

1 Article

2 The Validity of a Mixed Reality-based Automated 3 Functional Mobility Assessment

4 Ruopeng Sun ^{1, *}, Roberto G. Aldunate ¹ and Jacob J Sosnoff ¹

5 ¹ Department of Kinesiology and Community Health, University of Illinois at Urbana-Champaign;
6 rusun@illinois.edu, aldunate@illinois.edu, jsosnoff@illinois.edu;

7 * Correspondence: rusun@illinois.edu;

8

9 **Abstract:** Functional mobility assessments (i.e. Timed Up and Go) are commonly used clinical tools
10 for mobility and fall risk screening in the aging population. In this work, we proposed a new Mixed
11 Reality (MR)-based assessment that utilized a Microsoft HoloLens™ headset to automatically lead
12 and track the performance of functional mobility tests, and subsequently evaluated its validity in
13 comparison with reference inertial sensors. Twenty-two healthy adults (10 older, 12 young)
14 participated in this study. An automated functional mobility assessment app was developed based
15 on the HoloLens platform. Mobility performance was recorded with the headset built-in sensor and
16 validated with reference inertial sensor (Opal, APDM) taped on the headset and lower back. Results
17 indicate vertical kinematic measures by HoloLens was in good agreement with the reference sensor
18 (Normalized RMSE ~ 10%). Additionally, the HoloLens-based test completion time was in perfect
19 agreement with clinical standard stopwatch measure. Overall, our preliminary investigation
20 indicates that it is possible to use an MR headset to automatically guide users to complete common
21 mobility tests with good measurement accuracy, thus it has great potential to provide objective and
22 efficient sensor-based mobility assessment.

23 **Keywords:** Mixed Reality Headset; Mobility Assessment; Wearable Sensor; Fall Risk; Aging

24

25 1. Introduction

26 Falls are the leading cause of injury related death in older adults [1]. Over 1 in 4 older adults will
27 experience a fall in the next year and a significant portion of those that fall will suffer an injury [2],
28 resulting in more than \$50 billion annual medical costs [3]. Moreover, although most falls do not end
29 in death or result in significant physical injury, the psychological impact of a fall, such as fear of
30 falling, loss of confidence, often results in anxiety and depression that further decreased quality of
31 life [4]. The risk of falling increase with aging due to multiple risk factors, such as deficits in vision,
32 cognition, muscle strength and mobility [1]. Given the frequency and severe consequences of falls,
33 there is a critical need for early and regular monitoring on individual's fall risk to reduce falls and
34 fall-related injuries.

35 Indeed, the American Geriatrics Society (AGS) and the Centers for Disease Control and
36 Prevention (CDC) recommends annual fall risk screening for older adults [5, 6]. The most commonly
37 used fall risk screening tests include functional mobility tests such as the five time sit to stand test
38 (STS) [7] and timed up and go test (TUG) [8]. Both tests are valid and reliable clinical tests focusing
39 on assessing lower-limb muscle strength and mobility, which are both critical factors to the overall
40 risk of falling. The STS is a clinical test that asks participants to stand up and sit down from a chair
41 five times as quickly as possible without using the armrest. Whereas the TUG test requires participant
42 stand up from a chair, walks 3m at normal pace, turn around and return to the chair.

43 Even though such standardized tests are relatively easy to conduct, it is still underutilized and
44 not routinely integrated into clinical practice. Partially due to clinicians' time constraints and
45 competing medical priorities, lack of access to lab-grade advanced testing equipment (such as motion

46 capture device, force platform), as well as lack of clinical expertise [9], thus limiting access to fall risk
47 screening in the community dwelling older adults. Consequently, older adults remain unaware of
48 their individual fall risk, appropriate fall prevention approaches and at elevated risk of falls.

49 With the recent advancement in sensing technology, sensor-based fall risk assessment that can
50 efficiently capture and analyze quantitative mobility data have received a growing interest for its
51 portability, accessibility and inexpensiveness [9, 10]. More specifically, the use of wearable sensors
52 (inertial measurement unit, IMU) for mobility-related tracking has been the focus of these work, in
53 which miniature accelerometers and/or gyroscopes were used to quantify movement
54 pattern/abnormality by various time and frequency parameters [9, 11].

55 Two recent systematic reviews on sensor-based fall risk assessment [9, 12] identified over 50
56 investigations using IMUs for fall risk assessment in older adults. This body of literature highlights
57 that sensors are capable of accurately quantifying mobility and capable of distinctive in identifying
58 mobility impairment in high fall risk individuals. Overall it is concluded that wearable sensors are a
59 viable technology for fall risk assessment. It is worth noting that in most investigations, IMU sensor(s)
60 was attached to lower back and/or lower limb as stand-alone recording device, and still required
61 additional personnel to guide the wearer through the assessment protocols.

62 An alternative approach to further improve the efficiency of sensor-based assessment is to offer
63 direct technology interaction with the intended users, such as deliver demonstrations and
64 instructions, and receive user inputs. Such a system would be able to provide an automated and self-
65 guided assessment system that requires no additional personnel. To achieve this goal, an ideal device
66 should be able to communicate with the wearer with both visual and auditory prompt, and allow
67 user input through natural interactions (gesture, voice, gaze, etc.). Indeed, recent research has
68 highlighted that smartphone and tablets can provide valid and reliable fall risk assessment to older
69 adults [13, 14].

