

Who Are Virtual Reality Headset Owners? A Survey and Comparison of Headset Owners and Non-Owners

Jonathan W. Kelly* Lucia A. Cherep† Alex F. Lim‡ Taylor Doty§ Stephen B. Gilbert¶

Iowa State University



Figure 1: Examples of accepted (left three) and rejected (right two) photos submitted by participants to verify HMD ownership.

ABSTRACT

The number of people who own a virtual reality (VR) head-mounted display (HMD) has reached a point where researchers can readily recruit HMD owners to participate remotely using their own equipment. However, HMD owners recruited online may differ from the university community members who typically participate in VR research. HMD owners ($n=220$) and non-owners ($n=282$) were recruited through two online work sites—Amazon’s Mechanical Turk and Prolific—and an undergraduate participant pool. Participants completed a survey in which they provided demographic information and completed measures of HMD use, video game use, spatial ability, and motion sickness susceptibility. In the context of the populations sampled, the results provide 1) a characterization of HMD owners, 2) a snapshot of the most commonly owned HMDs, 3) a comparison between HMD owners and non-owners, and 4) a comparison among online workers and undergraduates. Significant gender differences were found: men reported lower motion sickness susceptibility and more video game hours than women, and men outperformed women on spatial tasks. Men comprised a greater proportion of HMD owners than non-owners, but after accounting for this imbalance, HMD owners did not differ appreciably from non-owners. Comparing across recruitment platform, male undergraduates outperformed male online workers on spatial tests, and female undergraduates played fewer video game hours than female online workers. The data removal rate was higher from Amazon compared to Prolific, possibly reflecting greater dishonesty. These results provide a description of HMD users that can inform researchers recruiting remote participants through online work sites. These results also signal a need for caution when comparing in-person VR research that primarily enrolls undergraduates to online VR research that enrolls online workers.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality

*e-mail: jonkelly@iastate.edu

†e-mail:lacherep@iastate.edu

‡e-mail: aflim@iastate.edu

§e-mail:tdoty@iastate.edu

¶e-mail:gilbert@iastate.edu

1 INTRODUCTION

Until recently, the high cost of virtual reality (VR) equipment necessitated that VR research be conducted in laboratories with significant financial resources. The advent of cheaper, consumer-oriented head-mounted displays (HMDs) has reduced the financial requirements for lab-based research, and has also led to a growing number of home users who can participate in research remotely. Furthermore, work sites like Prolific and Amazon’s Mechanical Turk (AMT) are ideally suited for recruitment and payment of such participants. However, remote testing of participants who own HMDs warrants caution. Important differences may exist between the well-controlled lab environment and the uncontrolled environments of remote participants. Furthermore, in-person research typically recruits participants from the university or from the local community, most of whom do not own an HMD. Therefore, participants recruited for lab-based studies may differ in important ways from HMD owners recruited online.

HMD owners and non-owners were recruited through two online work sites—AMT and Prolific—as well as through a university research participation pool composed of undergraduate students enrolled in introductory psychology courses. Participants completed an online questionnaire in which they provided demographic information and completed measures of HMD use, video game use, spatial ability, and propensity for motion sickness. The results provide a description and comparison of HMD owners and non-owners recruited from two online work sites and an undergraduate participant pool.

2 RELATED WORK

Demographic data on HMD owners is sparse, and most of the existing research is available only through expensive market research reports. However, available evidence indicates that men are more likely than women to own an HMD, and that HMD owners are more likely to play video games than non-owners.

This section reviews available data on relevant demographics of HMD owners, as well as research indicating that the demographic differences between HMD owners and non-owners could mean that HMD owners have higher spatial cognitive ability compared to non-owners. Spatial ability is an important consideration for many VR researchers, given the spatial nature of virtual environments and the tasks commonly performed within them, including locomotion, navigation, and object manipulation.

2.1 Characteristics of HMD owners

There is little published academic research describing the characteristics of HMD owners and whether they differ from the general population. In a recent study that recruited an online sample of 24 HMD owners [38], 78% were men, 74% were white, and 65% had education beyond high school. Some insights can also be gleaned through market research summary reports. One such report [31] indicates that HMD owners are composed of 57% men and 42% women. Another report [7] indicates that the vast majority (90%) of HMD owners play games on their headsets at least once per week, indicating that HMD owners are likely to be gamers. However, the report lacks data from possible comparison groups, so it is unclear whether HMD owners play video games more than non-owners.

Based on these reports, it is expected that HMD owners, when compared to non-owners, are more likely to be men and more likely to be gamers. Women are under-represented in lab-based (i.e., in-person) VR research [33], and if women are less likely to own an HMD then they may continue to be under-represented in online VR research.

2.2 Sex differences in spatial ability

Given the possible over-representation of men among HMD owners (section 2.1), it is important to consider how this could impact the results of online VR research. One of the largest and most reliable sex differences in cognitive psychology is that men outperform women on tests of mental rotation and other small-scale spatial transformation tests [26,46], although some of these differences have been shown to be stimulus-specific [13,37]. Larger-scale navigation tasks show a smaller but still reliable male advantage [30].

Considerable research has explored the possible underpinnings of these sex differences in spatial cognition. Some research indicates an impact of circulating sex hormones on spatial task performance [1,16], but other large-scale studies have failed to find such an effect [36]. One popular account holds that sex differences in spatial cognition are related to differences in experience with spatial activities that promote spatial learning [25]. Indeed, spatial experiences are associated with enhanced spatial reasoning abilities [2], and training on spatial tasks increases spatial ability of men and women alike [44], reflecting the malleability of spatial skills.

2.3 Video games and spatial ability

HMD owners may be more likely to play video games than non-owners [7], so it is important to consider possible correlates of video game experience. Gamers have been shown to differ from non-gamers in several important ways, including personality [6], attention [22], spatial attention [12], spatial working memory [4,48], and task switching [41]. Given the spatial nature of interactions with VEs, the focus here is on spatial cognitive correlates of video game experience.

