepancy in mean AI based on arm movement (arm-swinging = 30.30; holding controllers = 20.89). Therefore, using AI created by wrist-mounted e4 devices does not seem to be a useful measure of activity intensity in conditions where arm-swinging may vary systematically. For this reason, no inferential statistics are reported on the AI data.

MIMS

After learning that AI was not a viable accelerometry analytic method for our purposes, we conducted exploratory data analysis of the accelerometry data using MIMS (*68*). MIMS is an open-source raw accelerometry processing approach that is "device-independent" (*68*). Although, like AI, it does seem to be susceptible to arm-swinging, MIMS appears to be a less biased approach. (The same researcher, ANS, walking on a treadmill at a standardized speed while swinging their arms and, subsequently, while holding controllers resulted in MIMS values of 23.1 and 17.4, respectively.)

Karas et al. provide MIMS equivalents to activity count light physical activity (LPA) and moderate-to-vigorous physical activity (MVPA) adult cut-points (*69*), presented in Montoye et al. for accelerometry data processed over a 1-min. epoch (*70*). We employed these MIMS equivalent cut-points: LPA is 15.047 to 19.614 MIMs units, and MVPA > 19.614 MIMs units. First, we used the "MIMSunit" R package (version 0.11.2) (*68*) to calculate MIMS over a 1-min epoch from the raw accelerometry data between the determined walking session start and end points. Next, we applied Karas et al.'s thresholds to calculate the number of minutes spent in LPA and MVPA for each session (*69*).

Because the number of minutes spent in LPA and number of minutes spent in MVPA were positive integers (i.e., counts) and their mean was not sufficiently large enough for the distribution to approximate normality, multiple linear regression (MLR) was not appropriate. To determine whether the number of minutes spent in LPA during the VR session predicted the number of minutes spent in LPA during the RW session, a negative binomial model using a logistic link function was fit. Negative binomial regression was chosen over Poisson regression because the dispersion parameter (DP) was greater than 1 (DP = 1.81), indicating overdispersion and violation of the mean-variance relationship assumption of the Poisson distribution. To account for differences in walking duration, an offset of RW walking duration was included in the model. Including an offset alters the interpretation of exponentiated coefficients from multiplicative differences in the average count of the outcome variable to multiplicative differences in the rate of the outcome variable, meaning that with this offset, our model predicted number of minutes spent in LPA in the RW divided by walking duration in the RW. Lastly, control variables included weather in the RW, whether the participant completed all tasks in the RW (Y = 1, N = 0), and session number of the RW condition (Session 1 = 0, Session 2 = 1). Since temperature

Table 2. Negative Binomial Model Results Regressing Minutes Spent in LPA in the RW on Minutes Spent in LPA in VR, Heat Index, Completion of Tasks, and Session Number with an Offset for RW Walking Duration

Predictor	*b* [95% CI LL, UL]	b^2 [95% CI LL, UL]	df	LR statistic	*p*-Value
(Intercept)	−5.50 [−12.48, 0.80]	0.004 [0.000004, 2.23]			
LPA min in VR	−0.04 [−0.13, 0.04]	0.96 [0.88, 1.04]	1	0.91	0.34
Heat index	0.04 [−0.03, 0.11]	1.04 [0.97, 1.12]	1	1.01	0.32
All tasks completed	0.57 [−0.39, 1.55]	1.77 [0.68, 4.72]	1	1.34	0.25
Session number	−0.53 [−1.51, 0.43]	0.59 [0.22, 1.54]	1	1.16	0.28
McFadden's pseudo R^2 (72)	0.05				

Note: VR = virtual reality; RW = real world; LPA = light physical activity.
95% confidence intervals are presented by lower (*LL*) and upper (*UL*) boundaries. LR statistic presents the calculated statistic for likelihood ratio testing. Dummy codes are as follows: All tasks completed: Y = 1, N = 0; Session number: Session 1 = 0, Session 2 = 1. Pseudo R^2 was calculated using R package "DescTools" (*73*).

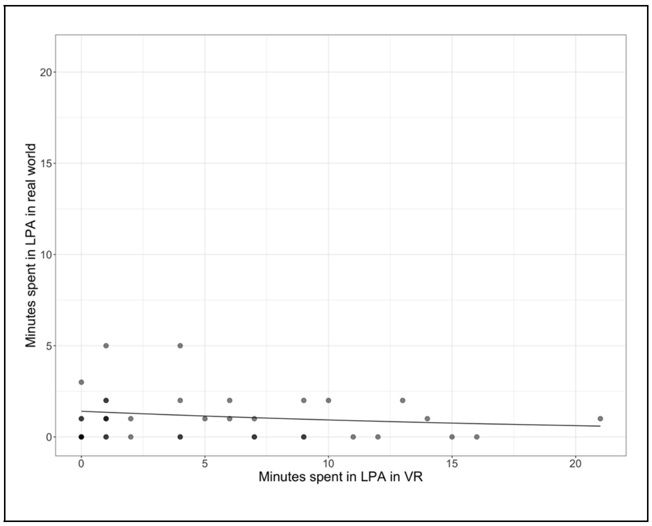

Figure 8. Real world (RW) light physical activity (LPA) × virtual reality (VR) LPA.
Note: Trend line represents the modeled relationship between minutes spent in LPA in VR and minutes spent in LPA in the RW at the median heat index, median RW walking duration, completion of all tasks, and first session number.

and humidity were moderately correlated ($r = -0.68$), they were transformed into a single heat index variable (°F), using the formulas available from the National Oceanic and Atmospheric Administration (*71*) to reflect weather. Evidence for an effect of number of minutes spent in LPA during the VR session was determined by a likelihood ratio test with alpha set at 0.05. The Pearson residual plot indicated a residual outlier. To check for robustness, this model was run with and without this outlier. Removing the outlier did not alter the results of likelihood ratio tests; thus, the model results are presented only for the model including the outlier.

There was no evidence that the number of minutes spent in LPA during the VR session predicted the rate of time spent in LPA during the RW session ($b = -0.04$, chi-square = 0.91, $p = 0.34$; Table 2, Figure 8). When controlling for heat index, completion of all tasks, and session number, each additional 1 min in LPA during the VR session was associated with a 4% decrease in the rate of LPA in the RW (95% CI 12% lower, 4% higher) though this relationship was not significant.

To determine whether number of minutes spent in MVPA during the VR session predicted number of minutes spent in MVPA during the RW session, a Poisson regression model using a logistic link function was fit. The dispersion parameter was less than 1 (DP = 0.66), indicating that the mean-variance relationship assumption was met. Identical to the LPA analysis, an offset of RW walking duration was included in the model to account for differences in walking duration. This inclusion in Poisson regression results in the same alterations to interpretation as in negative binomial models. Control variables were identical to the LPA analysis, and evidence of an effect was determined by a likelihood ratio test with alpha set at 0.05.

There was no evidence that the number of minutes spent in MVPA during the VR session predicted the rate of time spent in MVPA during the RW session ($b = -0.001$, chi-square = 0.01, $p = 0.91$; Table 3, Figure 9). Each additional 1 min in MVPA during the VR session was associated with a 0% change in rate of MVPA in the RW (95% CI 2% lower, 2% higher) when controlling for heat index, completion of all tasks, and session number.