70 Mixed reality head-mounted display (HMD, e.g. Microsoft HoloLens, Fig.1a) systems are also
71 uniquely fitted for technology-based fall risk assessment. For instance, the HoloLens uses a
72 transparent display with light projector to provide holograms on the lenses in front of user's eye that
73 blend the digital display with physical environments [15]. It contains multiple sensors to scan the
74 user's environment which enables the holograms to be placed at a specific location in real world [15].
75 By using such device in mobility assessment, the user can receive instructions and visual
76 demonstration, naturally interact with the virtual display through voice command, gesture control
77 and gaze, and complete mobility tests with full visibility of the surrounding environments. In
78 addition, the embedded IMU and depth sensor can be used to track the user's head movement during
79 balance and mobility tests. Although less commonly used than sensors mounted at lower limb and
80 lower back for mobility assessment, head movement has been used as approach for mobility
81 evaluation, based on the notion that head movement is linked to the trunk movement as well as gait-
82 related oscillations during locomotion [16, 17]. Additionally, given that head stabilization has been
83 shown as a critical component in maintaining upright posture [16], monitoring head movement may
84 provide novel insights to the mobility control and fall risk evaluation. These unique features of HMD
85 have the potential to enable older adults to complete fall risk screening intuitively and autonomously.

86 Although this Mixed Reality Headset holds promise for enhancing the fall risk assessment in
87 community dwelling older adult, its validity for objective mobility assessment have not been
88 investigated. Therefore, the aim of this study is to evaluate the validity of the mixed-reality headset
89 for automated mobility assessment in young and older adults.

90 2. Materials and Methods

91 2.1 Participants

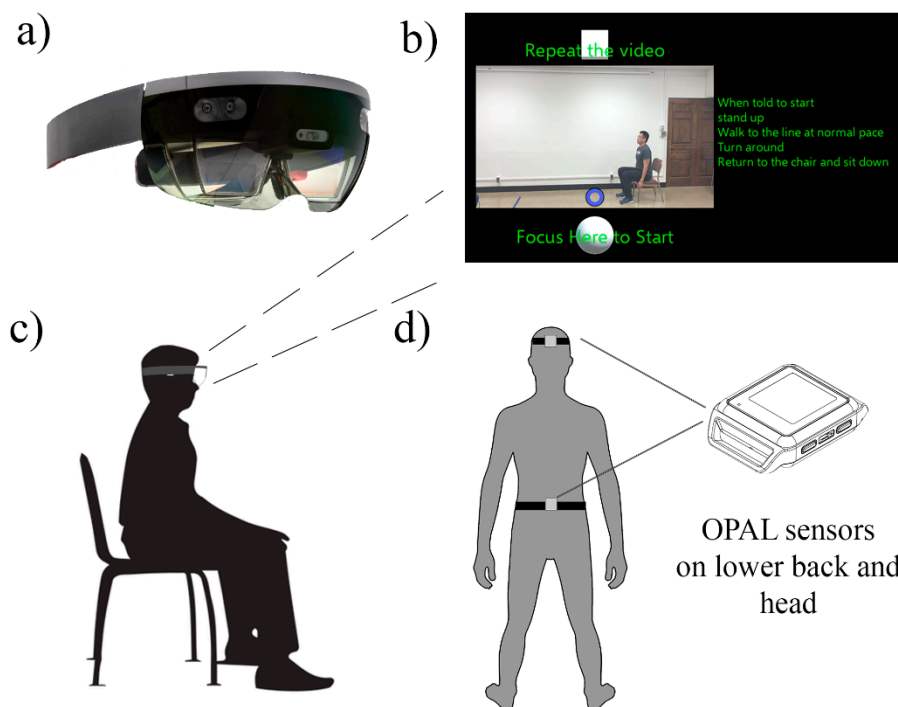
92 Twenty-two healthy adults (10 older adults -OA, 12 young adults-YA) participated in this study.
93 The inclusion criteria for participation were age between 18-30 years or 65 years and older, able to
94 stand 30s unaided, able to walk with or without aid, have normal or corrected to normal hearing and
95 vision, no history of neuro-muscular or cardiovascular disease, and no history of motion sickness,

96 chronic neck pain or seizure-related conditions. All procedures were approved by the University of
 97 Illinois at Urbana-Champaign Institutional Review Board, and all participants completed written
 98 informed consent prior to participation. Participant testing was performed at 2 sites selected for
 99 convenience for the participants. OA were tested at an unoccupied apartment setting at a local
 100 retirement community, while YA were tested at the university laboratory setting.

101 2.2 System Setup

102 A customize built Universal Windows Platform application was developed under Unity (2018
 103 2.6 personal) and Visual Studio (Microsoft Visual Studio 2017), and deployed on the Microsoft
 104 HoloLens head-mounted display operating under the Windows 10 system. The HoloLens features
 105 depth cameras for environment scanning and spatial mapping, as well as IMU for position and
 106 orientation estimation [15]. The transparent visor and light projector allow user to see high-definition
 107 virtual content (hologram) over real world objects (Fig.1b). The field of view from HoloLens was
 108 estimated as 30° H and 17.5° V [15]. The system can operate as a stand-alone device that requires
 109 neither PC nor smartphone. For this project, the onscreen display was also streamed on a laptop for
 110 monitoring participant's interaction with the system. The HoloLens features multimodal user
 111 interaction methods, such as finger pinch, voice command, and estimated gaze fixation, etc. In order
 112 to simplify user interaction and allow intuitive control for senior user, we chose to use gaze fixation
 113 (orientation estimated) to control the interface, i.e. user will control the system by fixating their gaze
 114 on a control button for 1-2 s (Fig 1b, purple circle as gaze cursor, white blocks as control buttons.
 115 Video in Supplementary S1). Additionally, to facilitate user onboarding and ensure self-guided test
 116 completion, participant watched a standardized tutorial video on a laptop explaining how to put on,
 117 adjust and control the headset (Supplementary S2). Participants were encouraged to ask questions
 118 before putting on the headset. Details about the user-interface design process and usability
 119 investigation will be reported in a separate publication.

