

# It's the Gesture That (re)Counts: Annotating While Running to Recall Affective Experience

Felwah Alqahtani

GEM Lab, Dalhousie University  
F1823899@dal.ca

Derek Reilly

GEM Lab, Dalhousie University  
reilly@cs.dal.ca

## ABSTRACT

We present results from a study exploring whether gestural annotations of felt emotion presented on a map-based visualization can support recall of affective experience during recreational runs. We compare gestural annotations with audio and video notes and a “mental note” baseline. In our study, 20 runners were asked to record their emotional state at regular intervals while running a familiar route. Each runner used one of the four methods to capture emotion over four separate runs. Five days after the last run, runners used an interactive map-based visualization to review and recall their running experiences. Results indicate that gestural annotation promoted recall of affective experience more effectively than the baseline condition, as measured by confidence in recall and detail provided. Gestural annotation was also comparable to video and audio annotation in terms of recollection confidence and detail. Audio annotation supported recall primarily through the runner’s spoken annotation, but sound in the background was sometimes used. Video annotation yielded the most detail, much directly related to visual cues in the video, however using video annotations required runners to stop during their runs. Given these results we propose that background logging of ambient sounds and video may supplement gestural annotation.

**Keywords:** Gesture annotation, emotional recall, running experience, remembering, visualization.

**Index Terms:** H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

## 1 INTRODUCTION

Recreational runners run for both fitness and pleasure. Not every run is as enjoyable as the next, however, and personal preferences for runners vary along many dimensions, including weather conditions, location, terrain, music, time of day, solo or group running. Our research is inspired by the idea that emotions play an important role in recreational physical activity such as jogging or running [3]. People who experience positive emotions such as enjoyment and happiness in recreational sport participate for longer [29]. Keeping track of one’s affective experience during runs, and correlating this with factors such as time of day, solo or group running, music, weather, etc., could give recreational runners better self-awareness regarding their running preferences—helping runners identify patterns that correlate with a positive running experience, enabling them to better choose the time, place, and circumstances of their leisure runs.

Commercial personal fitness informatics systems such as Runkeeper and Fitbit concentrate on performance metrics using passively captured data including heart rate, speed, distance, and elevation. Some provide the option to create a text annotation after a physical activity, allowing users to capture short reflections on the activity as a whole. Emotional state often varies during a run, however. Our research considers ways for runners to capture emotional state during runs, alongside physiological and performance data like heart rate and speed, and contextual data like route, weather, music playlist, etc.

Improving self-knowledge through logged emotional states would benefit from empathetic recall of those states—from the ability to relive the moments. In this way, an individual can reflect not only on how they were feeling, but importantly why they felt that way. Other researchers have considered the link between biometrics, affective experience, and recall in sport. Tholander and Nylander [28] interviewed 10 elite and recreational athletes who used a personal wearable sports technology. They found that combining biometric data with how one feels helps to analyze training sessions, and that personal recollections are routinely associated with data like heart rate and GPS.

In our work, we investigate how different techniques for capturing emotions while running affect recall of the affective experience of recreational runs, and how each technique is used alongside passively captured performance and contextual data during recollection. Specifically, we compare a gestural annotation technique—making simple gestures that represent emotional states on a touchscreen while running—against making audio or video annotations, in terms of amount of detail recalled, confidence in recall, and ease of use while running. In a between-subjects study, 20 runners were asked to record their emotional state at regular intervals while running a familiar route using one of the four methods, over four separate runs. Five days later, runners interacted with an interactive map-based visualization to review and recall their running experiences. In addition to the routes run, the visualization presented a set of cues that might support recall: weather, running speed, elevation, heart rate, and the location of each annotation.

Results indicate that gestural annotation promoted recall of affective experience more effectively than a baseline “mental note” condition, as measured by confidence in recall and detail provided. Gestural was also comparable to video and audio annotation in terms of time to recall, confidence, and detail, suggesting that gesture may be a good “low demand” technique for capturing effect during activities like running. Audio annotation supported recall primarily through the runner’s spoken annotation, but sound in the background (such as people, wind, fire engine sirens) was also used. Some participants expressed reservations about using audio annotations in public. Video annotation yielded the most detail, much directly related to visual cues in the video, however using video annotations required runners to stop during their runs. Given these results we suggest that background logging of ambient sounds and/or video may effectively supplement gestural annotation.