In one study [12], participants who self-identified as playing action video games outperformed non-gamers on a test of spatial attention known as the useful field-of-view task. The task requires participants to spread their attention broadly across the screen in order to identify a target that briefly appears in one of several possible locations. The authors argue that playing action video games trains gamers to spread their attention in this manner (e.g., to detect enemies), leading to superior performance on the spatial attention task.

Spatial attention is considered a key process underlying mental rotation. Individuals who excel at mental rotation tend to rotate the whole object at once, whereas poor performers tend to rotate the object one part at a time, in a piecemeal manner [17,21]. Whether participants use a holistic or piecemeal rotation strategy is thought to depend on the capacity of their spatial attention. Consistent with this notion, non-gamers who were trained on an action video game improved on tasks of spatial attention as well as mental rotation

[12]. In a related study, training participants on Tetris also led to improvements in mental rotation performance [42].

The literature on video games and cognitive functions is controversial. There are many reports of significant associations between video game experience and cognitive ability (see above). Yet, others have critiqued the methodologies of some of those studies while reporting no correlations among large participant samples [43]. Meta-analytic evidence is also mixed [3,39].

2.4 Research using online work sites

Online work sites such as AMT and Prolific provide researchers with convenient access to participants, and convenient payment systems for participant compensation. The large number of available workers often allows researchers to complete data collection quickly. The growing number of HMD owners means that AMT and Prolific may provide opportunities to recruit and compensate participants for VR research.

AMT and Prolific have been shown to reproduce findings from psychology that were originally established using in-person research [34], lending credibility to this method of participant recruitment. In some cases, AMT participants have been shown to outperform undergraduate participants compensated with course credit. For example, AMT participants passed more attention checks embedded within the task than did undergraduate student participants compensated with course credit [15]. In another study [50], participants were tasked with identifying and marking corn tassels within images of corn fields in order to create training data for a machine learning algorithm. AMT participants outperformed undergraduate students who were compensated with course credit, and AMT performance was close to that of a trained expert.

A common concern in paid online research is that participants will misrepresent themselves in order to gain access to studies with specific prerequisites [49]. For example, a participant might indicate that they own an HMD when they actually do not in order to gain access to the study. Such misrepresentation is less likely to be effective in VR research, since most VR studies would require the participant to actually own an HMD in order to install the VR application and complete the researcher's tasks. Because the current study used a questionnaire and did not require the use of an HMD, HMD owners were required to provide photographic evidence that they actually owned an HMD (see Figure 1). This method has been successfully used in past research to verify HMD ownership [27].

3 STUDY OVERVIEW

A survey was administered to HMD owners and non-owners for the purpose of comparing the two groups on dimensions relevant to VR research. The survey measured demographics, video game experience, spatial ability, and propensity for motion sickness. The survey was administered on AMT, Prolific, and through an undergraduate psychology participant pool.

Although the study was largely exploratory, the following hypotheses were pre-registered (see link provided in the Supplemental Material section at the end of the document). First, we hypothesized that the proportion of men would be larger among HMD owners compared to non-owners. Second, we hypothesized that HMD owners would report more video game hours than non-owners. Third, we hypothesized that HMD owners would perform better on tests of spatial ability than non-owners.

4 METHOD

4.1 Participants

A total of 688 participants completed the survey. Data collection occurred in summer and fall of 2020. Only US residents 18 years or older were eligible to participate. Undergraduate students were recruited through the psychology participant pool at Iowa State University. Following best practices for online data collection [35],

	AMT HMD owners	AMT non-owners	Prolific HMD owners	Prolific non-owners	Undergrad HMD owners	Undergrad non-owners
Submitted complete survey	205	112	112	97	40	122
Unverifiable HMD image	88	N/A	19	N/A	0	N/A
Did not follow task instructions	10	20	3	10	6	6
Failed attention check	1	0	1	0	2	4
Outlier data	4	4	3	2	0	3
Included in analyses	102	88	86	85	32	109
Percent retained	49.8%	78.6%	76.8%	87.6%	80.0%	89.3%

Figure 2: Visualization of data exclusion criteria and numbers of excluded participants.

AMT and Prolific participants were required to have an approval rating of 97% or greater and to have previously completed at least 100 jobs. Participants recruited through AMT and Prolific were paid \$3.50 (HMD owners) or \$2.50 (non-owners), and undergraduate participants received course credit.

Figure 2 shows the distribution of participants across HMD ownership category and recruitment platform, as well as reasons for data removal. Data from 186 participants were removed prior to data analysis, leaving a final sample of 502 participants. Data removal details are reported in section 5.

4.2 Materials

The full surveys distributed to HMD owners and non-owners are available as supplemental material. HMD owners were asked to verify HMD ownership and answer questions about HMD usage habits and experiences.

All participants answered self-report questions about motion sickness susceptibility, video game usage habits, and perceived sense of direction (SOD), followed by an attention check question. Next, participants completed two measures of spatial ability: a perspective-taking test (SOT: spatial orientation test [23]) followed by a mental rotation test (MRT [45]). Lastly, participants completed demographic questions.

4.2.1 HMD ownership verification and HMD experiences

HMD owners were asked to verify ownership by uploading a photo showing their HMD, and to include in the photo the last 4 digits of their user ID associated with the recruitment platform [27]. HMD owners then answered questions about their HMD usage, including preferred applications, usage frequency, and cybersickness experiences. Cybersickness experiences were assessed by having participants select the symptoms they experience from a list of 8 items, representing a subset of items from the Simulator Sickness Questionnaire [20].

4.2.2 Motion sickness

Motion sickness susceptibility was assessed using a single-item self-report rating on a 4-point scale [14]: "Do you regard yourself as susceptible to motion sickness?"