120 **Figure 1.** Illustration of system setup. (a) HoloLens headset; (b) Onscreen hologram instruction for



121 Timed Up and Go test (Shaded background, video animation, green font, white control button and
 122 purple gaze cursor); (c) Illustration of participant's starting position; (d) Reference IMU sensors
 123 placement.

124 Based on the CDC fall risk assessment recommendations (Stopping Elderly Accidents, Death &
125 Injuries - STEADI [6]) and feasibility of head-mount movement tracking, a set of valid and reliable
126 clinical tests focusing on mobility and muscle strength were selected and integrated in the automated
127 mobility assessment App. Muscle strength and coordination was assessed with the STS test, whereas
128 mobility was assessed with TUG test.

129 2.3 Test Procedure

130 Participants completed an MR-based mobility assessment as well as a clinical fall risk
131 assessment. During the MR-based assessment, participants were outfitted with two additional
132 APDM Opal IMU sensors (APDM, Inc.). One was secured to the top of the HoloLens visor – denoted
133 as HD sensor and one placed on participant's lower back via belt – denoted as LB sensor, Fig 1d).
134 After being fitted with the headset, participants were prompted to complete test sessions in the
135 following order: 1) STS, 2) TUG. For each test, a video recording with standard demonstration and
136 instruction was displayed on the headset (Fig. 1b, center display and text instruction). Upon video
137 completion, participants were provided the option to proceed to test or repeat the demonstration
138 (Fig. 1b, white blocks). During the test, a 5s count down with audio tone and text of the instruction
139 was displayed. Test completion button was prompted up after a 15s delay. After test completion,
140 participants had options to repeat the test if not satisfied with their performance. To ensure identical
141 test setup between participants, markings were placed on the ground (standing feet placement, chair
142 location and 3 m walking path). Research personnel offered safety spotting and minimal interaction
143 with the participant, unless asked by the participant to help. Participant who made errors during test
144 (i.e. false start, incorrect number of repetitions, etc.) were asked to repeat the test.

145 After completion of the self-guided mobility assessment, the physiological profile assessment
146 (PPA) [18] was administered by trained research personnel to evaluate the overall fall risk. The
147 Montreal Cognitive Assessment (MoCA) [19] and the Activity-specific Balance Confidence (ABC)
148 scale [20] were also administered to assess participant's cognitive function and balance confidence,
149 respectively. The MoCA test is a validated screening tool for detecting cognitive impairment,
150 whereas the ABC scale is a validated self-reported questionnaire of confidence in performing various
151 daily activities without losing balance. The PPA consists of a set of comprehensive tests assessing
152 vision, lower limb sensation, muscle strength, reaction time and balance that are associated with risk
153 of falling [18].

154 2.4 Data Processing

155 Due to HoloLens API setup on sensor data access, raw accelerometry/gyroscope data was not
156 accessible, thus only the processed head position and orientation was available for recording at a
157 dynamic sampling rate at/around 30Hz (variation due to windows internal clock frequency). Such
158 data was processed by its internal proprietary sensor fusion algorithm (IMU, depth and
159 environmental cameras) that output the 3D head spatial coordinate and gravity aligned orientation.
160 Acceleration and gyroscope data from the Opal sensor was recorded at 128Hz, and gravity corrected
161 after orientation estimation using an extended Kalman filter provided by APDM. Both HoloLens and
162 Opal data was segmented and synced to each task, resampled at 30 Hz and low-pass filtered (4th
163 order Butterworth) with cutoff frequency at 5 Hz [21-23].

164 For STS and TUG tests, the gravity corrected vertical (VT) acceleration data from Opal sensors
165 were double integrated over time to obtain the vertical displacement, with drift and integration
166 error corrected using 1) a high pass filter (4th order Butterworth, 0.1Hz) [23] and 2) drift
167 correction under the assumption that participants reach the same height when they make contact
168 with the chair (zero displacement update-ZDU) [24]. The processed VT displacement from HoloLens
169 and Opal sensors was then time-aligned using cross correlation analysis (calculating the similarity
170 and time lag between signal). Finally, time-aligned VT kinematic data (displacement, velocity, and
171 acceleration) profile were derived from Opal and HoloLens using numerical integration and
172 differentiation accordingly. The VT data were also used to calculate the following performance
173 features: 1) STS duration: time between the initiation of first chair rising and the completion of last

174 chair descend [21, 22]. 2) STS mean duration of the sitting and standing phase [21, 22]. 3) Maximum
 175 acceleration and velocity during STS. 4) TUG duration: time between the initiation of chair rising and
 176 the completion of chair descend. 5) Maximum acceleration and velocity during TUG. Due to
 177 significant signal drift over time in AP, ML direction using Opal sensor (see discussion), and lack of
 178 viable signal correction method, signal comparison in AP/ML direction was not performed for STS
 179 and TUG tests.

180 Pairwise signal agreement between HoloLens, HD and LB sensors were analyzed using the
 181 normalized root mean squared error (NRMSE – RMSE divided by signal amplitude range), as well
 182 as the cross-correlation coefficient (correlation coefficient at zero lag, denoted as Xcor). The NRMSE
 183 reports the error as a ratio of the measurement range, with lower values indicates better signal
 184 agreement [25]. The Xcor measure the signal similarity of two time series, with higher value (~1)
 185 indicates better signal agreement [23]. Mean and 95% confidence interval of NRMSE and Xcor were
 186 calculated by the functional test condition (STS, TUG).