## 2 RELATED WORK

### 2.1 Lifelogging

Digital sensors and mobile applications make it possible to store a vast amount of personal information and data about everyday activities either passively (e.g., GPS) or actively (e.g., tracking food purchases). Much lifelogging and personal informatics research explores tools to capture and store information about daily activities and personal events, and their impacts on behavior in diverse areas including work performance, health, and lifestyle. Accordingly, lifelogging data is used for a wide range of purposes. It can be used to cue one's memory regarding the location of an item, people, places, among other things. It can be used to manage one's time, share everyday events or experiences with others, regulate emotions, improve health by monitoring and promoting healthy eating habits and physical activity [8].

Some lifelogging systems are designed to help people capture memorable personal experiences. SenseCam, one of the earliest digital lifelogging devices, was originally intended to support those with memory impairments [12] both as a day-to-day memory prosthesis and as a repository of meaningful life events. Other lifelogging tools have sought to support personal and shared memories for the broader public. Such systems grapple with a tension between the high volumes of data captured on the one hand, and the desire to filter and preserve only personally meaningful moments on the other. Living Memory Box was designed to store family memories by capturing audio and video of specific events [26]. MemoryLane allows people to capture digital memories about places, people, objects in their home, etc., and visually arrange these memories in various ways [16].

Lifelogging systems for physical activity typically rely on environmental and biometric sensors. For example, TripleBeat is a mobile app that helps runners achieve exercise aims by selecting music to encourage pace goals dynamically, using data from a heart rate monitor and a 3-axis accelerometer (to track movement) [24]. Houston is another app that encourages physical activity by capturing daily step counts and sharing this with friends [6]. Li et al. [17] presented contextual data alongside step counts. In a study participants became more aware of the factors influencing their physical activity, and suggested that capturing mood would also be useful. While such systems sometimes provide facilities to make notes or share reactions, data capture is largely passive.

Other lifelogging systems track data that is difficult to capture passively, and rely on active capture by the user. Many applications that collect emotion or affect fall under this category. For instance, in Mood Map [10], the user captures their mood by clicking the corresponding area on a two-dimensional colour coded map, and adding a brief annotation. Other systems combine such annotation with passive data capture. Emotion Map [14] enables users to record emotions with location, activity information and time, and then displays the emotions on a map. Matassa and Rapp [20] present a mobile application that captures time, position, and weather alongside user annotations such as their feelings about a specific event, a picture, or video.

Gesture holds promise as a mechanism for active capture while mobile. In a lab experiment [5] Bin Hannan et al. showed that people can express the intensity of emotion using attributes of a gesture (e.g. pressure, length) in a consistent way, and they can recognize intensity when the gestures are played back. Other work has also illustrated that the manner continuous gestures are made on a mobile phone can encode additional information, such as placing emphasis on words in a text message [1].

### 2.2 Visualization for Recall

Personal visualizations let people explore and gain insight into their own data. Personal visualization tools support different

purposes including satisfying curiosity, achieving a personal goal, and sharing with friends [27]. Such tools are also useful for recalling past experience (e.g., special events [21] and trips [27]), and for gaining insight into one's own habits through reflection on past experience (e.g., computer [11] and mobile device [23] usage). Personal visualization can also represent both internal phenomena like goals, skills, and experiences, and a variety of phenomena external to the individual, from the physical environment to social impact [13]. These visualizations use data obtained from a range of sources and cover many areas of an individual's life such as health, behavioural change, and social life.

Visualization systems to support recall and reminiscence have been explored for many aspects of our life. For example, McDuff et al. [22] developed AffectAura to log audio, visual, physiological and contextual data related to users' workday activities (such as desktop activities, meetings, and locations) and predict users' affective states using a classification scheme. When presenting the data using AffectAura, participants were able to reconstruct a 'story' about their workday activities. Hailpern et al. [11] presented YouPivot, a contextual aide to support retrieval of browsing history. The tool logs computer state (open files, music, physical location) and a visualization interface displays the user's activities. In a comparative study, YouPivot was preferred and provided faster retrieval than traditional browser history.