4.2.3 Video game habits

Video game habits were assessed by asking participants how many hours they play video games on a typical weekday and on a typical weekend day. Responses were used to calculate hours played per week. Participants were also asked to select from a list of which, if any, video game genres they typically play. If they selected "Immersive Virtual Reality" games from the list, then they were coded as having VR experience, whether or not they were HMD owners.

4.2.4 Sense of direction

Sense of direction was assessed using a single-item self-report rating on a 7-point scale [24]: "My sense of direction is very good."

4.2.5 Attention check

The attention check question provided four options and directly instructed participants to choose a specific option.

4.2.6 Perspective-taking test

Perspective-taking performance was assessed using the spatial orientation test (SOT) [23]. Each of the 12 test items displayed the same overhead layout of 7 objects. Test items instructed the participant to take a specific perspective and to point to another object from that perspective (e.g., "Imagine you are standing at the flower, facing the cat. Point to the car."). The instructed perspective varied across trials. Participants mapped their directional responses onto a clock marked with minutes, and then typed their response direction in minutes into a response box. Following one practice item with feedback, participants completed as many items as they could in 5 minutes. Absolute angular error was calculated for each completed trial and averaged across trials to obtain a single mean error score.

4.2.7 Mental rotation test

Mental rotation performance was assessed using a 3D object rotation test [45]. Each of the 20 trials presented a 3D "standard" object made from several cubes, along with four "test" objects also made from cubes. Two of the test objects were rotated versions of the standard object. Following one practice trial with feedback, participants completed as many trials as they could in 6 minutes. A trial was scored as correct only if both correct options were selected. Total score was calculated by summing the correct trials.

4.2.8 Demographics

Participants completed demographic questions about gender, age, race, ethnicity, education, household size, and US state of residence.

5 RESULTS

The data removal process is shown in Figure 2. Prior to analysis, participants in the HMD owners group were first screened for submitting a verifiable HMD image (see Figure 1 for examples of accepted and rejected HMD images). Next, data from all remaining participants were screened for adherence to task instructions on the SOT and MRT. Then, data were screened for passing the attention check. Finally, outliers were identified and removed. Univariate outliers were indicated by scores greater than 3 standard deviations from the mean, and multivariate outliers were identified using the Mahalanobis Distance test. The AMT HMD owners group stands

Table 1: HMD ownership totals by participant recruitment platform. The 10 most frequent HMDs are shown. The complete list is available as supplemental material.

VR Devices	AMT	Prolific	Undergrad	Total
PlayStation VR	19	19	11	49 (22.3%)
Samsung Gear VR	13	13	4	30 (13.6%)
Oculus Quest	11	14	3	28(12.7%)
Oculus Rift	8	11	1	20 (9.1%)
HTC Vive	10	5	2	17(7.7%)
Cardboard VR	3	9	4	16 (7.3%)
Oculus Go	12	2	1	15 (6.8%)
Oculus Rift S	5	6	3	14 (6.4%)
Valve Index	4	3	2	9 (4.1%)
Google Daydream	5	1	2	8 (3.6%)
Multiple*	7	7	5	19 (8.7%)

*13 owned 2 headsets, 3 owned 3 headsets, 2 owned 4 headsets, and 1 owned 7 headsets.

out as having the lowest percentage of data retained (49.8%), due to the large number of rejected HMD images. All other groups had comparable rates of data retention, averaging 82.5%.

Table 1 shows which HMDs were owned, sorted by ownership frequency. PlayStation VR stands out as the most frequently owned HMD across all three participant platforms, followed by a mixture of phone-based and traditional HMDs.

Most HMD owners (74.5%) reported experiencing at least one symptom of cybersickness. The frequency of symptom occurrence is shown in Table 2. See supplemental material for cybersickness data presented separately for the most frequently owned HMDs. Differences in symptom frequency across displays should be considered in the context of the primary applications used on those devices (e.g., some displays are primarily used for immersive gaming, which may lead to greater reports of nausea associated with those devices).

Participant demographics separated by HMD ownership and recruitment platform can be found in Tables 4 and 5. Online workers were older and included more men compared to undergraduate participants. Prolific participants included a smaller proportion of people who identified as "White or Caucasian" than did undergraduate or AMT participants.

There were many significant correlations among variables (see Table 3), most notably between gender and several dependent variables, as well as correlations among spatial ability measures. These associations are considered further below in the context of the MANOVA results. The gender variable was included in analyses of HMD ownership and recruitment platform, due to the correlation between gender and other key variables, and also due to the uneven distribution of gender across the HMD ownership groups.

Table 2: Frequency of reported cybersickness symptoms

Symptoms	Response
Eyestrain	95 (43.2%)
Dizziness	71 (32.3%)
Sweating	58 (26.5%)
Nausea	52 (23.6%)
Headaches	51 (23.2%)
Fatigue	30 (13.6%)
Vertigo	24 (10.9%)
Vomiting	1 (0.5%)
Other*	4 (1.8%)
Symptoms	164 (74.5%)
No symptoms	56 (25.5%)

*Head pain, Head strain (Oculus Quest is too heavy and too tight against the face; Only dizzy and nauseous for certain games; Stiffness in neck and shoulders from the weight of the headset).

5.1 Comparison of HMD owners and non-owners

A MANOVA was conducted to statistically evaluate the influence of HMD ownership and gender on the dependent variables, with follow-up comparisons as needed to further evaluate specific predictions and significant effects. Descriptive statistics for these groups are shown in Table 4. There were numerous gender differences: men outperformed women on the MRT and SOT, men reported higher SOD than women, men reported more video game hours than women, and men reported lower susceptibility to motion sickness than women. When accounting for the unbalanced gender distribution across HMD ownership categories, virtually no differences remained between HMD owners and non-owners, save for the finding that female HMD owners played more hours of video games than female non-owners. The following results support these conclusions.