187 STS and TUG duration derived from HoloLens measure was compared with manual stopwatch
 188 tracking using Bland-Altman limit of agreement analysis [26]. The Bland-Altman limit of agreement
 189 is a robust statistical approach to indicate the level of agreement between any two measurements.
 190 Since a high correlation between any two methods does not necessarily mean that the two methods
 191 are in good agreement, the Bland-Altman technique is utilized in many studies to investigate the
 192 presence of absolute agreement between the two technologies.

193 For all derived mobility features from the HoloLens (STS duration, STS mean duration of the
 194 sitting and standing phase, Maximum acceleration and velocity during STS, TUG duration,
 195 Maximum acceleration and velocity during TUG), group (OA and YA) comparison was also
 196 conducted using two-tail student t-test. All data processing and statistical analysis was performed
 197 with customized MATLAB program (MathWorks, Inc.)

198 3. Results

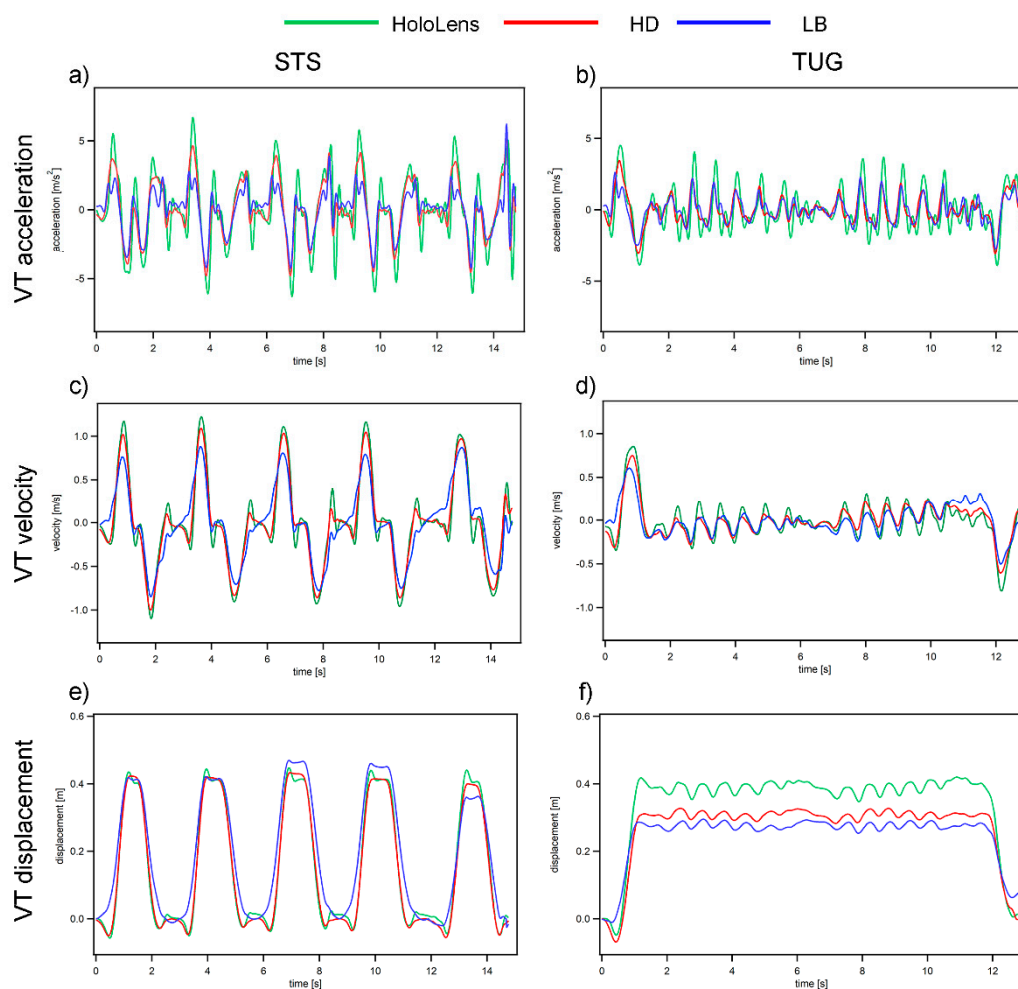
199 **Table 1.** Participant characteristics (mean and standard deviation). * indicates significant group
 200 difference ($p < 0.05$)

	OA n=8,6F	YA n=12, 6F
Age (yrs) *	78.2 (6.1)	24.4 (3.9)
BMI (kg/m ²)	23.9 (3.6)	24.5 (2.9)
MoCA *	26.2 (2.3)	28.6 (1.7)
ABC	88.8 (13.3)	96.0 (3.7)
MET *	19.9 (1.5)	21.2 (0.6)
RT (ms) *	257.7 (33.6)	217.5 (32.8)
Proprio	3.0 (1.2)	3.3 (3.5)
KneeMax (kgf) *	25.3 (9.6)	41.9 (8.5)
AP sway (mm)	27.2 (9.8)	20.8 (10.9)
ML sway (mm)	33.7 (18.9)	20.5 (12.3)
PPA *	0.9 (0.7)	-0.3 (0.7)

201 Two older participants (82 and 91 years old, both male and have significant fall risk -PPA >2)
 202 did not complete the self-guided assessment, due to balance/mobility deficits and need for assistance
 203 in challenging conditions. Therefore, only 8 OA and 12 YA were included in data analysis.
 204 Additionally, 2 OA participants were asked to repeat the test due to performance error (false start,
 205 wrong number of sit to stand repetition, etc.). Participants' demographic characteristics and
 206 physiological profile measured by PPA are presented in Table.1. As expected, significant difference
 207 in age, cognitive function (MoCA), contrast visual acuity (MET), reaction time, muscle strength, and

208 overall risk of falls were observed between OA and YA. It is worth noting only 1 of the remaining 8
 209 OA has significant fall risk (PPA > 2).

