Student number:	2467273
Course title:	COMPSCI4068 Mobile HCI (H)
Questions answered:	ALL

1.

a) There are many, many challenges to do with mid-air gestures, many of which could have contributed to this feature's removal.

Firstly, mid-air gestures can very easily be misconstrued by the software as intentional input when not intended. For example, when users put their phone down on a table and are using it to listen to music, but also talking with someone else and gesturing with their hands, some of these gestures could accidentally be taken as skipping a song, which would be extremely frustrating to a user and would decrease their trust in the device. It would also make users feel tense and nervous when moving around their stationary phone since they would not want to accidentally trigger an event.

Secondly, users needed to learn a whole new interaction technique, which is itself a demotivation for some users who do not want any significant changes to the way they interact with their devices. This learning might seem to some to not be worth it for the tiny increase in usability of their device.

Thirdly, a constant sensor would almost definitely decrease battery life, which is already a problem for most smartphones in the Google Pixel series. Users might value battery life much more than this tiny usability improvement, which might lead them to alternatives which do not sacrifice this important aspect.

Fourthly, another trade-off of this feature is in decreasing resources used for other applications. The Pixel 4's RAM might not be as impressive with the constant sensor being activated, which could frustrate users because they want their devices' memory to be focused on more important tasks and apps.

Fifthly, this feature seems to discriminate further in its accessibility, i.e., mid-air gestures are not very accessible to people with movement disabilities, which reduces the audience to whom this feature might be marketed.

b) While driving a car, one of the most important aspects of designing interactions with devices is to reduce the necessary attention to the device for any intended input (eyes-off-the-road time (EORT)), in which case attention can be kept on the road and risk of any accidents is minimised. Mid-air gestures are a way to reduce this attention by having "eyes-free" input, which means that (at least visual) attention is not divided as the user interacts with the device. Additionally, these gestures can be designed to be wildly different from each other, which would mean the error margin for them could be made very large, in which case these gestures could be completed without much effort on the user's side, which, again, could reduce the attention needed to interact.

c) Firstly, to perform mid-air hand movements, at least one hand has to be taken off the steering wheel, which could be argued as a worse impact than glancing for a second away from the road. Driving with one hand increases risk of the driver not being able to respond to unexpected circumstances on the road.

Secondly, feedback is a large aspect to consider when interacting with mid-air gestures: it is not as easy to give tactile/vibrational feedback to a user as to whether their gesture succeeded in doing the intended task or whether there was any problem. If feedback is done audially, it could be a nuisance to multiple people sitting in the car if, for example, they are engaged in conversation, which would be interrupted by the device, thus frustrating its users. If feedback is not implemented, however, the lack of it might increase ambiguity of the device's users and whether they succeeded in their task, which could also be a source of frustration.

Thirdly, if the gestures are not implemented correctly, the misinterpreting of them by the device could lead to much frustration and might even distract the driver, which is the worst thing to do in this situation.

2.

a) In my opinion, an augmented reality (AR) heads-up-display (HUD) would be the most appropriate in this situation. Since packing, storing, and picking up goods are the core operations of their business, an AR solution would be most helpful in interacting with all of the packages of the goods. A light AR solution seems most appropriate because 1) augmented reality is able to supplement reality but still ground itself in reality, which is needed because the nature of a delivery company is extremely physical, 2) there are usually many employees working in delivery companies, which means that any more expensive form factors for other devices would exponentially increase costs in acquiring these devices, which would be much mitigated by a low requirement, easy to develop for device, 3) HUDs are the lightest (in weight and size) of currently available HUD devices, which would enable employees to wear them throughout their working hours without much difficulty, compared to stronger AR solutions, and 4) there seem to be no obvious requirements for direct 3D interaction for such a company, which is not a capability that HUDs have (currently).

As for direct potential uses for HUDs, for example, when looking at some barcode on the packages, the HUD could show, either symbolically or realistically, the contents of the packages, their group/class in order to know where to deliver them, how exactly to pack them, or where to store them. These devices can also be tracked or interact with other employees' devices (like smartphones or smartwatches) to enable tracking, which in turn could be used to help employees see the route to a storage or delivery location with very low requirements. Additionally, AR could be used to share/stream an employee's vision with another employee to encourage collaboration between them and essentially increase the fidelity of their conversations, e.g., it is much easier to understand questions and give instructions to someone when they can share directly what they are seeing in case things are hard to describe/imagine.

b) One potential hazard is that of occlusion, i.e., when the HUD covers important objects in reality, causing the user not to notice these objects, which could lead to many different kinds of accidents. Two ways to solve this could be to either make the display content more transparent, which would improve visibility of real-life objects, but which would also decrease visibility of the HUD content, or make the HUD less cluttered and hide some content behind interactions (through the phone, watch, direct manipulation of the HUD, or otherwise), which would clear up visibility of reality and allow the users to notice dangerous objects/situations better.