Box's M value of 108.460 was associated with a p value $<.001$, which was interpreted as significant. Thus, the covariance matrices between the groups were not assumed to be equal and Pillai's Trace was used for the MANOVA. A MANOVA examined associations between the DVs (MRT, SOT, SOD, video game hours, and motion sickness susceptibility) and IVs (HMD ownership: owners vs. non-owners, and gender: men vs. women). The results showed a significant main effect of gender, $F(5, 494) = 18.182, p <.001$; Pillai's Trace = .155, $\eta_p^2 = .155$ (considered a large effect using the suggested cutoffs for η_p^2 of .01 for a small effect, .06 for a medium effect, and .14 for a large effect), and a significant main effect of HMD ownership, $F(5, 494) = 6.27, p <.001$; Pillai's Trace = .060, $\eta_p^2 = .060$. The interaction was not significant, $F(5, 494) = 1.669, p = .140$; Pillai's Trace = .017, $\eta_p^2 = .017$.

Univariate analyses showed that men ($M = 1.66, SD = .79$), compared to women ($M = 2.07, SD = .917$), reported significantly lower susceptibility to motion sickness, $F(1, 498) = 28.349, p <.001, \eta_p^2 = .054$. Welch's t -tests were conducted for DVs with unequal variances. Men ($M = 8.88, SD = 5.22$), compared to women ($M = 6.64, SD = 4.10$), performed significantly better on the MRT, $t(500) = 5.388, p <.001, d = .472$ (suggested cutoffs for Cohen's d are .2 for a small effect, .5 for a medium effect, and .8 for a large effect). Men ($M = 40.11, SD = 32.14$), compared to women ($M = 59.66, SD = 36.58$), also performed significantly better on the SOT, $t(488.025) = 6.373, p <.001, d = .582$. Men ($M = 5.47, SD = 1.315$), compared to women ($M = 4.63, SD = 1.684$), reported significantly higher sense-of-direction, $t(407.896) = 6.078, p <.001, d = .565$, and men ($M = 27.53, SD = 16.01$), compared to women ($M = 23.15, SD = 16.42$), also reported significantly more hours of video game play, $t(466.844) = 2.998, p = .003, d = .271$.

Univariate analyses showed that HMD owners ($M = 1.83, SD = .90$) and non-owners ($M = 1.85, SD = .85$) did not significantly differ on motion sickness susceptibility, $F(1, 498) = .947, p = .331, \eta_p^2 = .002$. Levene's test indicated unequal variances for the MRT, $F(3, 498) = 7.163, p <.001$, the SOT, $F(3, 498) = 4.482, p = .004$, SOD, $F(3, 498) = 4.682, p <.001$, and hours of video game play, $F(3, 498) = 4.682, p = .003$. Therefore, a Welch's t -test was reported for these DVs. HMD owners ($M = 8.21, SD = 4.93$), compared to non-owners ($M = 7.65, SD = 4.84$), did not significantly differ on the MRT, $t(466.465) = 1.281, p = .201, d = .115$. HMD owners ($M = 44.93, SD = 33.43$), compared to non-owners ($M = 51.67, SD = 36.02$) performed significantly better on the SOT, $t(485.154) = 2.168, p = .031, d = .193$, and HMD owners ($M = 5.36, SD = 1.40$), compared to non-owners ($M = 4.89, SD = 1.62$) reported significantly higher SOD, $t(494.422) = 3.508, p <.001, d = .308$. Finally, HMD owners ($M = 29.66, SD = 17.21$), compared to non-owners ($M = 22.43, SD = 14.87$) reported significantly more hours of weekly video game play, $t(433.758) = 4.957, p <.001, d = .455$.

Caution is warranted with interpreting the differences reported above between HMD owners and non-owners, which could be driven by gender differences due to the unequal distribution of men and women among HMD owners and non-owners. Specifically, there

Table 3: Correlations between study variables. * $p < .05$; ** $p < .01$. Gender was coded 0 = men and 1 = women; HMD Ownership was coded 0 = non-owner and 1 = owner; VR Experience was coded as 0 = no experience using VR and 1 = experience using VR. MRT = Mental Rotation Test, SOT = Spatial Orientation Test, SOD = Sense of Direction.

	M	SD	1.	2.	3.	4.	5.	6.	7.	8.
1. Gender	—	—								
2. HMD ownership	—	—	-0.25**							
3. VR experience	0.29	0.46	-0.19**	0.66**						
4. Age	29.84	11.10	-0.06	0.10*	0.01					
5. MRT	7.89	4.88	-0.23**	0.06	0.08	-0.14**				
6. SOT	48.72	35.04	0.28**	-0.10*	-0.13**	0.13**	-0.54**			
7. SOD	5.10	1.54	-0.27**	0.15**	0.10*	0.15**	0.09*	-0.10*		
8. Game hrs/wk	25.60	16.32	-0.13**	0.22**	0.19**	-0.05	0.01	0.06	0.07	
9. Motion sickness	1.84	0.87	0.23**	-0.02	0.07	-0.08	-0.07	0.07	-0.22**	-0.05

is a smaller proportion of women among HMD owners compared to non-owners (33% vs 55%). Given the significant gender differences already reported, it is possible that the unequal distribution of men and women among the HMD ownership groups underlies the significant effects of HMD ownership on SOT performance, self-reported sense of direction, and video game hours. Therefore, within-gender comparisons were conducted between the two HMD ownership groups on these three measures. HMD owners did not differ from non-owners on SOT or SOD. The Welch's t-test for video game hours showed that female HMD owners ($M = 31.08$, $SD = 20.00$) reported significantly more hours compared to female non-owners ($M = 19.77$, $SD = 13.34$), $t(90.618) = 4.210$, $p < .001$, $d = .622$. Men did not differ on this measure.

Further analyses were conducted to compare owners of gaming-oriented HMDs (e.g., those that can connect to Steam) to non-gaming HMDs (e.g., smartphone-based HMDs). The results indicate no appreciable differences between gaming and non-gaming HMD owners. Complete results are available as supplemental material.

