ADDITIONAL MATERIAL: USABILITY TESTS WITH THE SOUND OF VISION SYSTEM IN OUTDOOR ENVIRONMENTS - EXPLORATION OF THE INDIVIDUAL PERFORMANCE

USER 1

User 1 is middle aged, graduated high school and is congenitally blind. He is currently unemployed and lives with his spouse who is also blind. He is highly involved in different activities organized by the local blind community. He was born blind, being in category 4 of visual impairment (according to WHO), he can only distinguish between light and darkness. He travels is familiar environments with his partner using the white cane, but he is not an experienced white cane user. In unfamiliar places he goes only accompanied by sighted persons. He was involved in training and testing with the SoV device since the first prototype. He has an optimistic attitude regarding the possibility of using this device in the future, hoping that it will increase his mobility, independence and safety especially when travelling in unknown environments. He is a calm and patient person and he concentrates very well on the tasks he has to do, performing them meticulously. He likes to sing and seems to have a passion for smartphones and IT technologies, especially those which could help him.

Level of experience

User 1 was involved in training and testing the SoV device since its first prototype, so he is very familiar with the system and its improvements between the three tested prototypes.

Preferred encoding

User 1 preferred to use the expanding sphere model with the impact sounds in audio and closest point in haptics. He has had this preference since the testing sessions with the previous prototype. In real world indoor experiments, he also preferred to use the continuous audio encoding (bubblestream) for the mobility tasks. However, he found it more difficult to use it in real world outdoor as the maximum sound volume for this encoding was too low to be perceived efficiently in noisy environments. Although he could turn off the sounds in the environment (using the active noise cancellation feature of the audio headset), he preferred to change the encoding than to not hear the environmental sounds.

Performance

Being involved in training and testing since the first prototype, User 1 was already accustomed with the device. He successfully completed all the perception tests, except for PT I-3, where he correctly identified the number of objects, but he could not report the correct direction, distance and elevation for each of them (Figure S1).

The main difference between PT I and PT II perception scenarios is that in the second one special objects were added to the scene, beside static and dynamic objects. As it can be observed in Figure S2, user's accuracy regarding the count-static, distance and direction increases between the two testing sessions. The exceptions here are both count-dynamic and elevation. User 1 found it difficult to point if there is a dynamic obstacle in the scene, when a special object (hole in the ground) was also present. He said that the audio encoding for negative obstacles provides a very strong feedback, distracting his attention from the moving objects/persons. Regarding the elevation, even if the participant correctly identified the number of static objects, he could not tell for sure if there is a head height obstacle in the scene. Still, he pointed out that it is extremely important that he has information for that specific object (hanging) even if he can't tell its elevation.

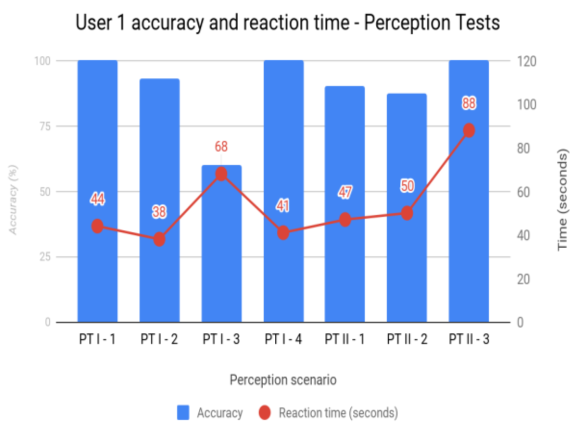

Figure S1. The individual accuracy and reaction time of User 1 in perception tests

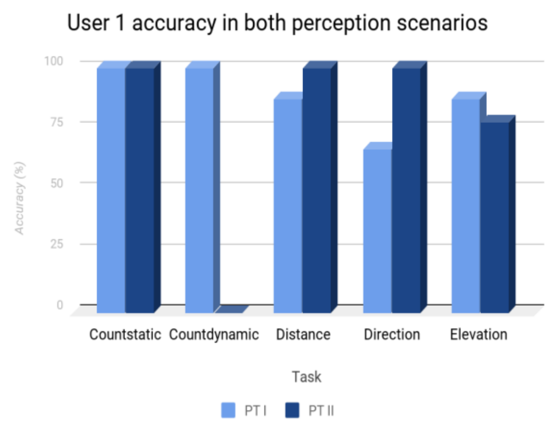

Figure S2. The individual accuracy of User 1 in both perception scenarios

User 1 **identified the special objects**, i.e., holes and walls, and their properties (type, number, distance, direction) with **100% accuracy** in all perception scenarios.