Walking Duration

The walking duration difference data conformed to a normal distribution according to a Shapiro test for normality ($p = 0.20$). Participants walked for about 13 min in the RW condition ($M = 777$ s, SD = 424) and for about 18 min in VR ($M = 1{,}085$, SD = 350). The results

Table 3. Poisson Model Results Regressing Minutes Spent in MVPA in the RW on Minutes Spent in MVPA in VR, Heat Index, Completion of Tasks, and Session Number with an Offset for RW Walking Duration

Predictor	*b* [95% CI LL, UL]	b^2 [95% CI LL, UL]	df	LR statistic	*p*-Value
(Intercept)	−0.76 [−2.09, 0.53]	0.47 [0.12, 1.70]			
MVPA min in VR	−0.001 [−0.02, 0.02]	1.00 [0.98, 1.02]	1	0.01	0.91
Heat index	0.01 [−0.01, 0.02]	1.01 [0.99, 1.02]	1	1.36	0.24
All tasks completed	0.10 [−0.11, 0.31]	1.10 [0.89, 1.36]	1	0.81	0.37
Session number	−0.04 [−0.25, 0.17]	0.96 [0.78, 1.19]	1	0.14	0.71
McFadden's pseudo R^2 (72)	0.01				

Note: VR = virtual reality; RW = real world; MVPA = moderate-to-vigorous physical activity.
95% confidence intervals are presented by lower (*LL*) and upper (*UL*) boundaries. LR statistic presents the calculated statistic for likelihood ratio testing. Dummy codes are as follows: All tasks completed: Y = 1, N = 0; Session number: Session 1 = 0, Session 2 = 1. Pseudo R^2 was calculated using R package "DescTools" (*73*).

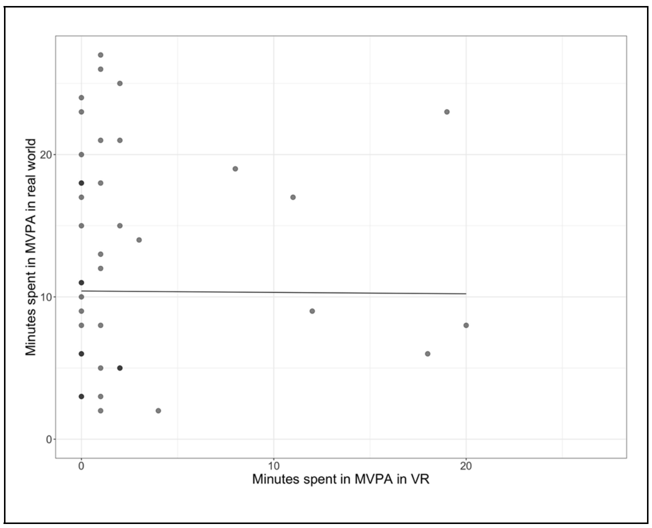

Figure 9. Real world (RW) moderate-to-vigorous physical activity (MVPA) × virtual reality (VR) MVPA.
Note: Trend line represents the modeled relationship between minutes spent in MVPA in VR and minutes spent in MVPA in the RW at the median heat index, median RW walking duration, completion of all tasks, and first session number.

of a paired *t*-test confirmed that this difference was statistically significant ($t = -4.14$, $p < 0.001$). This difference of 325.25 s (95% CI [166.21, 484.29]) or about 5.5 min was a moderate effect ($d = -0.65$).

To control for the completion of all tasks, session number, and heat index, mixed-effects models were performed. Table 4 presents the output of the mixed-effects model using the R packages "lme4" (*74*), "lmerTest" (*75*), and "sjPlot" (*76*). Condition was dummy-coded as 1 for VR and 0 for RW; session number (Session 1 = 0, Session 2 = 1), completion of all tasks (Y = 1, N = 0), and heat index (°F) were included as fixed effects. Participant number was included as a random effect. The QQ plot (quantile-quantile) for this model indicated that the distribution of residuals was approximately normal (Figure 10).

Controlling for heat index, if the participant completed all tasks, and if it was their first or second session, there was evidence of an effect of condition ($t = 2.35$, $p = 0.02$) as calculated using Satterthwaite's method. Based on the model estimate, participants walked 223 s (95% CI [39.29, 408.71]) or roughly 2 min more in VR as compared to the RW, accounting for heat index, task completion, and session number.

We visualized these effects in Figure 11. Visual analysis of this figure suggested order effects: the distributions of VR walking duration for the RW->VR group and VR->RW group as well as the RW walking duration for the RW->VR group seem similar, but the RW walking duration data from the VR->RW group seems to deviate from this distribution as it shows less spread and a concentration of relatively short durations around 500 s (˜8 min). We performed an exploratory analysis of these order effects using a mixed-effects model with an interaction term. Condition (VR = 1, RW = 0), session number (Session 1 = 0, Session 2 = 1), interaction between condition and session number, completion of all tasks (Y = 1, N = 0), and heat index (°F) were included as fixed effects. Participant number was included as a random effect. The QQ plot for this model indicated that the distribution of residuals was approximately normal (Figure 12). The results of this model are presented in Table 5. Importantly, we chose to dummy-code binary variables, meaning that condition and session number effects are simple effects and must be interpreted with reference levels (*77*). In this model, we did not find evidence of a condition simple effect ($t = 0.40$, $p = 0.69$) at the Session 1 reference level, meaning that there was no evidence of a difference between VR and RW conditions among Session 1 observations. There was evidence of a session number simple effect ($t = -2.48$, $p = 0.02$) at the RW reference level: participants completing their second

Table 4. Mixed Effect Model Results Regressing Walking Duration on Main Effects of Condition, Task Completion, Session Number, and Heat Index

Fixed effects					
Predictor	Estimate	95% CI [LL, UL]	df	*t*-Value	*p*-Value
(Intercept)	1138.88	[132.66, 2149.10]	69.09	2.20	0.03*
Condition	223.43	[39.29, 408.71]	60.04	2.35	0.02*
Heat index	0.69	[−11.04, 12.34]	70.50	0.12	0.91
All tasks completed	−551.36	[−678.33, −423.70]	74.31	−8.33	<0.001**
Session number	−95.53	[−202.52, 12.42]	49.29	−1.71	0.09
Marginal R^2	0.61				
Conditional R^2	0.75				

Note: 95% confidence intervals are presented by lower (*LL*) and upper (*UL*) boundaries. VR = virtual reality; RW = real world. Participant ID was included as a random effect. Dummy codes are as follows: Condition: VR = 1, RW = 0; All tasks completed: Y = 1, N = 0; Session number: Session 1 = 0, Session 2 = 1.
*$p < 0.05$; **$p < 0.01$.

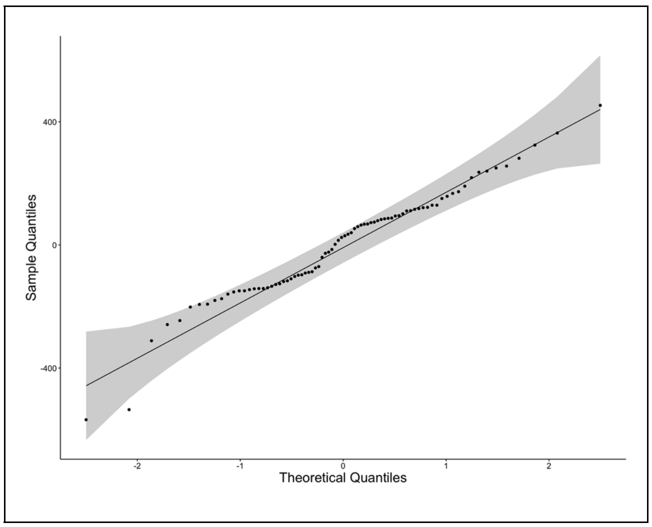

Figure 10. QQ plot for mixed effect model results regressing walking duration on main effects of condition, task completion, session number, and heat index.
Note: QQ = quantile-quantile.

session in the RW walked about 4 min less than participants completing their first session in the RW. This study was not powered to detect interaction as none was expected; nonetheless, the interaction term suggests there may be order effects ($t = 1.82$, $p = 0.08$). These results are best visualized in the bar graph in Figure 13.