210 Representative data traces of STS and TUG test from a young participant (27 years old, female)
 211 are illustrated in Fig 2. Vertical kinematic data comparison between HoloLens and Opal sensors using
 212 NRMSE and Xcor are presented in Table 2. Overall, data recorded from HoloLens (green line in Fig
 213 2) is in good agreement with the HD/LB sensors (red/blue line in Fig 2). In general, signal similarity
 214 as measured by the Xcor is good to excellent (0.740-0.998), with higher Xcor observed in displacement
 215 measure (0.965-0.998) in comparison to velocity (0.853-0.979) and acceleration (0.740-0.888). This
 216 finding can be explained as the displacement drift was corrected using the ZDU method, whereas
 217 velocity and acceleration signal from HD/LB sensors remain slightly affected by the
 218 integration/differential error. For STS task, signal agreement as measured by NRMSE were excellent
 219 (below 10%) for all measures except displacement comparison between HoloLens and LB sensor
 220 (11.88%). Whereas for TUG task, signal agreement was relatively low (close to 20% NRMSE for
 221 displacement comparison), likely due to ZDU drift correction only performed 1 time over the entire
 222 recording for HD/LB sensor (contrary to drift correction after each sitting cycle in STS task), resulting
 223 in more vertical displacement bias (Fig 2e).



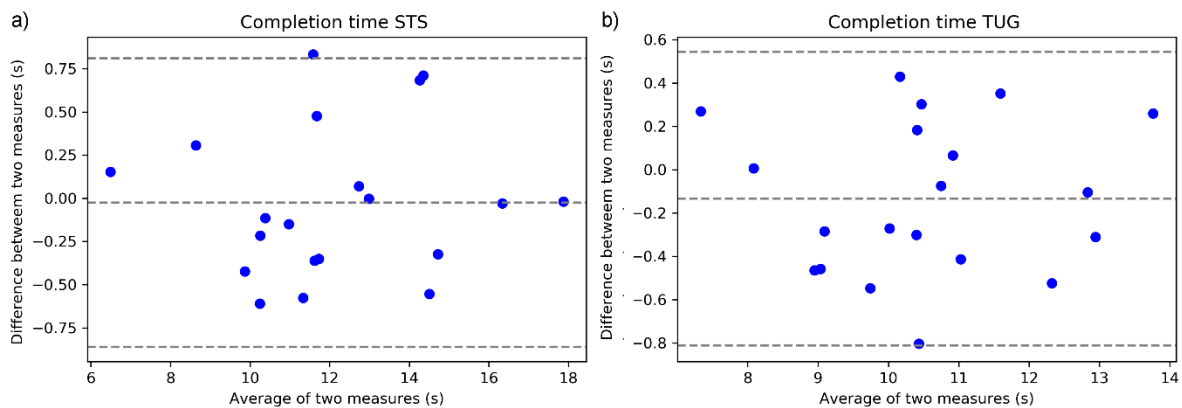
224

225 **Figure 2.** Sample kinematic profile from a young participant. Green denotes HoloLens, Red denotes
 226 HD sensor, Blue denotes LB sensor. (a, b) VT acceleration profile from STS and TUG task; (c, d) VT
 227 velocity profile from STS and TUG task; (e, f) VT displacement profile from STS and TUG task.

228 **Table 2.** Kinematic measurement (VT acceleration, velocity, and displacement) agreement between
 229 HoloLens and HD/LB sensors. NRMSE – Normalized Root Mean Squared Error. Xcor – Cross
 230 Correlation Coefficient. All values reported as mean and 95% confidence interval.

HoloLens vs HD				
		A	V	D
STS	NRMSE	9.60 (8.70,10.51)	4.83(4.22,5.45)	5.58 (4.29, 6.87)
	Xcor	0.888(0.872,0.904)	0.979(0.975,0.983)	0.993 (0.989 0.997)
TUG	NRMSE	10.53 (9.60,11.46)	6.16 (5.57,6.76)	19.56 (17.24,21.87)
	Xcor	0.802 (0.770,0.834)	0.926(0.918,0.934)	0.998 (0.997 0.999)
HoloLens vs LB				
		A	V	D
STS	NRMSE	9.77 (8.29,11.25)	8.55 (6.98,10.12)	11.88 (9.72,14.03)
	Xcor	0.765 (0.704,0.827)	0.900(0.851,0.949)	0.965 (0.949,0.982)
TUG	NRMSE	8.48 (7.56,9.41)	7.68 (7.05,8.31)	14.07 (11.86,16.28)
	Xcor	0.740 (0.695,0.786)	0.853 (0.835,0.872)	0.986 (0.978 0.993)

231 Figure 3a and 3b shows the Bland-Altman plot for the agreement in STS and TUG completion
 232 time between HoloLens and manual stopwatch recording. The absolute difference between each data
 233 pair is plotted against their mean. The two horizontal lines represent the 95% limits of agreement
 234 (range of error) calculated as 1.96 times the standard deviation from the mean differences between
 235 two methods. The figure illustrates that the mean difference between two methods is less than 0.02 s
 236 for STS and 0.13 s for TUG measure, with range of error within ± 0.8 s.