Mathur et al. [21] designed a visualization tool (LifeView) that can convert lifelog data into a cartoon strip and make contextual hyperlinks between the various events in the lifelogs. The study shows that the cartoon-ized interface provided better recall of affective experience than text. Kalnikaitė et al. [15] compared three visualizations for recall of life events: Snap (visual images), Track (location data) and SnapsTrack (a combination), finding that Snap generate recall of more detail than Track, while SnapsTrack promoted inferences about an individual's habitual patterns. Sellen et al. [25] found that SenseCam images assisted recollection and event reconstruction in the absence of recollection, and that this was not affected by the passage of time. Selfies triggered at transitional points in mobile app usage (screen unlock, app launch, etc.) and explicitly intended to capture emotion through facial expression can also be effective cues supporting recollection of how one was feeling [23]. A range of personal visualization systems allow reflection on physical activity. Commercial platforms like Runkeeper, Strava, and Fitbit tend to provide environmental and physiological indicators in a map-based dashboard visualization. Researchers have explored the benefits of alternative visualizations, including cartoons [2], gardens [7], and fish tanks [18], providing evidence that the visualization used can impact the amount and nature of personal reflection that occurs. In our work, we consider ways in which annotation of emotional state can be integrated into map-based visualizations used by runners.

## 3 STUDY

The primary goal of our study was to explore differences in recall of affective experience by recreational runners when different annotation techniques are used (specifically gesture, audio, video, and mental notes). A secondary goal was to gather feedback about the suitability of each technique for annotating while running.

We employed a between-subjects design. While this increased the potential impact of individual differences due to sample size, it allowed us to see how participants used a method over a series of runs and whether they could detect patterns in affective experience across runs using the visualization interface corresponding to their capture technique. These are critical aspects to consider when exploring the relationship between

annotation and recall. The study was divided into two phases as follows:

**Capture phase:** Each participant ran 4 times in 10 days, capturing their emotional state using the method assigned to them. The participants were assigned to the conditions randomly.

**Recall phase:** Five days after the last run, runners interacted with an interactive map-based visualization to review and recall their running experiences.

### 3.1 Capture Apparatus

All participants used the same Samsung Galaxy smartphone and Mio heart rate wristband. The smartphone captured running data in the background using the Runkeeper app. All participants held the phone in hand while running. We explored headworn camera, armband, and watch options, but settled on this configuration as it is not uncommon for recreational runners to hold their phones while running. While we acknowledge that hands-free techniques are possible for audio, video, and even gestural annotation, the common form factor used in each condition facilitated comparison of participant feedback about the annotation techniques.

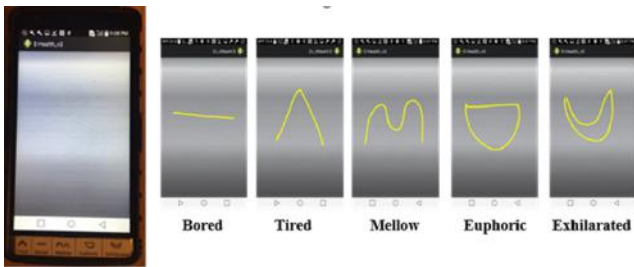


Figure 1: JogChalker app with Gesture/Emotion set used

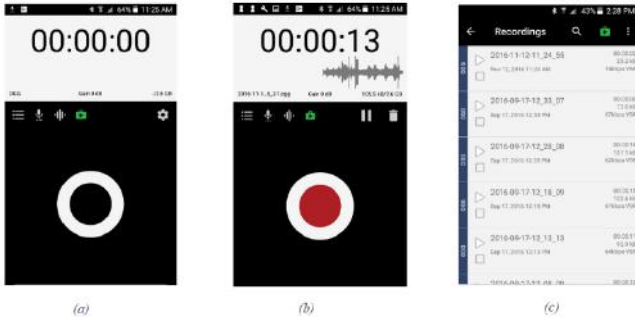


Figure 2: ASR audio recorder, (a) before recording, (b) during recording, (c) timestamp with audio recordings

**Gestural annotation:** the JogChalker capture application was used for gestural annotation [4]. It is written in Java for the Android platform, and provides a simple touchscreen interface for making gestures. Gesture data (touch location/trajectory, width, pressure) are encoded into SVG format. The app also passively captures GPS coordinates and time (Figure 1). Two researchers from our lab tested the capture application over several weeks, making their own custom annotations while running. Based on their recommendations we identified a set of emotional states and designed simple corresponding gestures. These emotional states

and their gestures were then considered by a group of recreational runners in a series of participatory design sessions [4], leading to five running-related emotional states (bored, tired, mellow, euphoric, exhilarated) (Figure 1), recorded using simple gestures. While we don't suggest that these are the most important emotional states to capture while running, they were deemed in this preliminary work to be a reasonable set for this study. The gestures were listed on the bottom of the phone for easy reference. Recorded gestures were timestamped and geocoded.