We ran pairwise comparisons to follow up on these results using the package "emmeans" (*78*) (see Table 6). Participants in the VR->RW group walked on average 293 s more (95% CI [30.8, 555]) or roughly 5 min more in VR than they walked in the RW (VR = 1,070 s, RW = 777, $t = 2.97$, $p = 0.02$; see Figure 13). This discrepancy was not present among participants in the RW->VR group (RW = 1,016, VR = 1,091, $t = -0.59$, $p = 0.93$).

To provide a useful effect size (i.e., the degree to which walking duration in VR can predict walking duration in the RW), we conducted exploratory analysis using MLR. A Shapiro–Wilks test indicated that the walking duration data in the RW did not conform to a normal distribution ($W = 0.91$, p-value = 0.004); however, MLR is robust against violations of normality. To verify our results using MLR, we employed bootstrapped confidence intervals. Model 1 used walking duration in VR (s), heat index in the RW (°F), task completion in the RW (Y = 1, N = 0), and session number (Session 1 = 0, Session 2 = 1) in the RW to predict walking duration in the RW (s). Model 2 added an interaction term between walking duration in VR and session number, thereby, accounting for order effects. The interaction model (Model 2) did not fit the data significantly better than the model without the interaction ($F = 0.001$, $p = 0.97$). Model comparison indicated that the inclusion of an interaction term explained no variance above and beyond the main effects for predictors (partial $R^2 < 0.001$). Concerning the best fitting model (Model 1 without the interaction), we did not observe evidence that walking duration in VR predicted walking duration in the RW when controlling for session number, task completion, and heat index ($b = 0.14$, $F = 1.33$, $p = 0.26$). A 1-s increase in walking duration in VR was associated with about a tenth of a second increase in walking duration in the RW, though this relationship was not significant. Walking duration in VR accounted for 1% of the variance in walking duration in the RW not accounted by session number, task completion, and heat index (partial $R^2 = 0.012$). Bootstrapped confidence intervals confirmed these results.

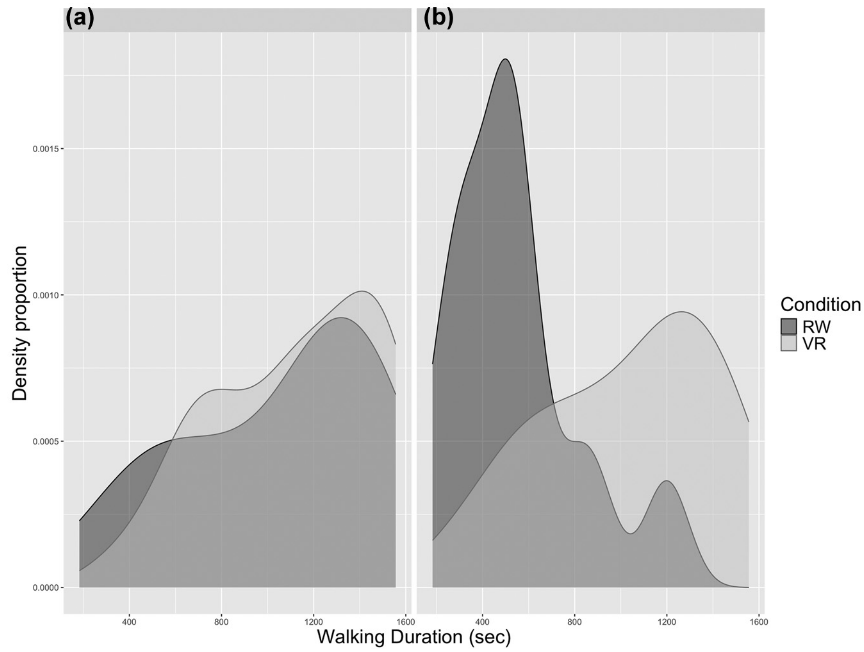

Figure 11. Walking duration distribution by condition and session number: (*a*) Session 1 and (*b*) Session 2.

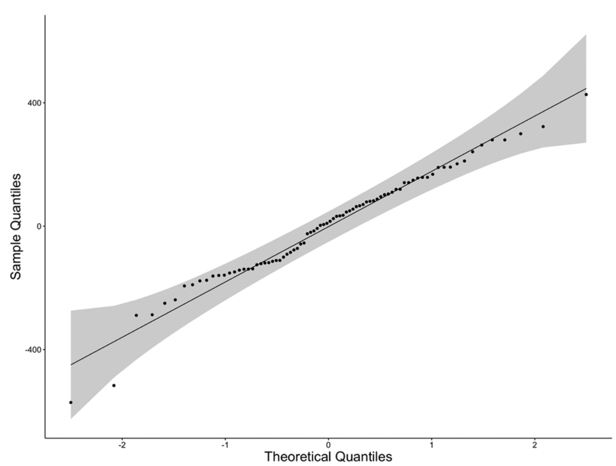

Figure 12. QQ plot for mixed effect model results with added condition-by-session number interaction.
Note: QQ = quantile-quantile.

Discussion

The current study examined how walking decisions in the RW compared with a VR replica of the same environment. If environmental planning decisions are to be informed by data gathered in VR models, it is necessary to know how RW behavior compares with behavior in a VR setting. Poisson and negative binomial regressions did not provide evidence of a relationship between intensity of walking in VR and intensity of walking in the RW when controlling for session number and heat index. The lack of an association between time spent in different PA zones in VR and in the RW may be partially the result of placement of the accelerometry device on the wrist. When choosing this device, we prioritized participant comfort and privacy as they donned the device in public; however, wrist-worn accelerometers are generally less accurate than hip-worn devices (*70*). In their establishment of the PA intensity cut-points used in the present study, Montoye and colleagues discuss how wrist-worn accelerometers can be susceptible to local movement of the arm leading to higher rates of incorrect PA intensity classification (e.g., playing cards categorized as MVPA) (*70*). In the current study, participants' local arm movement was constrained in the VR condition by holding hand controllers, which may explain why there was no association between PA intensity in VR and in the RW.

The results of a *t*-test, mixed-effects model, and MLR indicated that people do not always choose to walk for the same duration in VR as they do in the RW. That is, walking duration was found to be shorter, on average, in the RW than in VR. This may be a factor of perceptual differences in VR versus the RW. As noted, the VR environment was created to match the RW environment complete with building shapes, façades, and legible signage; nonetheless, participants recommended that we

Table 5. Mixed Effect Model Results with Added Condition × Session Number Interaction

Fixed effects					
Predictor	Estimate	95% CI [LL, UL]	df	*t*-value	*p*-value
(Intercept)	1,466.98	[421.32, 2,505.37]	66.96	2.70	<0.01**
Condition	53.73	[−200.80, 307.98]	73.81	0.40	0.69
Heat index	−2.36	[−14.15, 9.50]	67.77	−0.38	0.70
All tasks completed	−537.18	[−662.74, −411.53]	73.71	−8.17	<0.001**
Session number	−239.17	[−424.40, −55.28]	68.81	−2.48	0.02*
Condition × Session	260.41	[−13.89, 537.88]	41.14	1.82	0.08
Marginal R^2	0.63				
Conditional R^2	0.75				

Note: 95% confidence intervals are presented by lower (*LL*) and upper (*UL*) boundaries. VR = virtual reality; RW = real world. Participant ID was included as a random effect. Dummy codes are as follows: Condition: VR = 1, RW = 0; All tasks completed: Y = 1, N = 0; Session number: Session 1 = 0, Session 2 = 1.
*p < 0.05; **p < 0.01.