237 **Figure 2.** Bland-Altman plot of sensor-derived and stopwatch timed task completion time. (a)
 238 STS completion time; (b) TUG completion time.

239

240 Table 3 shows the group comparison (OA vs YA) in functional performance outcomes derived
 241 from the HoloLens measures. Overall, there is no significant difference between groups, with
 242 marginal significant difference observed in max velocity in STS and TUG task, reflecting the healthy
 243 nature of the OA samples (only 1 of 8 participants who completed the assessment have significant
 244 fall risk).

245 **Table 3.** Group differences of key outcome measures (mean and standard deviation).

Task	Outcome Measures	OA	YA	<i>p</i>
STS	Total Time (s)	12.22 (3.61)	12.08 (1.99)	0.922
	Mean Stand Time (s)	0.52 (0.18)	0.64 (0.22)	0.198
	Mean Sitting Time (s)	1.15 (0.55)	1.03 (0.23)	0.575
	Max Acceleration (m/s ²)	4.75 (1.81)	6.22 (2.03)	0.108
	Max Velocity (m/s)	1.02 (0.16)	1.20 (0.28)	0.087
TUG	Total Time (s)	10.61 (2.37)	10.56 (1.00)	0.96
	Max Acceleration (m/s ²)	3.98 (0.92)	3.97 (0.62)	0.961
	Max Velocity (m/s)	0.69 (0.10)	0.81 (0.15)	0.059

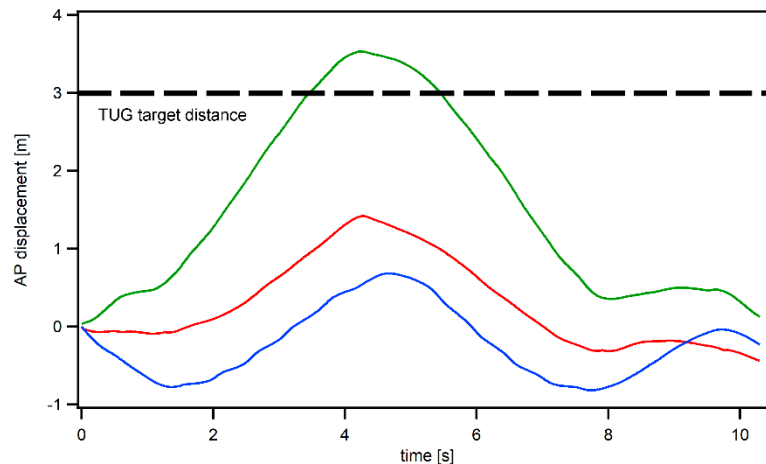
246

247 4. Discussion

248 Mixed Reality headset holds great promises for enabling portable, self-guided mobility
 249 assessment that can be undertaken more regularly without clinician oversight, and subsequently
 250 increase the efficiency of current healthcare practice. This investigation is the first to evaluate the
 251 validity of MR headset (HoloLens) for mobility assessment in young and older adults. Given the
 252 unique advantage of multi-modal user interaction methods (visual/audio/gesture/gaze), this device
 253 enables users to initiate and complete a set of valid mobility assessment with step by step guidance,
 254 and record head movement as a mean for objective measure of performance.

255 Overall, our preliminary investigation indicates that it is possible to use a mixed reality headset
 256 to automatically guide both young and old user to complete common functional mobility tests (TUG
 257 and STS) with good measurement accuracy in comparison to industry standard inertial sensors. More
 258 specifically, by comparing the vertical kinematic measures (displacement, velocity, and acceleration)
 259 derived from the HoloLens and Opal sensors, we found good to excellent signal agreement for the
 260 majority of STS and TUG measures (X_{cor} 0.74-0.99, NRMSE ~10%), with better signal agreement
 261 observed in STS task, as each sit-to-stand cycle allows for displacement calibration adjustment (Zero
 262 Displacement Update) [24]. For sensor signal comparison in TUG task, however, due to the
 263 integration and drift error associated with IMU sensor and lack of viable calibration during the
 264 walking period (10s or longer), the vertical displacement derived from IMU sensors was biased in
 265 comparison to HoloLens output (Fig.2e), resulting unsatisfactory NRMSE measures (14.07-19.56%).
 266 Moreover, because of HoloLens utilized both depth sensor and IMU sensor to derive the
 267 displacement measure (the depth sensor on HoloLens is similar to the Kinect sensor, which has been
 268 extensively validated for accuracy in kinematic measure [27, 28]), we would expect the HoloLens
 269 output in displacement measure is more trustworthy than IMU derived displacement measure.
 270 Although AP displacement comparison between sensors was not analyzed statistically (due to lack
 271 of viable drift correction method for IMU sensors), our exploratory investigation found that only
 272 HoloLens AP displacement measure matches the standard 3m walking distance utilized in the TUG,
 273 whereas the IMU derived AP displacement measure severely underestimate the walking distance
 274 (evident by the displacement measure in AP direction as shown in Fig. 3).

275



276 **Figure 3.** Sample AP displacement in TUG task. Red denotes HD sensor, Blue denotes LB sensor.
 277 Note only HoloLens AP displacement measure correctly match the 3m walking path utilized in
 278 TUG task.

279 Additionally, the sensor-derived completion time of STS and TUG was also compared with
 280 current gold standard manual stopwatch method using Bland-Altman agreement plot. Excellent
 281 agreement with stopwatch timing was found for both STS and TUG completion time, with HoloLens
 282 measure demonstrate less than 0.2s in measurement bias and less than 0.8s in range of error).