He successfully completed all the mobility tasks, except for MT I-3, when performed with SoV & white cane. The user did not identify the last car on the left due to a false positive object detection by the system (the system output was constantly inspected by one test assistant).

In MT I-1, the user chose to walk close to the wall in both scenarios that involved the white cane. However, **he reported to rely more on the system when he tested with both SoV and cane**. This is consistent with the number of wall hits with the cane in the two situations (11 when testing only with cane, 0 when testing with SoV & cane). Moreover, he bumped into the static obstacle on the path, but only when testing in condition 1 (with cane only).

He **successfully identified and used the wall as shoreline in navigation with the SoV system**, in both MT I-1 and MT I-2. He could also easily detect the corner of the building in both scenarios.

He **successfully identified the bus stop** in MT II-1, which he found to be a very helpful feature. He also **emphasized the added value of SoV compared to the cane in detecting hanging objects**. Although he scored below 85% accuracy in perceiving the elevation of objects in ego-static tests, he said that having the hanging objects specifically encoded by the danger mode was very helpful in avoiding them in the mobility scenarios. Moreover, he felt that when using the system, understanding their elevation is not that important as long as the objects are reported by the system.

Although it took him more time to complete the two courses, MT I-1 and MT II-2, with SoV (with or without cane) than with the cane, the difference is not striking (Figure S3). Moreover, the adjusted completion time, i.e., weighted by the number of obstacles, in MT II-2, is even lower when walking only with SoV than with the cane, given that in the former case there were 3 times more extra obstacles on the path.

He had no collisions in any of the mobility tasks with SoV. Although he had one collision when using only the cane still, on average, walked faster in this condition than in the other two (cane - **0.89m/s**, cane&SoV - **0.6m/s**, SoV - **0.53 m/s**). Not surprisingly, when **using both modalities the user walked slower than when only using the SoV system**, as the amount of information coming from both devices can be overwhelming for users unexperienced with any of them.

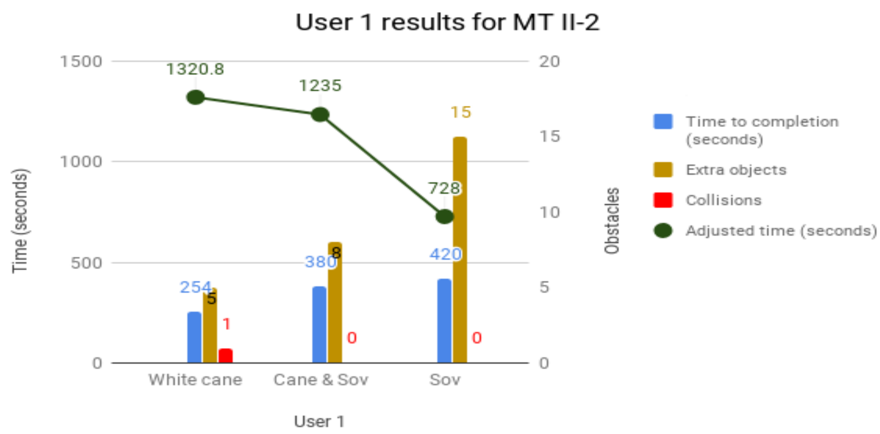

*Figure S3. **User 1** results, **walking on the sidewalk** mobility scenario with the 3 modalities*

USER 2

User 2 is of older age, graduated high school and is currently unemployed. He is married, living with his spouse who is also blind. He was not born completely blind, in childhood he had rest sight (only one eye) and was able to navigate alone. He became completely blind at the age of 12, so he has been living with blindness for a long time, belonging to category 5 of visual impairment. He worked for some years in a factory. He is also an active member of the local blind community, being eager to participate in different competitions and social activities dedicated to visually impaired people. He was very interested in the SoV device from the first prototype. He hopes that this device will become a commercial one and that he will be able to afford it. He really thinks that this kind of device will change his life. He also took part in all training and testing sessions (VTE, indoor, outdoor) with the SoV device. He is a restless person; thus, he sometimes loses his ability to concentrate and to focus on learning or doing new activities, he also expects immediate results. He does not walk alone in unfamiliar places being always accompanied by a sighted person. He is using the white cane but is not an experienced user.