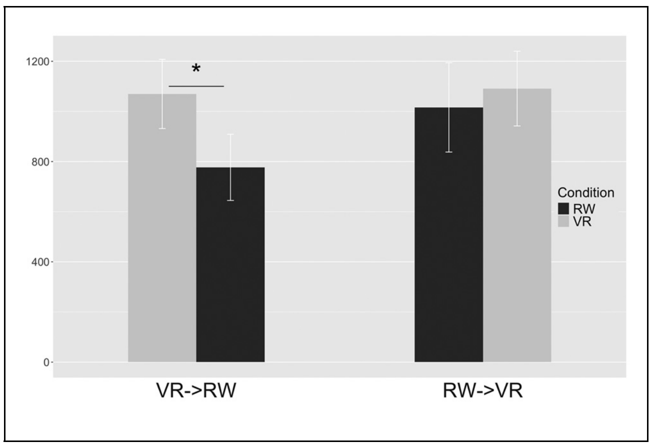

Figure 13. Estimated marginal means of walking duration × condition and randomized order.
Note: Error bars represent 95% confidence intervals.
*p < 0.05; **p < 0.01.

increase the clarity of the simulation environment in the free-response survey. Not only would the fidelity of the environmental design have a potential effect on walkability perceptions, but there are inherent perceptual differences in viewing VR environments over RW. For instance, Bach et al. reported individuals consistently underestimating distances in VR: a perception that became exacerbated as the spatial distances increased (*79*). Also, Bogon et al. demonstrated that individuals expect the duration of events (e.g., how long it would take for a car to drive down the street) to be longer in VR compared with the RW even though individuals experience time similarly in both environments (*80*). Further exploration of perceptual differences may help to explain the subtle variations in walking behavior between VR and RW. Additionally, there was an observed divide in VR walking duration between participants with and without previous VR experience. Familiarity with VR technology may have influenced behavior in VR, including participant motivation.

HR could not be incorporated owing to an error associated with the Empatica e4 devices. This study was important in understanding the need for clarification of the context in which repeated measures may contribute to fatigue and order effects, and the need for understanding HR through more reliable measures of PA when comparing behavior in a virtual and RW environment. Thus, further investigation is needed to understand the

Table 6. Pairwise Comparison Results of Condition and Session Number on Walking Duration

Contrast	Estimate	95% CI [LL, UL]	df	*t*-value	*p*-value
RW Session 1—VR Session 2 (RW -> VR group)	−75.0	[−408.6, 259]	59.9	−0.59	0.93
VR Session 1—RW Session 2 (VR -> RW group)	292.9	[30.8, 555]	49.1	2.97	0.02*

Note: 95% confidence intervals are presented by lower (*LL*) and upper (*UL*) boundaries. RW = real world; VR = virtual reality.
*p < 0.05.

circumstances in which these differences exist before definitive conclusions can be drawn concerning the use of VR in decision making for urban planning.

Analysis of free-response items in the postexperience survey demonstrated that participants were similarly distracted by cars in RW and VR environments, but the primary distractor in the RW was other people, which were not present in VR, though participants commonly suggested that we add them. Participants were also distracted by the presence of curbs in VR, possibly because they received no tactile feedback to match the visual curbs (i.e., participants saw a curb in VR but physically were walking on a flat plane). When asked what would encourage them to spend more time walking in the VR environment, participants focused primarily on simulation limitations, such as the lack of interaction, graphic features, and the size mismatch between the VR environment and the physical room where the VR session was conducted. These priorities carried into their responses for suggested changes to the VR environment, including increasing the resolution and depth of the simulation. There was less consensus among participants in the RW concerning what would encourage them to walk longer, with the most common answer being better weather. Future researchers may consider conducting qualitative work to gain more understanding of how the experiences of an urban environment in VR and in the RW compare.

Exploratory analyses suggested that the relationship between condition and walking duration may be more complex than a simple increase in duration when navigating a VR environment. Moreover, walking duration in VR may match walking duration in the RW under certain conditions. A mixed-effects model revealed that there was no longer evidence of a condition simple effect when the interaction between condition and session number was included in the model. This implies order effects: the sequence by which one walks in VR and walks in the RW may matter. Pairwise comparisons of this interaction showed that the RW walking duration of the VR->RW group was significantly shorter than their VR walking duration. Conversely, VR walking duration and RW walking duration did not differ among participants in the RW->VR group. However, these results must be replicated before strong conclusions can be drawn.

Possible explanations were investigated for these order effects. First, participants completed the RW condition in different weeks depending on their assigned order owing to logistical limitations with regard to renting a large enough space to conduct the VR condition. Order randomization was completed before participants were provided with the opportunity to sign up for a session. It is possible that this resulted in unknown but meaningful differences between groups. The participants who were available during the first week of data collection (RW->VR group) may have differed from the participants who were available during the last week of data collection (VR->RW group). Notably, the proportion of participants reporting previous experience with VR was equivalent across the RW->VR group (47%) and the VR->RW group (48%). Second, weather differed between the first week of data collection and the last week of data collection. This difference seems unlikely to explain our observed order differences, as the weather in Week 1 was presumably less pleasant for walking outside (mean heat index = 87.3°F) than in Week 3 (mean heat index = 81.2°F) when participants walked less. Moreover, Martins et al.'s results depict that heat is negatively correlated with engagement in PA during summer (*81*); thus, weather did not explain our order effects. Third, history effects may explain these results. We primarily recruited employees from a large university in the Western United States (83%): the third week of our data collection coincided with the penultimate week of this institution's summer courses. Participants walking in Week 3 (VR->RW group) may have been short on time given this pending deadline and chose to end their walking session sooner.

Lastly, there may be true order effects in the comparability of RW and VR walking. Order effects have been shown when measuring recall of distance perception in VR, and it is possible that memory plays a role when estimating how long it takes to walk certain distances depending on the order of RW and VR walking scenarios (*82*). However, other previous studies did not find differences in walking duration (*83*) or order effects of RW and VR walking (*41*). Pastel and colleagues randomized order of participating in RW and VR walking (*41*). They also mirrored the walking route in VR to minimize learning effects, which may have contributed to this lack of order effects (*41*, *82*). In our study, we can identify two potential sources of true order effects. First, one of the study tasks asked participants to search for an obscure location (i.e., the shipping courier); the difficulty of navigating and therefore finding this location may have differed in VR than in the RW, leading to order effects. A larger proportion of participants found the tasks to be more challenging in VR. In direct contradiction, the proportion of participants who completed all tasks in their first session was comparable regardless of whether that first session was in VR (43%) or the RW (41%). An alternative explanation is that the experience of finding this specific location transferred better from VR to the RW compared with the RW to VR. In other words, for the VR->RW group, the skill of finding a location may have transferred well to the RW, leading to a large decrease in walking duration for the RW condition. However, for the RW->VR group, this skill may not have transferred well, so their second session (VR) was a comparable length as their first (RW). Of those in the VR->RW group who did not complete all tasks in their first session (13), 92% completed all tasks in their second session in the RW (12). In contrast, of those in the RW->VR group who did not complete all tasks in their first session (10), only 40% completed all tasks in their second session (4).

There were some mixed results with regard to the differences in VR versus RW navigation. Some have speculated that navigating in VR is more difficult than in the RW because of difficulty moving one's head to walk in the desired direction, which often requires more coordination than RW walking (*84*). Similarly, participants in VR conditions have demonstrated weaker navigational skills in regard to pointing to start and finish lines as well as drawing a map of the route they walked compared with those in RW conditions (*85*). To mitigate head movement effects while navigating, Drewes et al. had participants hold a system controller close to their body to orient themselves while making walking decisions (*83*). Interestingly, they found that there were no significant navigational differences between VR and RW walking. These varied findings suggest that following directions and navigating in VR may pose more challenges but do not necessarily play a large role in order effect differences.