**Audio annotation:** We used ASR, a voice and audio recorder app to record participants describing their affective experience of the run at prescribed points (Figure 2). ASR provides a large toggle button, making audio recording straightforward while running. Recorded audio is timestamped.

**Video annotation:** We used a video recorder feature in the JogChalker capture application for video annotation (Figure 3). A large toggle button starts/stops video recording. Resulting video annotations were timestamped and geocoded.

**Data merging:** a script combined the Runkeeper output with the annotation files. The resulting file was used by the visualization interface to present running data.

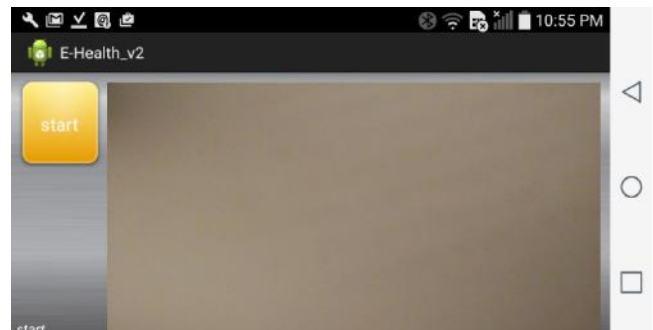


Figure 3: The video recorder in JogChalker. The button toggles recording, and a preview window appears on the right.

### 3.2 Visualization Interface

Building on our participatory design sessions with recreational runners [4], we developed four variations of visualization interfaces for the study (one per annotation condition) using Mapbox Studio [19]. Each interface has five tabs: four presenting the captured data for each run individually, and a “report” tab presenting all the data combined. Along the route teardrop markers indicate annotation locations; clicking these causes corresponding emotion annotation to be displayed alongside running data (speed, heart rate, and elevation). The day and weather are always visible in a small box in the upper right corner of the interface. The gesture visualization presents the gesture annotation of their affective experiences as an animation. This replayed the gesture at the same rate as it was recorded. Stroke width matched touch area and pen colour was matched to gesture pressure. This was done as related work indicates that the way a gesture is made can express emotional intensity [5]. The audio visualization plays the audio annotation while displaying running data. The video visualization plays the video annotation alongside the running data. For the mental note condition, only the running data is displayed (Figure 4).



Figure 4: Route annotation popups: (A) heart rate, (B) speed, (C) elevation, (D) gestural annotation, (E) video annotation, (F) audio annotation. The leftmost popup was shown for the “mental note” condition and the remaining for their respective conditions.

Finally, clicking the Report tab displays all the data together (Figure 5). The left side of the interface shows the routes chosen by the participant, displayed in different colours to distinguish them. We also displayed average speed for the four runs, average heart rate for the four runs, total distance and differing emotions during the four runs.

### 3.3 Participants

Our study population was drawn broadly from the Dalhousie community. We recruited 20 runners for the study (10 males, 10 females, aged 18-64, (mean 30)). Thirteen participants had run recreationally for more than 3 years; five had 1-3 years of running experience, and two had run for less than one year. One participant regularly ran >10k per run, eight ran 5-10k, nine 2-5k, and one <2k. Because some runners listen to music while others do not, and since music can directly impact emotional state, due to the time required to run each participant we decided to keep our sample size manageable by screening for runners who never or rarely listened to music while running. 7 of 20 participants had at least some prior experience using personal informatics apps/devices for running such as Runkeeper, Strava, Fitbit, or Garmin connect. Since capturing emotional state during recreational runs may appeal to runners who are less performance-oriented, we wanted to collect feedback from runners who use personal analytics for performance and personal goals and those who do not. Participants received \$30 for taking part in the study.

### 3.4 Procedure

At the beginning of the study, the study was explained and participants underwent a training session on how to express emotions using their assigned technique: using the JogChalker application for gesture, ASR for audio, JogChalker for video, or making a mental note (in this condition they were simply prompted to think about their emotional state without recording it and without taking notes at the end of the run). Participants were told that they could stop or slow down when annotating if they wished. In the audio and video conditions participants were asked to record how they were feeling in a few words; participants using video were also asked to capture the surrounding environment. The study was conducted for two weeks and it was divided into two phases: capture and visualization.