5.2 Comparison across recruitment platform

A MANOVA was conducted to statistically evaluate the influence of recruitment platform (AMT, Prolific, and undergraduate) and gender on the dependent variables, with paired comparisons to evaluate specific predictions and significant effects. Descriptive statistics for these groups are shown in Table 5. In addition to the gender differences described in section 5.1, the primary results of this analysis are that undergraduate men outperformed online workers on spatial tasks, although AMT participants self-reported higher SOD than Prolific or undergraduate participants. Furthermore, undergraduate women reported fewer video game hours than AMT women and Prolific women. The following results support these conclusions.

Box's M value of 112.897 was associated with a p value = .005, which was interpreted as significant. Therefore, Pillai's Trace was used for the MANOVA. A MANOVA examined associations between the DVs (MRT, SOT, SOD, video game experience, and mo-

tion sickness susceptibility) and IVs (platform: AMT vs. Prolific vs. undergraduate, and gender: men vs. women). The results showed a significant main effect of platform, $F(10, 986) = 4.240$, $p < .001$; Pillai's Trace = .082, $\eta_p^2 = .041$, and a significant main effect of gender, $F(5, 492) = 23.282$, $p < .001$; Pillai's Trace = .191, $\eta_p^2 = .191$. The main effects were qualified by a significant interaction between platform and gender, $F(10, 986) = .050$, $p = .005$; Pillai's Trace = .050, $\eta_p^2 = .025$. The focus here is on the interactions between recruitment platform and gender because the main effects of gender were described thoroughly in section 5.1, and because the main effects can only be interpreted in the context of the interaction.

Levene's test indicated unequal variances for the MRT, $F(5, 496) = 2.681$, $p = .021$, the SOT, $F(5, 496) = 4.381$, $p = .001$, and SOD, $F(5, 496) = 5.234$, $p < .001$. Therefore, the Games-Howell post-hoc correction was applied to account for unequal variances and sample sizes.

MRT performance showed that undergraduate men outperformed AMT men and Prolific men, $F(2, 278) = 10.925$, $p < .001$, $\eta_p^2 = .073$, who did not significantly differ from one another ($p = .885$). In contrast, MRT performance for women did not differ across recruitment platform, $F(2, 218) = 1.178$, $p = .310$, $\eta_p^2 = .011$.

SOT performance showed that undergraduate men outperformed AMT men, $F(2, 278) = 5.311$, $p = .005$, $\eta_p^2 = .037$. Prolific men ($M = 40.62$, $SD = 31.99$) did not significantly differ from undergraduate men ($p = .071$) or AMT men ($p = .934$) on the SOT. In contrast, SOT performance for women did not differ across recruitment platform, $F(2, 218) = 1.672$, $p = .190$, $\eta_p^2 = .015$.

Self-reported SOD was significantly higher for AMT men ($M = 5.73$, $SD = 1.25$) compared to Prolific men ($M = 5.20$, $SD = 1.30$), $F(2, 278) = 5.066$, $p = .007$, $\eta_p^2 = .035$. Undergraduate men ($M = 5.47$, $SD = 1.32$) were not significantly different from AMT men ($p = .143$) or Prolific men ($p = .833$) on SOD. AMT women ($M = 5.11$, $SD = 1.74$) reported a significantly higher SOD compared to undergraduate women ($M = 4.36$, $SD = 1.68$), $F(2, 218) = 3.888$, $p = .022$, $\eta_p^2 = .034$. Prolific women ($M = 4.52$, $SD = 1.71$) were not significantly different from AMT women ($p = .116$) or undergraduate women ($p = .816$) on SOD.

Video game hours reported by men did not differ across recruitment platform, $F(2, 278) = 1.033$, $p = .357$, $\eta_p^2 = .007$. In contrast, undergraduate women ($M = 17.95$, $SD = 13.63$) reported significantly fewer video game hours compared to Prolific women ($M = 26.47$, $SD = 18.40$) and AMT women ($M = 26.05$, $SD = 15.79$), $F(2, 218) = 7.035$, $p = .001$, $\eta_p^2 = .061$, who did not significantly differ from one another ($p = 1.00$).

Motion sickness susceptibility ratings by men did not significantly differ across platform, $F(2, 278) = 1.769$, $p = .172$, $\eta_p^2 = .013$. Nor did sickness ratings by women significantly differ across platform, $F(2, 218) = .374$, $p = .688$, $\eta_p^2 = .003$.

Table 4: Demographics and other dependent variables by HMD ownership group and gender

	HMD owners		Non-owners	
	Female (n=66)	Male (n=154)	Female (n=155)	Male (n=127)
Age*	32.45 (11.48)	30.49 (9.23)	27.58 (12.18)	30.45 (11.24)
White†	53 (80.30%)	121 (78.57%)	123 (79.35%)	85 (66.93%)
Game hrs/wk*	31.08 (20.00)	29.06 (15.90)	19.77 (13.34)	25.67 (16.01)
Sickness*	2.12 (0.97)	1.70 (0.83)	2.05 (0.87)	1.61 (0.74)
SOD*	4.92 (1.77)	5.55 (1.17)	4.50 (1.64)	5.36 (1.47)
MRT*	6.53 (4.20)	8.93 (5.06)	6.68 (4.07)	8.82 (5.42)
SOT*	61.31 (34.02)	37.91 (0.71)	58.95 (36.31)	42.79 (33.73)

*Mean(SD); †N (%)

Table 5: Demographics and other dependent variables by participant recruitment platform and gender