The second source of true order effects could be that participants may have volunteered for the study because they were excited to experience VR. By their second session, participants in the VR->RW group may have already performed the more attractive session of the study. Knowing the study session ends when they complete the walking session, they may have chosen, consciously or not, to end the walking session sooner. If future confirmatory studies indicate that there are true order effects either because of navigational differences or participant motivation, practitioners interested in using VR to examine potential built environmental interventions might consider strategies to minimize these effects, such as providing participants directions in VR, recruiting participants through diverse channels, or selecting repeated-measure study designs that account for carry-over effects.

Strengths

This study successfully compared actual decisions about walking behaviors in VR with those in the RW using the VR locomotive technique (i.e., overground walking) that individuals use in the RW. Prior work has focused on walking-related perceptions such as judgments of pleasantness, safety, or intention to walk (*37*, *48*, *49*) instead of actual walking. Other researchers compared the kinematics between walking in a VR room with walking in an RW room identical to the VR room (*39*, *47*); however, this does not address the decision to continue or to stop walking in an urban environment, which is central to walkability research. This study is unique in that it empirically examined the similarity of actual walking decisions made in VR and the RW at an urban scale using overground walking. Furthermore, the walking decisions made by participants in this study mirrored the types of decisions made while walking in urban environments. We believe this study brings PA researchers one step closer to using VR to address the causal gap between built environmental factors and the decision to engage in walking.

Limitations

The findings of this study must be understood in the context of its limitations. The manufacturer of the HMD used in the present study discourages its use outdoors as sunlight can harm the device; consequently, the VR condition took place inside a climate-controlled room, whereas the RW condition used an outdoor location where the heat index could vary throughout the walking area. That is to say, condition was confounded by weather. To reduce the effect of this confounding, heat index was included in multilevel and regression models; nonetheless, the results of statistical models with intercorrelated predictors can be unstable (*86*, *87*). Weather variability presents a persistent challenge for individuals aiming to utilize VR technology to simulate RW behaviors, as the current limitations of VR technology (i.e., the sensitivity of VR equipment to both light and heat) restrict its application to indoor locations. Also, because the historic district and its simulation are larger than the room in which participants walked in VR, participants needed to pause their walking, turn their bodies, and reorient the VR environment. This added a step in VR, which was not necessary in the RW, and may have affected walking decisions. Additionally, owing to the limited availability of the large indoor space, all VR sessions occurred within the second week of data collection whereas the RW sessions either occurred in the first or third week. This necessity created the possibility of selective attrition bias, weather differences, and the history effects discussed above. Furthermore, using cut-points to categorize time epochs into PA intensity levels has inherent limitations (*88*), including compounding measurement error from the study that employs cut-points with that of the cut-point validation study. Separating LPA and MVPA into two separate analytic models may have masked a relationship between them owing to range restriction.

Gait-based locomotion for VR has been shown to reduce the risk of motion sickness and is preferred by users over other locomotion techniques (*89*, *90*). As this was the primary mode of locomotion for the user the effects of simulation sickness were minimal. The other two locomotion techniques integrated into the experience, snap turns and teleportation, would be used minimally by the user. These techniques have also been shown to reduce discomfort for users (*91*). However, despite these design considerations there is a chance that participants were acutely aware of the VR headset, which in turn could have increased perceived levels of fatigue, resulting in a desire to remove the HMD as soon as possible (*92*). This may be reflected in the slightly higher reported symptoms in relation to the statements about VR, presented in Figure 4b.

Future Directions

Future researchers should avoid using the Empatica e4 wristband for measuring HR under movement conditions. Although the Empatica e4 wristband contains an accelerometer and is marketed as research-grade, its HR functionality is not validated for use when the wrist is moving. Moreover, researchers should abstain from using wrist-worn accelerometers to measure PA intensity for VR research as arm, and thereby, wrist, movement in VR (i.e., holding controllers) does not parallel that of RW walking (i.e., arm-swinging).

Further research is needed to explore the order effects suggested by the exploratory analyses of this study. Research using a similar procedure as the current study but completing both VR and RW sessions within the same time frame could determine whether the order effects observed here are meaningful to all researchers using VR to study walkability or an artifact of our research design. Research teams might additionally consider highly powered study designs to enable examination of individual factors affecting the similarity of VR and RW walking, such as VR experience and age. Lastly, as VR hardware and software continue to develop, it is important to continue testing the comparability of VR walking decisions with RW walking decisions if VR is to be used as a highly controlled and ecologically valid technique for studying walking behavior.

Conclusions

Walking decisions in VR may differ from those made in the RW. These differences may only be present under specific circumstances. We encourage researchers to investigate the conditions under which walking decisions in VR may equate to those made in the RW and the conditions under which they are not similar. Through understanding that there might be differences in how individuals interact with VR (versus RW) environments, simulations can be adjusted or interpreted with these discrepancies in mind. This research is necessary for the emerging use of VR to evaluate the causal influence of specific built environmental factors on walking behaviors.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: A. N. Spitzer, K. M. Oselinsky, D. J. Graham; data collection: A. N. Spitzer, M. R. Ramey, Y. "S." Yu; analysis and interpretation of results: A. N. Spitzer, M. R. Ramey, Y. "S." Yu; draft manuscript preparation: A. N. Spitzer, K. M. Oselinsky, D. J. Graham, M. R. Ramey, Y. "S." Yu, K. McMahon, B. Kelley, D. Dean, S. B. LoTemplio, D. Rojas-Rueda, F. R. Ortega. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a College of Natural Sciences Dean's Faculty Success Initiative, Colorado School of Public Health Faculty Seed grant, and by the City of Fort Collins.

ORCID iDs

Amanda N. Spitzer https://orcid.org/0000-0002-3045-5626
Matea R. Ramey https://orcid.org/0009-0000-9945-320X
Yiqing "Skylar" Yu https://orcid.org/0009-0009-6512-2803
Katrina M. Oselinsky https://orcid.org/0009-0005-3757-7860
Katie McMahon https://orcid.org/0009-0005-0918-4023
Brendan Kelley https://orcid.org/0000-0002-3415-3836
David Rojas-Rueda https://orcid.org/0000-0001-5854-2484
Daniel Dean https://orcid.org/0000-0002-6276-2524
Sara B. LoTemplio https://orcid.org/0000-0003-0142-3702
Dan J. Graham https://orcid.org/0000-0001-9993-2585

Data Accessibility Statement

The data that support the findings of this study will be openly available in OSF at https://osf.io/s395w/.