283 Because of the relatively healthy nature of the OA participant who completed the test without
 284 any assist (only 1 OA participant has significant fall risk, PPA >2), no group difference for sensor-
 285 derived performance measure was found between YA and OA, with only marginal difference
 286 observed in maximum ascending velocity in STS and TUG task. The two excluded OA participants
 287 who have significant fall risk, however, could not complete the STS and TUG task without requiring
 288 physical assist, indicating the stand-alone HoloLens device and self-guided functional mobility tests
 289 may not be suitable for those who already have severe mobility deficits.

290 We acknowledge certain limitations for this investigation, most of which related to the
 291 pioneering use a novel technology. First, due to the lack of optical motion tracking equipment for
 292 portable/community testing, industry standard IMUs were utilized for this validation study.
 293 Therefore, due to IMU's inherent integration/drift error and the limited data access to HoloLens
 294 displacement data, horizontal (AP/ML) kinematic measurement comparison was not conducted.
 295 Secondly, the relatively small sample size, and healthy nature of OA participants who can complete
 296 the functional tests without any assist, preclude the investigation to detect the diagnostic power of
 297 using head-mount device for fall risk/mobility deficit screening in older adults. Therefore, future
 298 studies should incorporate optical motion tracking and larger heterogeneous samples to investigate
 299 the use of HoloLens for fall risk screening.

300 **Supplementary Materials:** The following are available online, Video S1: Demonstration recording of the MR-
 301 based TUG task. Video S2: Tutorial video shown to user on how to put on and adjust the headset.

302 **Author Contributions:** conceptualization, R.S and J.S.; methodology, R.S.; software, R.S. and R.A.; formal
 303 analysis, R.S.; writing—original draft preparation, R.S.; writing—review and editing, R.S., R.A and J.S.;
 304 visualization, R.S.; supervision, J.S.; project administration, J.S.; funding acquisition, R.S. and J.S

305 **Funding:** This research was funded by University of Illinois at Urbana Champaign CHART Seed grant

306 **Acknowledgments:** The authors would like to thank Microsoft for donating the HoloLens for this research
 307 project. The authors would also like to acknowledge Mr. Shrey Pareek for assistance in the software
 308 development.

309 **Conflicts of Interest:** The authors declare no conflict of interest.

310

311 **Appendix A**312 **List of abbreviations**

313	IMU – Inertial Measurement Unit
314	HMD – Head-mounted Display
315	MR – Mixed Reality
316	MoCA- Montreal Cognitive Assessment,
317	ABC- Activity-specific Balance Scale
318	STS - Five time Sit to Stand test
319	TUG – Timed Up and Go test
320	PPA – Physiological Profile Assessment
321	RT – Reaction Time
322	MET – Melbourne Edge Test
323	Proprio- Proprioception
324	KneeMax – Maximal isometric knee extension force
325	AP – Anterior Posterior
326	ML – Medial Lateral
327	VT - Vertical
328	OA - Older Adults
329	YA – Young Adults
330	HD – Head IMU sensor
331	LB – Lower Back IMU sensor
332	A – Acceleration
333	V – Velocity
334	D – Displacement
335	NRMSE – Normalized Root Mean Squared Error
336	Xcor – Cross Correlation Coefficient
337	ZDU – Zero Displacement Update
338	

339 **References**

- 340 1. Rubenstein, L. Z., Falls in older people: epidemiology, risk factors and strategies for prevention. *Age*
341 *and ageing* **2006**, 35, (suppl_2), ii37-ii41.
- 342 2. Bergen, G., Falls and fall injuries among adults aged ≥ 65 years—United States, 2014. *MMWR. Morbidity*
343 *and mortality weekly report* **2016**, 65.
- 344 3. Florence, C. S.; Bergen, G.; Atherly, A.; Burns, E.; Stevens, J.; Drake, C., Medical costs of fatal and
345 nonfatal falls in older adults. *Journal of the American Geriatrics Society* **2018**, 66, (4), 693-698.
- 346 4. Ory, M. G.; Schechtman, K. B.; Miller, J. P.; Hadley, E. C.; Fiatarone, M. A.; Province, M. A.; Arfken, C.
347 L.; Morgan, D.; Weiss, S.; Kaplan, M., Frailty and injuries in later life: the FICSIT trials. *Journal of the*
348 *American Geriatrics Society* **1993**, 41, (3), 283-296.
- 349 5. Panel on Prevention of Falls in Older Persons, A. G. S.; Society, B. G., Summary of the updated
350 American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls
351 in older persons. *Journal of the American Geriatrics Society* **2011**, 59, (1), 148-157.
- 352 6. Stevens, J. A.; Phelan, E. A., Development of STEADI: a fall prevention resource for health care
353 providers. *Health promotion practice* **2013**, 14, (5), 706-714.
- 354 7. Buatois, S.; Miljkovic, D.; Manckoundia, P.; Gueguen, R.; Miget, P.; Vançon, G.; Perrin, P.; Benetos, A.,
355 Five times sit to stand test is a predictor of recurrent falls in healthy community-living subjects aged 65
356 and older. *Journal of the American Geriatrics Society* **2008**, 56, (8), 1575-1577.