	AMT		Prolific		Undergrad		All platforms	
	Female (n=63)	Male (n=127)	Female (n=75)	Male (n=96)	Female (n=83)	Male (n=58)	Female (n=221)	Male (n=281)
Age*	38.41 (11.25)	34.63 (9.17)	32.08 (11.69)	31.80 (9.45)	19.17 (1.70)	19.16 (1.23)	29.04 (12.15)	30.47 (10.17)
White†	54 (85.71%)	96 (75.59%)	50 (66.67%)	63 (65.63%)	72 (86.75%)	47 (81.03%)	176 (79.63%)	206 (73.31%)
Game hrs/wk*	26.05 (15.79)	26.02 (14.81)	26.47 (18.40)	28.74 (17.56)	17.95 (13.63)	28.83 (15.82)	23.45 (16.42)	27.53 (16.01)
Sickness*	2.00 (0.95)	1.58 (0.78)	2.07 (0.88)	1.78 (0.80)	2.13 (0.93)	1.64 (0.79)	2.07 (0.92)	1.66 (0.79)
SOD*	5.11 (1.74)	5.73 (1.25)	4.52 (1.71)	5.20 (1.30)	4.36 (1.55)	5.33 (1.38)	4.63 (1.68)	5.47 (1.32)
MRT*	6.00 (4.31)	8.31 (5.10)	7.05 (4.08)	7.98 (5.18)	6.75 (3.95)	11.62 (4.65)	6.64 (4.10)	8.88 (5.22)
SOT*	66.20 (35.04)	44.97 (33.40)	58.78 (36.36)	40.62 (31.99)	55.48 (34.98)	28.64 (26.79)	59.66 (35.58)	40.11 (32.14)

*Mean(SD); †N (%)

6 DISCUSSION

Research using VR has traditionally been conducted in laboratories using lab-owned equipment. Participants in these studies have traditionally been recruited from the university community, including university undergraduates. The growing number of HMD owners makes it possible to conduct VR research remotely, by recruiting HMD owners to participate from home using their own equipment [38, 40]. Online work platforms like AMT and Prolific further simplify recruitment and payment of HMD owners by providing access to a large number of potential participants. This project recruited participants from AMT, Prolific, and an undergraduate research pool in order to evaluate and compare characteristics of HMD owners and compare them to non-owners.

The results show that there is a larger proportion of men among HMD owners compared to non-owners, as hypothesized. Women are already under-represented in lab-based VR research [33], so researchers conducting online research with HMD owners will need to make efforts to ensure appropriate representation of women in their studies.

Within-gender comparisons showed that HMD owners and non-owners did not perform differently on spatial tasks, failing to support the hypothesis. Owners and non-owners also did not differ on motion sickness susceptibility. The only significant difference was that female HMD owners reported playing more video game hours than did female non-owners, which partially supports the hypothesis that HMD owners would play more video game hours compared to non-owners. But even this difference is mitigated by the lack of significant correlation between video game hours and spatial ability measures in the current data set (but see section 2.3 for a summary of evidence for such an association). Given the fundamentally spatial nature of virtual environments and the spatial demands of locomotion interfaces for VR [5, 9, 28, 29, 32], the similarity between HMD owners and non-owners on spatial measures is good news for researchers who hope that their research involving participants from one population (e.g., HMD non-owners) will generalize to the other population (e.g., HMD owners).

The results also show that workers recruited through online work sites are older, include more men, and are more racially diverse than undergraduate students recruited from Iowa State University. Undergraduate women played fewer hours of video games compared to women from online work sites. Furthermore, undergraduate men performed better on spatial tasks compared to men from online work sites. However, an important caveat is that the undergraduate participants were enrolled at a university with an engineering emphasis. Although student major was not recorded, it is likely that many of the students came from STEM majors, and spatial ability is positively correlated with success in STEM generally [47] and engineering in particular [18]. The gender imbalance in some STEM majors could account for why only male undergraduates outperformed their online worker counterparts. On the other hand, online workers may be more

computer proficient than the general population, and their spatial ability may also be elevated as a result [11, 19]. Further research is needed to determine the generality of the superior performance by male undergraduates compared to online workers.

The superior spatial ability of undergraduate men compared to men from online work sites is concerning given the spatial nature of virtual environments and locomotion interfaces. For example, navigation performance when using the ubiquitous teleporting interface is associated with performance on the MRT and SOT [8, 10]. Researchers collecting data from online samples should therefore be cautious when comparing their results to results from in-person studies, especially when spatial tasks are involved.

The data removal rate was higher among AMT participants compared to Prolific participants and undergraduate participants. In this study, AMT participants were removed primarily due to rejected HMD images. Rejected images were most often stock HMD images, which participants presumably downloaded and submitted in hopes of qualifying for a study that they were actually ineligible for. This may not be a major issue when participants are required to use their HMD to complete the study, because lack of an HMD would prevent cheaters from even attempting to participate. However, this does suggest that dishonesty is higher among AMT participants than other participation platforms.

The results provide a snapshot of the most frequently owned HMDs, which could help researchers decide which equipment to plan for when developing their studies. The results also show that cybersickness symptoms among HMD owners are commonplace, and that experienced symptoms vary across displays. Further research in which participants rate the magnitude of experienced symptom could provide a more detailed comparison across different displays and activities, and could be helpful for further elucidating the causes of cybersickness.

These results must be considered within the context of the populations that were sampled in this study, as they may not generalize to the broader populations of HMD owners and non-owners. The current study recruited online workers and undergraduate students. Recruitment of HMD owners through other means may produce samples with different characteristics than those recruited here. Therefore, future work should evaluate HMD owners recruited through other popular methods, such as social media.

This study was conducted relatively early in the technology adoption life cycle for HMDs. Consumer-oriented HMDs became available in 2016, and despite their recent rise in popularity they are owned primarily by innovators and early adopters. Characteristics of HMD owners may shift as the HMD adoption life cycle continues. Therefore, it will be worth reassessing the current findings after HMD ownership becomes more widespread.