References

1. Seefeldt, V, R. M Malina, and M. A Clark. Factors Affecting Levels of Physical Activity in Adults. *Sports Medicine*, Vol. 32, 2002, pp. 143–168. https://doi.org/10.2165/00007256-200232030-00001
2. Heath, G. W, R. C Brownson, J Kruger, R Miles, K. E Powell, and L. T Ramsey. The Effectiveness of Urban Design and Land Use and Transport Policies and Practices to Increase Physical Activity: A Systematic Review. *Journal of Physical Activity and Health*, Vol. 3, 2006, pp. S55–S76. https://doi.org/10.1123/jpah.3.s1.s55

3. Ding, D, A Ramirez Varela, A. E Bauman, U Ekelund, I.-M Lee, G Heath, P. T Katzmarzyk, R Pratt, and M Reis. Towards Better Evidence-Informed Global Action: Lessons Learnt from the Lancet Series and Recent Developments in Physical Activity and Public Health. *British Journal of Sports Medicine*, Vol. 54, 2020, pp. 462–468. https://doi.org/10.1136/bjsports-2019-101001
4. Murtagh, E. M, M. H Murphy, and J Boone-Heinonen. Walking – the First Steps in Cardiovascular Disease Prevention. *Current Opinion in Cardiology*, Vol. 25, 2010, pp. 490–496. https://doi.org/10.1097/HCO.0b013e32833ce972
5. Wilson, D. B, J. S Porter, G Parker, and J Kilpatrick. Anthropometric Changes Using a Walking Intervention in African American Breast Cancer Survivors: A Pilot Study. *Preventing Chronic Disease*, Vol. 2, 2005, p. A16.
6. Yaffe, K, D Barnes, M Nevitt, L. Y Lui, and K Covinsky. A Prospective Study of Physical Activity and Cognitive Decline in Elderly Women: Women Who Walk. *Archives of Internal Medicine*, Vol. 161, 2001, pp. 1703–1708. https://doi.org/10.1001/archinte.161.14.1703
7. Gatrell, A. C Therapeutic Mobilities: Walking and 'Steps' to Wellbeing and Health. *Health and Place*, Vol. 22, 2013, pp. 98–106. https://doi.org/10.1016/j.healthplace.2013.04.002
8. Bratman, G. N, G. C Daily, B. J Levy, and J. J Gross. The Benefits of Nature Experience: Improved Affect and Cognition. *Landscape and Urban Planning*, Vol. 138, 2015, pp. 41–50. https://doi.org/10.1016/j.landurbplan.2015.02.005
9. De Brito, J. N, Z. C Pope, N. R Mitchell, I. E Schneider, J. M Larson, T. H Horton, and M. A Pereira. Changes in Psychological and Cognitive Outcomes after Green versus Suburban Walking: A Pilot Crossover Study. *International Journal of Environmental Research and Public Health*, Vol. 16, 2019, p. 2894. https://doi.org/10.3390/ijerph16162894
10. Gidlow, C. J, M. V Jones, G Hurst, D Masterson, D Clark-Carter, M. P Tarvainen, G Smith, and M Nieuwenhuijsen. Where to Put Your Best Foot Forward: Psycho-Physiological Responses to Walking in Natural and Urban Environments. *Journal of Environmental Psychology*, Vol. 45, 2016, pp. 22–29. https://doi.org/10.1016/j.jenvp.2015.11.003
11. Scott, E. E, K. W Crabtree, A. S McDonnell, S. B LoTemplio, G. D Mc Nay, and D. L Strayer Measuring Affect and Complex Working Memory in Natural and Urban Environments. *Frontiers in Psychology*, Vol. 14 2023, P. 1039334. https://doi.org/10.3389/fpsyg.2023.1039334
12. Ewing, R Can the Physical Environment Determine Physical Activity Levels? *Exercise and Sport Sciences Reviews*, Vol. 33, 2005, p. 69.
13. Hickman, R, O Ashiru, and D Banister. Transport and Climate Change: Simulating the Options for Carbon Reduction in London. *Transport Policy*, Vol. 17, 2010, pp. 110–125. https://doi.org/10.1016/j.tranpol.2009.12.002
14. Kim, H, and S Yang Neighborhood Walking and Social Capital: The Correlation between Walking Experience and Individual Perception of Social Capital. *Sustainability*, Vol. 9, 2017, P. 680. https://doi.org/10.3390/su9050680
15. Stroope, J Active Transportation and Social Capital: The Association between Walking or Biking for Transportation and Community Participation. *Preventive Medicine*, Vol. 150, 2021, P. 106666. https://doi.org/10.1016/j.ypmed.2021.106666
16. Arellana, J, V Alvarez, D Oviedo, and L. A Guzman Walk This Way: Pedestrian Accessibility and Equity in Barranquilla and Soledad, Colombia. *Research in Transportation Economics*, Vol. 86, 2021, P. 101024. https://doi.org/10.1016/j.retrec.2020.101024
17. Kang, C.-D Spatial Access to Pedestrians and Retail Sales in Seoul, Korea. *Habitat International*, Vol. 57, 2016, pp. 110–120. https://doi.org/10.1016/j.habitatint.2016.07.006
18. Soni, N, and N Soni. Benefits of Pedestrianization and Warrants to Pedestrianize an Area. *Land Use Policy*, Vol. 57, 2016, pp. 139–150. https://doi.org/10.1016/j.landusepol.2016.05.009
19. De Vries, S. I, I Bakker, W Van Mechelen, and M Hopman-Rock. Determinants of Activity-Friendly Neighborhoods for Children: Results from the Space Study. *American Journal of Health Promotion*, Vol. 21, 2007, pp. 312–316. https://doi.org/10.4278/0890-1171-21.4s.312
20. Miyake, K. K, A. R Maroko, K. L Grady, J. A Maantay, and P. S Arno. Not Just a Walk in the Park: Methodological Improvements for Determining Environmental Justice Implications of Park Access in New York City for the Promotion of Physical Activity. *Cities and the Environment*, Vol. 3, 2010, pp. 1–17. https://doi.org/10.15365/cate.3182010
21. Estacio, I, R Hadfi, A Blanco, T Ito, and J Babaan. Optimization of Tree Positioning to Maximize Walking in Urban Outdoor Spaces: A Modeling and Simulation Framework. *Sustainable Cities and Society*, Vol. 86, 2022, P. 104105. https://doi.org/10.1016/j.scs.2022.104105
22. Besser, L. M, and D. P Mitsova. Neighborhood Green Land Cover and Neighborhood-Based Walking in U.S. Older Adults. *American Journal of Preventive Medicine*, Vol. 61, 2021, pp. e13–e20. https://doi.org/10.1016/j.amepre.2021.01.013
23. Sun, G, E. Y Choe, and C Webster. Natural Experiments in Healthy Cities Research: How Can Urban Planning and Design Knowledge Reinforce the Causal Inference? *Town Planning Review*, Vol. 94, 2023, pp. 87–108. https://doi.org/10.3828/tpr.2022.14
24. Nakamura, K Experimental Analysis of Walkability Evaluation Using Virtual Reality Application. *Environment and Planning B: Urban Analytics and City Science*, Vol. 48, No. 8, 2020, pp. 2481–2496. https://doi.org/10.1177/2399808320980747
25. Oselinsky, K, A. N Spitzer, Y Yu, F. R Ortega, L. H Malinin, K. A Curl, H Leach, and D. J Graham. Virtual Reality Assessment of Walking in a Modifiable Urban Environment: A Feasibility and Acceptability Study. *Scientific Reports*, Vol. 13, 2023, P. 5867. https://doi.org/10.1038/s41598-023-32139-w
26. Berkman, M. I, and E Akan. Presence and Immersion in Virtual Reality. In *Encyclopedia of Computer Graphics and Games* (N Lee ed.), Springer International Publishing, Cham, 2019, pp. 1–10.
27. Meier, B. P, S Schnall, N Schwarz, and J. A Bargh. Embodiment in Social Psychology. *Topics in Cognitive Science*,