Level of experience

User 2 was involved in training and testing the SoV device since the first prototype, so he is very familiar with the system and its improvements between the three prototypes.

Preferred encoding

User 2 preferred to use the expanding sphere model with the bubble sounds in audio and closest point in haptics. He has had this preference since the previously testing session he participated. However, in real world indoor experiments he also preferred to use either the continuous audio encoding (bubblestream) or the flashlight audio encoding for the mobility tasks. Like User 1, User 2 also complained regarding the maximum sound volume for the continuous audio encoding as being too low to be perceived efficiently in noisy environments. Although he could turn off the sounds in the environment (using the active noise cancellation feature of the audio headset), he preferred to change the encoding than to not hear the environmental sounds.

Performance

User 2 successfully completed all the perception tests, except for PT I-2 (Figure S4). Introducing new elements in the scene (especially dynamic objects) require more training time for User 2. Dynamic objects caused confusion such that he could not correctly describe the scene.

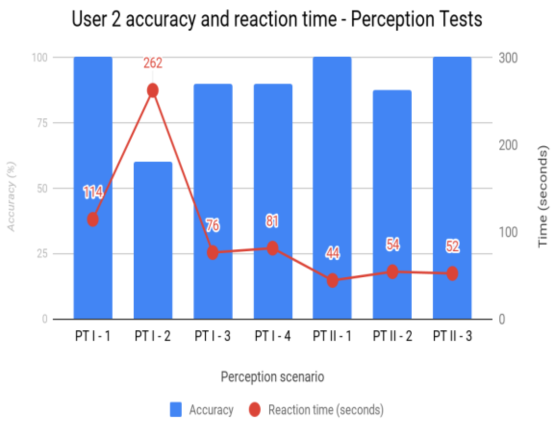

Figure S4. The individual accuracy and reaction time of User 2 in perception tests

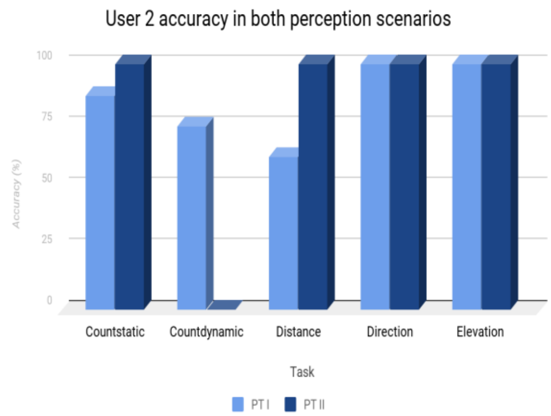

Figure S5. The individual accuracy of User 2 in both perception scenarios

User 2, had an accuracy of 100% in pointing the direction and elevation of an object. The participant claimed the same thing like User 1, that indicating the presence of a dynamic object is hampered by the audio encoding for negative obstacles. As it can be observed in Figure S5, besides identifying dynamic objects, user's accuracy regarding count-static, distance, direction and elevation increased from PT I to PT II. This can again be explained by relying on additional training time accumulated between PT I and PT II tests. User 2 emphasized the importance in detection of negative obstacles and of elevated objects and was very enthusiastic about how the system performed in these situations.

Regarding special objects, User 2 obtained an accuracy of 100% by indicating the number, direction and distance to the special objects, thanks to the intuitive audio encodings for special objects, as he commented.

User 2 successfully completed all the **mobility tasks**, except for MT I-1 with cane and MT I-2. In MT I-1 he could not identify the corner of the building just using the cane. In MT I-2 he was confused by a false positive wall detection on the opposite side of the real wall and he left the testing area. However, he managed to avoid all the hanging obstacles (tree branches) on the path, and reported to rely much on the danger mode for this.

When using both the cane and the system, he relies mostly on the cane. This fact is also illustrated by the number of wall hits with the cane in MT I-1, which was much the same for both conditions. However, he managed to successfully avoid the obstacles on the sidewalk when walking with the

SoV system. Moreover, he registered a similar adjusted completion time with only SoV in finalizing the MT II-2 course as with the white cane (Figure S6).