These results provide a description of HMD users that can inform recruitment of remote participants who own VR equipment. When recruiting HMD owners, researchers should be prepared to make

efforts to recruit a sufficient number of women. Few differences exist between HMD owners and non-owners, but caution is warranted when recruiting from online work sites such as AMT and Prolific, where spatial ability may be lower than in the undergraduate population, particularly among men. Such differences in spatial ability could have wide-ranging effects on task performance in virtual environments, given their fundamentally spatial nature. Future work comparing undergraduates with online workers, especially on spatial tasks such as locomotion, is essential to understanding the potential consequences of conducting VR research using online participants.

SUPPLEMENTAL MATERIAL

Pre-registration, complete survey questions, data, and supplemental analyses are available on the Open Science Framework: <https://osf.io/843ec/>.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. CHS-1816029.

REFERENCES

- [1] A. Aleman, E. Bronk, R. P. C. Kessels, H. P. F. Koppeschaar, and J. Honk. A single administration of testosterone improves visuospatial ability in young women. *Psychoneuroendocrinology*, 29(5):612–617, 2004. doi: 10.1016/S0306-4530(03)00089-1
- [2] M. Baenninger and N. Newcombe. The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, 20:327–344, 1989. doi: 10.1007/BF00287729
- [3] B. Bediou, D. M. Adams, R. E. Mayer, E. Tipton, C. S. Green, and D. Bavelier. Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin*, 144(1):77–110, 2018. doi: 10.1037/bul0000130
- [4] K. J. Blacker and K. M. Curby. Enhanced visual short-term memory in action video game players. *Attention, Perception, and Psychophysics*, 75:1128–1136, 2013. doi: 10.3758/s13414-013-0487-0
- [5] D. A. Bowman, D. Koller, and L. F. Hodges. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Proc. Virtual Reality Annual International Symposium*, pp. 45–52. IEEE Computer Society, Washington, D.C., 1997. doi: 10.1109/VRAIS.1997.583043
- [6] B. Braun, J. M. Stopfer, K. W. Muller, M. E. Beutel, and B. Egloff. Personality and video gaming: Comparing regular gamers, non-gamers, and gaming addicts and differentiating between game genres. *Computers in Human Behavior*, 55A:406–412, 2016. doi: 10.1016/j.chb.2015.09.041
- [7] CCS. Ccs insight: Virtual and augmented reality headset shipments ready to soar, 2019.
- [8] L. Cherep, J. Kelly, A. J. Miller, A. Lim, and S. B. Gilbert. Individual differences in teleporting through virtual environments. *PsyArXiv*, Dec 2020. doi: 10.31234/osf.io/b6cyd
- [9] L. A. Cherep, A. F. Lim, J. W. Kelly, D. Acharya, A. Velasco, E. Bustamante, A. G. Ostrander, and S. B. Gilbert. Spatial cognitive implications of teleporting through virtual environments. *Journal of Experimental Psychology: Applied*, 26(3):480–492, 2020. doi: 10.1037/xap0000263
- [10] L. A. Cherep, A. F. Lim, J. W. Kelly, A. Miller, and S. B. Gilbert. Individual differences in teleporting through virtual environments: A latent profile analysis. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 724–725, 2020. doi: 10.1109/VRW50115.2020.00213
- [11] R. De Lisi and D. M. Cammarano. Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior*, 12(3):351–361, 1996. doi: 10.1016/0747-5632(96)00013-1
- [12] J. Feng, I. Spence, and J. Pratt. Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, 18(10):850–855, 2007. doi: 10.1111/j.1467-9280.2007.01990.x
- [13] M. L. Fisher, T. Meredith, and M. Gray. Sex differences in mental rotation ability are a consequence of procedure and artificiality of stimuli. *Evolutionary Psychological Science*, 4:124–133, 2018. doi: 10.1007/s40806-017-0120-x
- [14] J. F. Golding. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Research Bulletin*, 47(5):507–516, 1998. doi: 10.1016/S0361-9230(98)00091-4
- [15] D. J. Hauser and N. Schwarz. Attentive turkers: Mturk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, 48:400–407, 2016. doi: 10.3758/s13428-015-0578-z
- [16] M. Hausmann, D. Slabbekoorn, S. H. M. V. Goozen, P. T. Cohen-Kettenis, and O. Güntürkün. Sex hormones affect spatial abilities during the menstrual cycle. *Behavioral Neuroscience*, 114(6):1245–1250, 2000. doi: 10.1037/0735-7044.114.6.1245
- [17] M. Heil and P. Jansen-Osmann. Sex differences in mental rotation with polygons of different complexity: Do men utilize holistic processes whereas women prefer piecemeal ones? *Quarterly Journal of Experimental Psychology*, 61(5):683–689, 2008. doi: 10.1080/17470210701822967
- [18] S. Hsi, M. C. Linn, and J. E. Bell. The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education*, 86(2):151–158, 1997. doi: 10.1002/j.2168-9830.1997.tb00278.x
- [19] S. Jones and G. Burnett. Spatial ability and learning to program. *Human Technology*, 4(1):47–61, 2008.
- [20] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993. doi: 10.1207/s15327108ijap0303_3
- [21] P. Khooshabeh, M. Hegarty, and T. F. Shipley. Individual differences in mental rotation: Piecemeal versus holistic processing. *Experimental Psychology*, 60(3):164–171, 2013. doi: 10.1027/1618-3169/a000184
- [22] M. Kowal, A. J. Toth, C. Exton, and M. J. Campbell. Different cognitive abilities displayed by action video gamers and non-gamers. *Computers in Human Behavior*, 88:255–262, 2018. doi: 10.1016/j.chb.2018.07.010
- [23] M. Kozhevnikov and M. Hegarty. A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29(5):745–756, 2001. doi: 10.3758/BF03200477
- [24] L. T. Kozlowski and K. J. Bryant. Sense of direction, spatial orientation, and cognitive maps. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4):590–598, 1977. doi: 10.1037/0096-1523.3.4.590
- [25] J. E. Lauer, E. Yhang, and S. F. Lourenco. The development of gender differences in spatial reasoning: A meta-analytic review. *Psychological Bulletin*, 145(6):537–565, 2019. doi: 10.1037/bul0000191
- [26] M. C. Linn and A. C. Petersen. Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, 56:1479–1498, 1985. doi: 10.2307/1130467
- [27] X. Ma, M. Cackett, L. Park, E. Chien, and M. Naaman. Web-based vr experiments powered by the crowd. In *Proceedings of the 2018 World Wide Web Conference, WWW '18*, p. 33–43. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 2018. doi: 10.1145/3178876.3186034
- [28] W. E. Marsh, J. W. Kelly, V. J. Dark, and J. H. Oliver. Cognitive demands of semi-natural virtual locomotion. *Presence: Teleoperators and Virtual Environments*, 22(3):216–234, 2013. doi: 10.1162/PRES_a.00152
- [29] K. R. Moghadam, C. Banigan, and E. D. Ragan. Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness. *IEEE Transactions on Visualization and Computer Graphics*, 2018. doi: 10.1109/TVCG.2018.2884468
- [30] A. Nazareth, X. Huang, D. Voyer, and N. Newcombe. A meta-analysis of sex differences in human navigation skills. *Psychonomic Bulletin and Review*, 26:1503–1528, 2019. doi: 10.3758/s13423-019-01633-6
- [31] Nielsen. Nielsen games 360 u.s. report, 2018.
- [32] R. Paris, J. Klag, P. Rajan, L. E. Buck, T. P. McNamara, and B. Bodenheimer. How video game locomotion methods affect navigation in virtual environments. In *Proc. Symposium on Applied Perception*, pp.