Vol. 4, 2012, pp. 705–716. https://doi.org/10.1111/j.1756-8765.2012.01212.x

28. Cmentowski, S, and J Krueger. Exploring the Potential of Vertical Jump Training in Virtual Reality. *Proc., Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play, ACM: Virtual Event Austria*, October 15, 2021, pp. 179–185
29. Schulz, P, D Alexandrovsky, F Putze, R Malaka, and J Schöning The Role of Physical Props in VR Climbing Environments. Proc., *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, Glasgow, Scotland, UK*, May 2, 2019, pp. 1–13.
30. Epic Games. Unreal Engine: The Most Powerful Real-Time 3D Creation Tool. https://www.unrealengine.com/en-US. Accessed June 8, 2023.
31. Unity. Unity Real-Time Development Platform: 3D, 2D, VR & AR Engine. https://unity.com. Accessed June 8, 2023.
32. Bebko, A. O, and N. F Troje. BmlTUX: Design and Control of Experiments in Virtual Reality and Beyond. *i-Perception*, Vol. 11, 2020, 204166952093840. https://doi.org/10.1177/2041669520938400
33. Chung, S, and J. V Sparks. Motivated Processing of Peripheral Advertising Information in Video Games. *Communication Research*, Vol. 43, 2016, pp. 518–541. https://doi.org/10.1177/0093650214566623
34. Wilson, C. J, and A Soranzo The Use of Virtual Reality in Psychology: A Case Study in Visual Perception. *Computational and Mathematical Methods in Medicine*, Vol. 2015, 2015, Pp. 151702. https://doi.org/10.1155/2015/151702
35. Slater, M Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, Vol. 364, 2009, pp. 3549–3557. https://doi.org/10.1098/rstb.2009.0138
36. Deb, S, D. W Carruth, R Sween, L Strawderman, and T. M Garrison. Efficacy of Virtual Reality in Pedestrian Safety Research. *Applied Ergonomics*, Vol. 65, 2017, pp. 449–460. https://doi.org/10.1016/j.apergo.2017.03.007
37. Bhagavathula, R, B Williams, J Owens, and R Gibbons. The Reality of Virtual Reality: A Comparison of Pedestrian Behavior in Real and Virtual Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62, 2018, pp. 2056–2060. https://doi.org/10.1177/1541931218621464
38. Feldstein, I. T, F. M Kölsch, and R Konrad. Egocentric Distance Perception: A Comparative Study Investigating Differences between Real and Virtual Environments. *Perception*, Vol. 49, 2020, pp. 940–967. https://doi.org/10.1177/0301006620951997
39. Janeh, O, N Katzakis, J Tong, and F Steinicke. Infinity Walk in VR: Effects of Cognitive Load on Velocity during Continuous Long-Distance Walking. Proc., ACM Symposium on Applied Perception 2019, Association for Computing Machinery, New York, NY, September 19, 2019, pp. 1–9.
40. Horsak, B, M Simonlehner, L Schöffer, B Dumphart, A Jalaeefar, and M Husinsky. Overground Walking in a Fully Immersive Virtual Reality: A Comprehensive Study on the Effects on Full-Body Walking Biomechanics. *Frontiers in Bioengineering and Biotechnology*, Vol. 9, 2021, p. 780314.
41. Pastel, S, C. H Chen, D Bürger, M Naujoks, L. F Martin, K Petri, and K Witte. Spatial Orientation in Virtual Environment Compared to Real-World. *Journal of Motor Behavior*, Vol. 53, 2021, pp. 693–706. https://doi.org/10.1080/00222895.2020.1843390
42. Bühler, M. A, and A Lamontagne. Coordinating Clearance and Postural Reorientation When Avoiding Physical and Virtual Pedestrians. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 30, 2022, pp. 1612–1620. https://doi.org/10.1109/TNSRE.2022.3181817
43. Kelly, J. W, T. A Doty, M Ambourn, and L. A Cherep. Distance Perception in the Oculus Quest and Oculus Quest 2. *Frontiers in Virtual Reality*, Vol. 3, 2022, p. 850471.
44. Mason, A. H, A. S Padilla, A Peer, M Toepfer, K Ponto, and K. A Pickett. The Role of the Visual Environment on Characteristics of Over-Ground Locomotion in Natural and Virtual Environments. *International Journal of Human-Computer Studies*, Vol. 169, 2023, P. 102929. https://doi.org/10.1016/j.ijhcs.2022.102929
45. Rhiu, I, Y. M Kim, W Kim, and M. H Yun. The Evaluation of User Experience of a Human Walking and a Driving Simulation in the Virtual Reality. *International Journal of Industrial Ergonomics*, Vol. 79, 2020, P. 103002. https://doi.org/10.1016/j.ergon.2020.103002
46. Varas-Diaz, G, S Paralkar, S Wang, and T Bhatt. Kinematic Gait Adjustments to Virtual Environments on Different Surface Conditions: Do Treadmill and Over-Ground Walking Exhibit Different Adaptations to Passive Virtual Immersion? *Rehabilitation Research and Practice*, Vol. 2020, 2020, p. e8901973. https://doi.org/10.1155/2020/8901973
47. Canessa, A, P Casu, F Solari, and M Chessa. Comparing Real Walking in Immersive Virtual Reality and in Physical World Using Gait Analysis. *Proc., 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, SCITEPRESS - Science and Technology Publications, Prague, Czech Republic*, 2019, pp. 121–128.
48. Nishio, S, and F Ito. Statistical Validation of Utility of Head-Mounted Display Projection-Based Experimental Impression Evaluation for Sequential Streetscapes. *Environment and Planning B: Urban Analytics and City Science*, Vol. 47, 2020, pp. 1167–1183. https://doi.org/10.1177/2399808318821693
49. Shi, J, T Honjo, K Zhang, and K Furuya. Using Virtual Reality to Assess Landscape: A Comparative Study Between On-Site Survey and Virtual Reality of Aesthetic Preference and Landscape Cognition. *Sustainability*, Vol. 12, 2020, p. 2875. https://doi.org/10.3390/su12072875
50. Doesken, N. J, R. A Pielke, and A.P. B Bliss. *Climate of Colorado*. Colorado Climate Center, Atmospheric Science Department, Colorado State University, Fort Collins, CO, 2003.
51. Farmani, Y, and R Teather. Viewpoint Snapping to Reduce Cybersickness in Virtual Reality. *Proc., 44th Graphics Interface Conference 2018*, Toronto, 2018, pp. 168–175.