Like User 1, he had less collisions in both of the mobility tasks with SoV than with the white cane, but on average, walked faster in the latter condition (cane - 0.61m/s, cane & SoV - 0.38m/s, SoV - 0.25m/s).

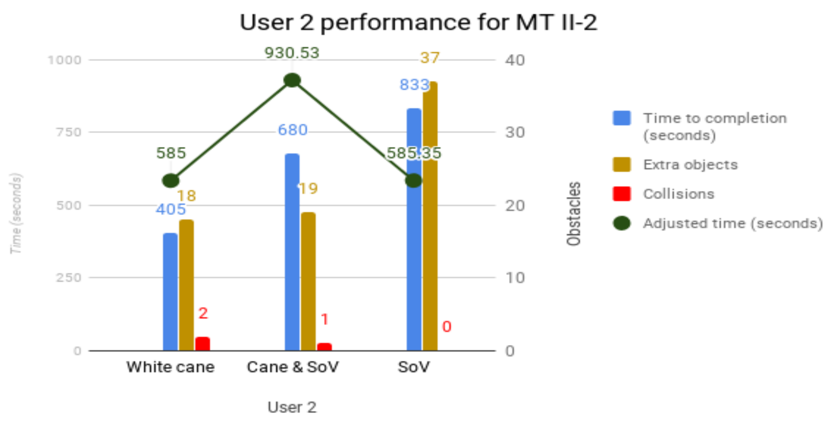

*Figure S6. **User 2** results, **walking on the sidewalk** mobility scenario with the 3 modalities*

User 3

User 3 is young, he graduated high school and now is studying foreign languages at the Faculty of Letters. He is currently living with his family, being always accompanied by a relative. Thus, he does not use the white cane. He lost sight of one of his eyes at birth and later he lost sight of the other eye due to a stupid accident while being at school. He belongs to category 5 of visual impairment. His family is very protective which can also explain him being a rather introverted and shy person. He was interested in the SoV device, participated actively in the VTE and both indoor and outdoor training and testing sessions, although sometimes he either lost his concentration either he waved before answering. Starting with outdoor training, he suddenly became more interested and enthusiastic about the device seeing the true advantages of the final prototype. He thinks that the device could give him the independence he lost when he became completely blind. As a matter of fact, he asked, at the end of all the testing sessions, whether he could take the device home.

Level of experience

User 3 was involved in training and testing with the SoV device starting with the second prototype.

Preferred encoding

User 3 preferred to use the expanding sphere model with the bubble sounds in audio and closest point in haptics. He has had this preference since the testing session with the previous prototype. However, in real world indoor experiments he also preferred to use either the continuous audio encoding (bubblestream) or the flashlight audio encoding for the mobility tasks.

Performance

Given his performance in the perception tests (Figure S7), we conclude that User 3 would need more training time. He had the shortest outdoor training session amongst all users, as he is currently studying. Yet, his results are promising.

Maybe due to fatigue or lack of interest in the system, he could not concentrate well, because as shown in the Figure S8, sometimes he could not identify the elevation, direction, in other scenarios the elevation, direction and distance or dynamic objects. There are several scenes (PT I-3, PT II-1, PT II-2) in which he could not identify 2 or 3 objects.

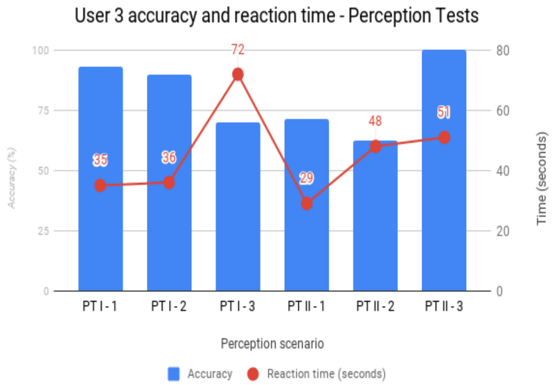

Figure S7. The individual accuracy and reaction time of User 3 in perception tests