- 12.1–12.7. ACM, New York City, NY, 2019. doi: 10.1145/3343036.3343131
- [33] T. C. Peck, L. E. Sockol, and S. M. Hancock. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1945–1954, 2020.
- [34] E. Peer, L. Brandimarte, S. Samat, and A. Acquisti. Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70:153–163, 2017. doi: 10.1016/j.jesp.2017.01.006
- [35] E. Peer, J. Vosgerau, and A. Acquisti. Reputation as a sufficient condition for data quality on amazon mechanical turk. *Behavior Research Methods*, 46(4):1023–1031, 2014. doi: 10.3758/s13428-013-0434-y
- [36] D. A. Puts, R. A. Cárdenas, D. H. Bailey, R. P. Burriss, C. L. Jordan, and S. M. Breedlove. Salivary testosterone does not predict mental rotation performance in men or women. *Hormones and Behavior*, 58(2):282 – 289, 2010. doi: 10.1016/j.yhbeh.2010.03.005
- [37] M. Rahe, V. Ruthsatz, and C. Quaiser-Pohl. Influence of the stimulus material on gender differences in a mental-rotation test. *Psychological Research*, 2020. doi: 10.1007/s00426-020-01450-w
- [38] D. Saffo, S. Di Bartolomeo, C. Yildirim, and C. Dunne. Remote and collaborative virtual reality experiments via social vr platforms. *Open Science Framework*, Jan 2021. doi: osf.io/c2amz
- [39] G. Sala, K. S. Tatlidil, and F. Gobet. Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin*, 144(2):111–139, 2018. doi: 10.1037/bul0000139
- [40] A. Steed, S. Frlston, M. M. Lopez, J. Drummond, Y. Pan, and D. Swapp. An ‘in the wild’ experiment on presence and embodiment using consumer virtual reality equipment. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1406–1414, 2016. doi: 10.1109/TVCG.2016.2518135
- [41] T. Strobach, P. A. Frensch, and T. Schubert. Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychologica*, 140:13–24, 2012. doi: 10.1016/j.actpsy.2012.02.001
- [42] M. S. Terlecki, N. S. Newcombe, and M. Little. Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Applied Cognitive Psychology*, 22:996–1013, 2008. doi: 10.1002/acp.1420
- [43] N. Unsworth, T. S. Redick, B. D. McMillan, D. Z. Hambrick, M. J. Kane, and R. W. Engle. Is playing video games related to cognitive abilities? *Psychological Science*, 26(6):759–774, 2015. doi: 10.1177/0956797615570367
- [44] D. H. Uttal, N. G. Meadow, E. Tipton, L. L. Hand, A. R. Alden, C. Warren, and N. S. Newcombe. The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2):352–402, 2013. doi: 10.1037/a0028446
- [45] S. G. Vandenberg and A. R. Kuse. Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, 47(2):599–604, 1978. doi: 10.2466/pms.1978.47.2.599
- [46] D. Voyer, S. Voyer, and M. P. Bryden. Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117:250–270, 1995. doi: 10.1037/0033-2909.117.2.250
- [47] J. Wai, D. Lubinski, and C. P. Benbow. Spatial ability for stem domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4):817–835, 2009. doi: 10.1037/a0016127
- [48] O. Waris, S. M. Jaeggi, A. R. Seitz, M. Lehtonen, A. Soveri, K. M. Lukasik, U. Söderström, R. A. C. Hoffing, and M. Laine. Video gaming and working memory: A large-scale cross-sectional correlative study. *Computers in Human Behavior*, 97:94 – 103, 2019. doi: 10.1016/j.chb.2019.03.005
- [49] K. S. Wessling, J. Huber, and O. Netzer. Mturk character misrepresentation: Assessment and solutions. *Journal of Consumer Research*, 44(1):211–230, 2017. doi: 10.1093/jcr/ucx053
- [50] N. Zhou, Z. D. Siegel, S. Zarecor, N. Lee, D. A. Campbell, C. M. Andorf, D. Nettleton, C. J. Lawrence-Dill, B. Ganapathysubramanian, J. W. Kelly, and I. Friedberg. Crowdsourcing image analysis for plant phenomics to generate ground truth data for machine learning. *PLoS Computational Biology*, 14(7):e1006337, 2018. doi: 10.1371/journal.pcbi.1006337