52. Delgado, M, and A. N Spitzer. *Fort Collins Post Office, 1911*. Matterport, 2022. https://www.chhs.colostate.edu/dm/programs-and-degrees/galleries-libraries-and-laboratories/architectural-virtual-library/fort-collins-buildings/
53. Empatica. E4 Data - HR.Csv Explanation. https://support.empatica.com/hc/en-us/articles/360029469772-E4-data-HR-csv-explanation. Accessed August 31, 2022.
54. Menghini, L, E Gianfranchi, N Cellini, E Patron, M Tagliabue, and M Sarlo. Stressing the Accuracy: Wrist-Worn Wearable Sensor Validation over Different Conditions. *Psychophysiology*, Vol. 56, 2019, P. e13441. https://doi.org/10.1111/psyp.13441
55. Empatica. How is the Acceleration Data Formatted in E4 Connect? https://support.empatica.com/hc/en-us/articles/202028739-How-is-the-acceleration-data-formatted-in-E4-connect. Accessed August 31, 2022.
56. Spitzer, A. N. How Does Walking down a Virtual Reality Street Compare to Walking down the Same Street in Real Life? *Presented at the 44th Annual Meeting & Scientific Sessions of the Society of Behavioral Medicine*, Phoenix, AZ, 2023
57. Meta. Avoid Sunlight Damage with Your Meta Quest Headset. https://www.meta.com/help/quest/articles/headsets-and-accessories/product-care-and-best-practices/avoiding-sunlight-damage/. Accessed February 10, 2023.
58. VIVE. Caring for Your VIVE. https://www.vive.com/us/support/vive/category_howto/caring-for-your-vive.html. Accessed February 10, 2023.
59. Posit team (2022) RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA
60. R Core Team(2024) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
61. Karas, M, M Straczkiewicz, W Fadel, J Harezlak, C. M Crainiceanu, and J. K Urbanek. Adaptive Empirical Pattern Transformation (ADEPT) with Application to Walking Stride Segmentation. *Biostatistics*, Vol. 22, 2021, pp. 331–347. https://doi.org/10.1093/biostatistics/kxz033
62. Karas, M, J Urbanek, and W Fadel *Adeptdata: Accelerometry Data Sets. CRAN, Vienna, 2021.*
63. Karas, M, C Crainiceanu, and J Urbanek Walking Strides Segmentation with Adept. https://cran.r-project.org/web/packages/adept/vignettes/adept-strides-segmentation.html. Accessed November 11, 2022.
64. Bai, J, C Di, L Xiao, K. R Evenson, A. Z LaCroix, C. M Crainiceanu, and D. M Buchner. An Activity Index for Raw Accelerometry Data and Its Comparison with Other Activity Metrics. *PLoS One*, Vol. 11, 2016, p. e0160644. https://doi.org/10.1371/journal.pone.0160644
65. Khan, N. C, V Pandey, K. Z Gajos, and A. S Gupta. Free-Living Motor Activity Monitoring in Ataxia-Telangiectasia. *The Cerebellum*, Vol. 21, 2022, pp. 368–379. https://doi.org/10.1007/s12311-021-01306-y
66. Empatica. E4 Data - IBI Expected Signal. https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal.
67. Empatica. E4 Wristband | Real-Time Physiological Signals and Wearable PPG, EDA, Temperature, Motion Sensors. https://www.empatica.com/research/e4. Accessed August 31, 2022.
68. John, D, Q Tang, F Albinali, and S Intille. An Open-Source Monitor-Independent Movement Summary for Accelerometer Data Processing. *Journal for the Measurement of Physical Behaviour*, Vol. 2, 2019, pp. 268–281. https://doi.org/10.1123/jmpb.2018-0068
69. Karas, M, J Muschelli, A Leroux, J. K Urbanek, A. A Wanigatunga, J Bai, C. M Crainiceanu, and J. A Schrack. Comparison of Accelerometry-Based Measures of Physical Activity: Retrospective Observational Data Analysis Study. *JMIR MHealth and UHealth*, Vol. 10, 2022, P. e38077. https://doi.org/10.2196/38077
70. Montoye, A. H. K, K. A Clevenger, K. A Pfeiffer, M. B Nelson, J. M Bock, M. T Imboden, and L. A Kaminsky. Development of Cut-Points for Determining Activity Intensity from a Wrist-Worn ActiGraph Accelerometer in Free-Living Adults. *Journal of Sports Sciences*, Vol. 38, 2020, pp. 2569–2578. https://doi.org/10.1080/02640414.2020.1794244
71. NOAA. Heat Index Equation. https://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml. Accessed June 16, 2023.
72. McFadden, D Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments. In *Behavioural Travel Modelling* (D. A. Hensher, and P. R. Stopher, eds.), Routledge, London, 2021, pp. 279–318.
73. Signorell, A *DescTools: Tools for Descriptive Statistics*. CRAN, Vienna, 2025.
74. Bates, D, M Mächler, B Bolker, and S Walker. Fitting Linear Mixed-Effects Models Using Lme4. *Journal of Statistical Software*, Vol. 67, 2015, pp. 1–48. https://doi.org/10.18637/jss.v067.i01
75. Kuznetsova, A, P. B Brockhoff, and R. H. B Christensen. LmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, Vol. 82, 2017, pp. 1–26. https://doi.org/10.18637/jss.v082.i13
76. Lüdecke, D *sjPlot: Data Visualization for Statistics in Social Science*. CRAN, Vienna, 2024.
77. Brehm, L, and P. M Alday Contrast Coding Choices in a Decade of Mixed Models. *Journal of Memory and Language*, Vol. 125, 2022, p. 104334. https://doi.org/10.1016/j.jml.2022.104334
78. Lenth, R. V *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*. 2024.
79. Bach, M, P Jensen, A Lucaci, D Pupczyk, and C Madsen. Effect of Environment Size on Spatial Perception in Virtual Reality: *Proc., 6th International Conference on Computer-Human Interaction Research and Applications, SCITEPRESS - Science and Technology Publications*, Valletta, Malta, 2022, pp. 97–104.
80. Bogon, J, J Högerl, M Kocur, C Wolff, N Henze, and M Riemer. Validating Virtual Reality for Time Perception Research: Virtual Reality Changes Expectations about the Duration of Physical Processes, but Not the Sense of Time. *Behavior Research Methods*, Vol. 56, 2023, pp. 4553–4562. https://doi.org/10.3758/s13428-023-02201-6
81. Martins, R. C, F. F Reichert, R. M Bielemann, and P. C Hallal. One-Year Stability of Objectively Measured

Physical Activity in Young Brazilian Adults. *Journal of Physical Activity and Health*, Vol. 14, 2017, pp. 208–212. https://doi.org/10.1123/jpah.2015-0384
82. Ziemer, C. J, J. M Plumert, J. F Cremer, and J. K Kearney. Estimating Distance in Real and Virtual Environments: Does Order Make a Difference? *Attention, Perception, & Psychophysics*, Vol. 71, 2009, pp. 1095–1106. https://doi.org/10.3758/APP.71.5.1096
83. Drewes, J, S Feder, and W Einhäuser. Gaze During Locomotion in Virtual Reality and the Real World. *Frontiers in Neuroscience*, Vol. 15, 2021, P. 656913. https://doi.org/10.3389/fnins.2021.656913
84. Clay, V, P König, and S. U König. Eye Tracking in Virtual Reality. *Journal of Eye Movement Research*, Vol. 12, 2019, pp. 1–18. https://doi.org/10.16910/jemr.12.1.3
85. Van Der, Ham, I J. M, A. M. E Faber, M Venselaar, M. J Van Kreveld, and M Löffler. Ecological Validity of Virtual Environments to Assess Human Navigation Ability. *Frontiers in Psychology*, Vol. 6, 2015, P. 637. https://doi.org/10.3389/fpsyg.2015.00637
86. Allen, M. P The Problem of Multicollinearity. In *Understanding Regression Analysis*, Springer US, Boston, MA, 1997, pp. 176–180.
87. Penn State Eberly College of Science. 10.4 – Multicollinearity. https://online.stat.psu.edu/stat462/node/177/. Accessed May 19, 2023.
88. A.V Rowlands, C. L Edwardson, M. J Davies, K Khunti, D. M Harrington, and T Yates. Beyond Cut Points: Accelerometer Metrics That Capture the Physical Activity Profile. *Medicine & Science in Sports & Exercise*, Vol. 50, 2018, pp. 1323–1332. https://doi.org/10.1249/MSS.0000000000001561
89. Kelly, J. W, T. A Doty, S. B Gilbert, and M. C Dorneich. Field of View Restriction and Snap Turning as Cybersickness Mitigation Tools. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 31, No. 9, 2024, pp. 5704–5712. https://doi.org/10.1109/TVCG.2024.3470214
90. Farmani, Y, and R. J Teather. Evaluating Discrete Viewpoint Control to Reduce Cybersickness in Virtual Reality. *Virtual Reality*, Vol. 24, 2020, pp. 645–664. https://doi.org/10.1007/s10055-020-00425-x
91. Wilson, P. T, W Kalescky, A Mac, and Williams B Laughlin. VR Locomotion: Walking > Walking in Place > Arm Swinging. *Proc., 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1*, ACM, Zhuhai China, December 3, 2016, pp. 243–249.
92. Kuber, P. M, and E Rashedi. Alterations in Physical Demands During Virtual/Augmented Reality-Based Tasks: A Systematic Review. *Annals of Biomedical Engineering*, Vol. 51, 2023, pp. 1910–1932. https://doi.org/10.1007/s10439-023-03292-0