- 357 8. Shumway-Cook, A.; Brauer, S.; Woollacott, M., Predicting the probability for falls in community-
358 dwelling older adults using the Timed Up & Go Test. *Physical therapy* **2000**, *80*, (9), 896-903.
- 359 9. Sun, R.; Sosnoff, J. J., Novel sensing technology in fall risk assessment in older adults: a systematic
360 review. *BMC geriatrics* **2018**, *18*, (1), 14.
- 361 10. Shany, T.; Wang, K.; Liu, Y.; Lovell, N. H.; Redmond, S. J., Are we stumbling in our quest to find the
362 best predictor? Over-optimism in sensor-based models for predicting falls in older adults. *Healthcare
363 technology letters* **2015**, *2*, (4), 79-88.
- 364 11. Shany, T.; Redmond, S.; Marschollek, M.; Lovell, N., Assessing fall risk using wearable sensors: a
365 practical discussion. *Zeitschrift für Gerontologie und Geriatrie* **2012**, *45*, (8), 694-706.
- 366 12. Howcroft, J.; Kofman, J.; Lemaire, E. D., Review of fall risk assessment in geriatric populations using
367 inertial sensors. *Journal of neuroengineering and rehabilitation* **2013**, *10*, (1), 91.
- 368 13. Hsieh, K. L.; Fanning, J. T.; Rogers, W. A.; Wood, T. A.; Sosnoff, J. J., A Fall Risk mHealth App for Older
369 Adults: Development and Usability Study. *JMIR Aging* **2018**, *1*, (2), e11569.
- 370 14. Hsieh, K. L.; Roach, K. L.; Wajda, D. A.; Sosnoff, J. J., Smartphone technology can measure postural
371 stability and discriminate fall risk in older adults. *Gait & posture* **2019**, *67*, 160-165.
- 372 15. Kress, B. C.; Cummings, W. J. In *11-1: Invited Paper: Towards the Ultimate Mixed Reality Experience:
373 HoloLens Display Architecture Choices*, SID Symposium Digest of Technical Papers, 2017; Wiley Online
374 Library: 2017; pp 127-131.
- 375 16. Menz, H. B.; Lord, S. R.; Fitzpatrick, R. C., Acceleration patterns of the head and pelvis when walking
376 on level and irregular surfaces. *Gait & posture* **2003**, *18*, (1), 35-46.
- 377 17. Psarakis, M.; Greene, D. A.; Cole, M. H.; Lord, S. R.; Hoang, P.; Brodie, M., Wearable technology reveals
378 gait compensations, unstable walking patterns and fatigue in people with multiple sclerosis.
379 *Physiological Measurement* **2018**, *39*, (7), 075004.
- 380 18. Lord, S. R.; Menz, H. B.; Tiedemann, A., A physiological profile approach to falls risk assessment and
381 prevention. *Physical therapy* **2003**, *83*, (3), 237-252.
- 382 19. Nasreddine, Z. S.; Phillips, N. A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings,
383 J. L.; Chertkow, H., The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild
384 cognitive impairment. *Journal of the American Geriatrics Society* **2005**, *53*, (4), 695-699.
- 385 20. Powell, L. E.; Myers, A. M., The activities-specific balance confidence (ABC) scale. *The Journals of
386 Gerontology Series A: Biological Sciences and Medical Sciences* **1995**, *50*, (1), M28-M34.
- 387 21. Ejupi, A.; Brodie, M.; Gschwind, Y. J.; Lord, S. R.; Zagler, W. L.; Delbaere, K., Kinect-based five-times-
388 sit-to-stand test for clinical and in-home assessment of fall risk in older people. *Gerontology* **2016**, *62*, (1),
389 118-124.
- 390 22. Sun, R.; Aldunate, R. G.; Paramathayalan, V. R.; Ratnam, R.; Jain, S.; Morrow, D. G.; Sosnoff, J. J.,
391 Preliminary evaluation of a self-guided fall risk assessment tool for older adults. *Archives of gerontology
392 and geriatrics* **2019**.
- 393 23. Sun, R.; Moon, Y.; McGinnis, R. S.; Seagers, K.; Motl, R. W.; Sheth, N.; Wright, J. A.; Ghaffari, R.; Patel,
394 S.; Sosnoff, J. J., Assessment of postural sway in individuals with multiple sclerosis using a novel
395 wearable inertial sensor. *Digital Biomarkers* **2018**, *2*, (1), 1-10.
- 396 24. Millor, N.; Lecumberri, P.; Gómez, M.; Martínez-Ramírez, A.; Izquierdo, M., Drift-free position
397 estimation for periodic movements using inertial units. *IEEE journal of biomedical and health informatics*
398 **2014**, *18*, (4), 1131-1137.

- 399 25. Walgaard, S.; Faber, G. S.; van Lummel, R. C.; van Dieën, J. H.; Kingma, I., The validity of assessing
400 temporal events, sub-phases and trunk kinematics of the sit-to-walk movement in older adults using a
401 single inertial sensor. *Journal of Biomechanics* **2016**, *49*, (9), 1933-1937.
- 402 26. Bland, J. M.; Altman, D., Statistical methods for assessing agreement between two methods of clinical
403 measurement. *The lancet* **1986**, *327*, (8476), 307-310.
- 404 27. Clark, R. A.; Bower, K. J.; Mentiplay, B. F.; Paterson, K.; Pua, Y.-H., Concurrent validity of the Microsoft
405 Kinect for assessment of spatiotemporal gait variables. *Journal of biomechanics* **2013**, *46*, (15), 2722-2725.
- 406 28. Clark, R. A.; Pua, Y.-H.; Fortin, K.; Ritchie, C.; Webster, K. E.; Denehy, L.; Bryant, A. L., Validity of the
407 Microsoft Kinect for assessment of postural control. *Gait & posture* **2012**, *36*, (3), 372-377.