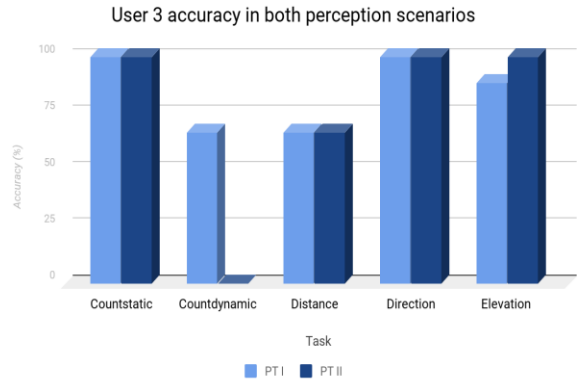

Figure S8. The individual accuracy of User 3 in both perception scenarios

User 3 suddenly expressed more interest and enthusiasm in using the system when he reached the mobility phase. He successfully completed all the **mobility tasks** but with many collisions on the sidewalk in the MT II-2. He walked the MT I-1 course by the wall with no collisions in either of the three conditions (cane, cane & SoV, SoV). Moreover, he reported to mostly rely on the SoV device in the mobility scenarios performed with both SoV and the cane. His feedback is also supported by the number of wall hits with the cane in MT I-1, which was 31 in the case of using only the cane and 0 when also using the SoV system, while keeping a ~1.5m distance to the wall on the entire course. He successfully avoided elevated objects in MT I-2 and identified the special objects (walls, bus stop).

The average number of collisions over all mobility tests, counted when the user was using only the SoV system is similar to the white cane condition. He had comparable walking speeds with the three modalities (cane - 0.74m/s, cane & SoV - 0.68m/s, SoV - 0.62m/s). He considered walking with either SoV or the cane of comparable difficulty, as he is inexperienced with either of them (Figure S9). He recently became blind and he is always accompanied by a family member when travelling in either familiar or unfamiliar environments. **He successfully walked the MT II-2 course which gave him the most confidence that he could walk alone**.

His lack of experience in travelling alone had a significant impact on his mobility behavior. Overall, he gained a lot of confidence during the mobility training and testing sessions, being more enthusiastic. He is also very confident that he could improve his performance with the system with more training in operational environments.

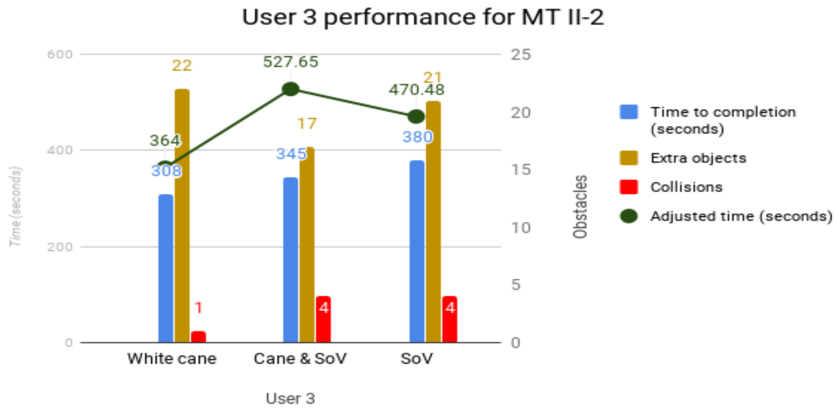

*Figure S9. **User 3** results, **walking on the sidewalk** mobility scenario with the 3 modalities*

USER 4

User 4 is also young. He was born blind and belongs to category 4 of visual impairment, being able to distinguish between light and darkness. He does not use a white cane, as he never travels alone, being always accompanied by a sighted person. He is a very optimistic person. He loves music. He likes to sing and to play music. He is also a composer. He plays seven instruments (e.g., flute, pan flute, oboe). He graduated high school and now is studying pan flute at the Arts University. He is very active and has the ability to learn quickly. Due to his talents, he participated in many cultural activities, playing one of his favorite instruments. He is a proactive person. He was involved in training and testing only with the Final SoV prototype. In a very short period, he learned the audio and haptic encodings and successfully passed training and testing sessions in VTE. He was interested in the SoV capabilities and even if he was the newest user to come, he obtained by far the best results both in perception and mobility scenarios.

Level of experience

User 4 was a newcomer with no previous experience with the SoV device. However, he was the fastest learner out of all users.

Preferred encoding

User 4 preferred the impact sounds for audio encoding claiming that the marker ticks are very helpful for him to perceive the distance to the encoded obstacles. This is also a feature that he could instantly train given his musical abilities.