Another potential hazard is that the HUD content might be distracting, not just large. If the elements are too colourful or with too many animations/flashes, these elements might distract the user and cause them to not notice an important object/situation in front of them when walking or interacting with packages or their equipment.

c) An ethically friendly metric to measure could be some sort of cognitive load metric, like the NASA Task-Load Index (TLX), which could be used to ask employees themselves how they physically, mentally, and temporally demanding their tasks were with the device, how much effort they had to put in, how they themselves would rate their performance, along with any frustrations they might have experienced on the way. This metric could be asked at the end of a task, a working day or a working week, and making this questionnaire voluntary would increase employees' satisfaction of their company because of the unenforced nature of the metric and the genuine interest about how the employees themselves would rate their performance.

An alternative, less ethically friendly metric would be something like the monitoring of employees' completed deliveries (or, in general, completed tasks) in a working day/week. Monitoring employees in this way (seemingly enforced and involuntary) would be an ethically ambiguous area since the AR glasses could accidentally monitor private conversations (both physical and digital) the employees have while, for example, on a lunch break, which would be a violation of the employees' privacy. Another ethically negative impact of constant monitoring could be on the employees' mental health, which could suffer because of the pressure constant monitoring puts on people's mental wellbeing.

3.

a) Two interaction challenges that touchscreen smartwatches suffer from because of their small screens is the "fat finger" problem (as for input) and not being very legible (as for output).

The "fat finger" problem occurs when the intended input targets, like icons or digital buttons, are smaller than the fingertip, which could cause users to miss their intended target, like a landmark or sight (an issue exacerbated by the smartwatch usually used in motion/varied contexts). This issue could be solved by including fewer elements on the screen, which would thus allow making the icons bigger, for example, by categorising landmarks and sights by location or type, in which case users could go down a hierarchy of related sights.

Legibility is also a big issue that the company needs to think about since smartwatches do not have a lot of screen to convey large amounts of information, which results in almost anything being very difficult to see and read from a smartwatch, especially if the user has any sight problems. A method to mitigate this problem would be to focus on very meaningful one- or two-element displays, like an arrow showing the way to the landmark a tourist has chosen to see.

b) One of the interaction techniques would be double tapping the touchscreen either on the right of the screen to accept a notification or the left to dismiss it. The other technique would be to gesture the hand over the smartwatch: away from the body to dismiss it and towards the body to accept a notification.

The double tapping interaction would need only touchscreen sensing in order to be implemented, which could be turned on or off based on the smartwatch's already implemented gyroscopic capabilities from the turning of the wrist. The pros of this technique would be that the interaction would be very hard to misinterpret since it is a very precise movement, and users are already familiar with touchscreen technology from use of smartphones, so there would not be a large learning curve, unlike the hand gesture interaction. The cons would be that this technique requires a very precise movement and does not have a large room for error, i.e., users could easily double tap the wrong region of the watch since the display is so small, that these interactions would always have to have the opposite hand free and unburdened to interact with a touchscreen.

The hand gesture interaction would require movement sensors on the smartwatch, which would be hard to choose in the resource use vs. simplicity trade-off since any depth, ultrasonic sensors or cameras would require a lot of resources and drain battery, but simpler sensing modalities, like infrared sensors. The pros of this technique would be that the gestures would be extremely easy for users to do, especially since they do not have to look at the screen if they are, for example, on the move/sightseeing and are not interested in any new notifications, that hand gesture movements are very translatable interaction metaphors (a swipe away for shooing something/someone away and a swipe towards oneself to invite them over), which would help users interact naturally with the app. The cons would be that interactions could easily be misconstrued by the watch, for example, when another person is moving their hands right next to them, that the new interactions would have a higher learning curve for users than more conventional techniques, and that it is not a very socially acceptable technique since hand gestures are not commonly for wearable technology and could be misconstrued by other people as interacting with them.

c) A visual approach could be to show an arrow on the display towards the intended location of a user walking towards a landmark. In this case, the arrow could be very easily seen on the watch and would require very little attention to the device itself when navigating to the landmark.

A non-visual approach would be with haptic feedback, i.e., if the watch vibrates once every time the user needs to turn right or twice for turning left. This would enable the user not to need to look at the watch at all after starting to navigate to the landmark, and it would be hard to miss these vibrations and miss out on a turn.

d) A group of around 10 users of varying ages and technological abilities will be given two days in a new city to test out the app while sightseeing. 5 of them will be given the visual approach (group A) and 5 of them the non-visual (group B).

They will be given a short presentation on how to use the watch and the app and then will have to learn on the spot on how to apply this new knowledge directly to see the sights around them. After the 2 days are up, interviews will be conducted of all of the participants, gaining their general sentiment towards the app, while also acquiring usability through Likert-scale questionnaires given at the end. The interview questions will ask how quickly they became familiar with the use of the app, any frustrations they experienced while using it and how much they would be willing to use this app in the future compared to traditional approaches to sightseeing with tour groups/physical maps/alternative smartphone apps. Comparing results from both groups, coding the interviews and aggregating results from the questionnaires, general sentiments on both navigation approaches and the more technical statistics from the questionnaires can be gained and contrasted to evaluate both approaches.