Performance

Even if User 4 benefited from the shortest total training time with the system, he obtained an accuracy of 100% in almost all perception scenarios, except PT I-1 and PT I -3 in which he incorrectly identified the elevation and the dynamic object, respectively (Figure S10).

As it can be observed in Figure S11, he successfully increases his accuracy to 100% from PT I to PT II scenarios, including the dynamicity and elevation of objects, which were the most challenging tasks for the previous users. This can be explained by the additional training time with the system between the two tests.

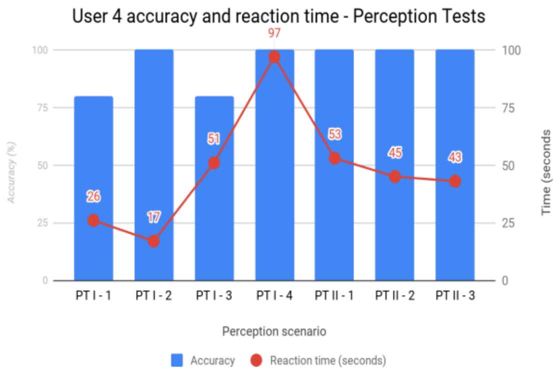

Figure S10. The individual accuracy and reaction time of ***User 4*** *in perception tests*

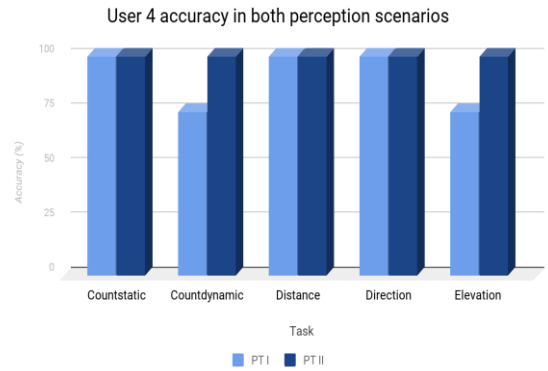

Figure S11. The individual accuracy of ***User 4*** *in both perception scenarios*

User 4 successfully completed all mobility tasks with 100% accuracy. He had the best performance in mobility with the SoV system, although he never walks alone. Moreover, he encountered the highest number of additional obstacles on the sidewalk path in MT II-2, i.e., 61 obstacles. Even so, he had no collisions and walked the course in the shortest time, out of all users, considering both measured and adjusted completion time (Figure S12).

He mostly relied on the system, when both white cane and SoV were available. In MT I-1 he hit the wall with the cane 35 times when using the white cane only, and not at all when using cane & SoV, choosing to walk at ~1.5m from the wall.

He had no collisions in any of the mobility tasks and, on average, walked slightly faster when using the SoV system than in the other two conditions (cane - 0.65m/s, cane & SoV - 0.64m/s, SoV - 0.66m/s).

He is very confident and enthusiastic about the system and had a very positive feedback regarding the detection of negative obstacles, hanging obstacles and signs. He suggested adding face recognition to the system, which would be a powerful feature for him.

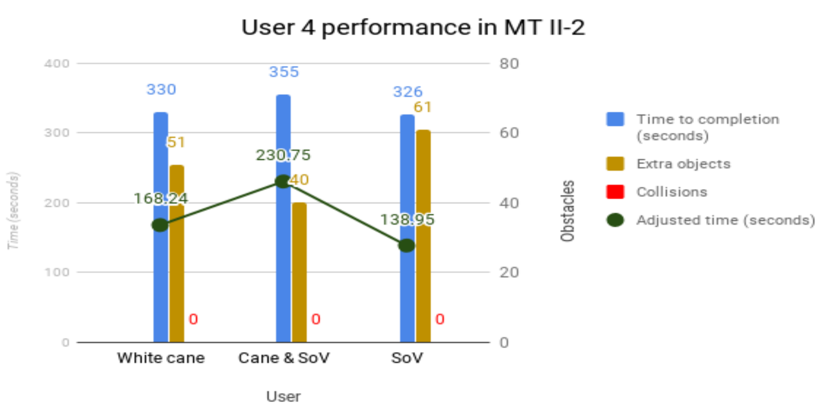

Figure S12. User 4 results, walking on the sidewalk mobility scenario with the 3 modalities