In the capture phase, each runner used one method to capture emotion over four runs. Participants went out for a run four times in 10 days for approximately 30 minutes on any route they chose, but from the same starting/ending point (to facilitate equipment dropoff). Participants annotated at prescribed times, spaced in 5-

minute intervals, giving 6 annotations per run. The device vibrated when it was time to create an annotation. We did this to avoid wide variation in annotation frequency between participants, something we observed in the preliminary design work. Prompting for annotation follows the Experience Sampling method [4], asking participants to reflect on their current emotional state rather than capturing annotations when the participant feels compelled to do so. We felt this was important as different participants may be more compelled to express different emotional states, potentially impacting comparison of annotation conditions, and also because we wanted to sample affective experience over the duration of each run.

We met the participants before each run to set up all devices and apps, and met them again after each run to collect the devices. After finishing all runs, participants had a five-day break before the recall phase. We waited five days to mitigate a potential recency effect on recollection, and to simulate a review of several runs completed in the recent past.

In the recall phase, participants interacted with the visualization interface corresponding to their annotation condition. Each run is presented in a separate tab, and a synthesis of all runs is presented in another tab. After participants finished familiarizing themselves with the interface, we asked the participants to describe in as much detail as possible how they felt at each annotation point and why they felt that way, starting with the first run and moving to the fourth. For each run, participants selected each annotation in sequence and attempted to give detailed descriptions of their affective experience at that moment. Specifically, they were asked to explain how they were feeling and why they felt that way

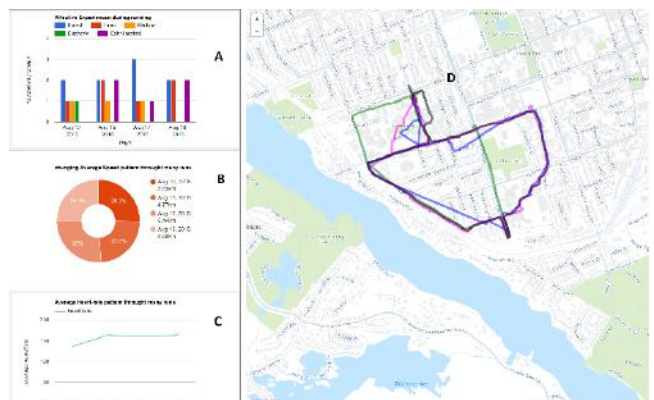


Figure 5: Report interface: (A) emotional state over 4 runs, (B) speed over 4 runs, (C) heart rate over 4 runs, (D) routes run.

(participants who used gesture annotation were able to remember the meaning of each gesture). After each recollection, we asked participants what data they used to help them remember and how confident they were about their recollection on a scale from 1 (not at all confident) to 5 (very confident). Each participant provided 24 recollections (6 annotations X 4 runs). After they finished the recall phase, they opened the report/summary tab. They were then asked if they could detect any patterns in their affective experiences across runs, and if so to describe them. Participants completed a post-study questionnaire that contained two questions to see whether capturing emotions during their runs enhances their recall of their running experience and which cues in the interface helped to trigger recall of their running experience. Semi-structured interview then was conducted in which they answered 11 questions related to aspects of their annotation condition that did and did not help them recall their affective experience, how they used the annotation technique, and questions eliciting feedback about the technique. For example, what challenges did you encounter while you were recording your emotions?

## 4 ANALYSIS

In addition to qualitative analysis of participant recollections and of feedback given in the questionnaire and interview, we considered the time taken to recall running experience associated with each annotation, the amount of detail recalled for each annotation, and a self-reported measure of confidence in the recollection provided for each annotation. We calculated number of details for each recollection as a rough quantitative measure for comparison between conditions. Participants were given one point for each emotion described, and one point for each reason provided. For example, 'I was feeling tired, because I was running uphill and the weather was hot' would be assigned 3 points. We were conservative with points allocation: points were not awarded for commentary not directly tied to affective experience (e.g. 'I just started running at this street' would be assigned 0 points). Our method of calculating the number of details given after probing recall is similar to that used in the lifelogging and memory research conducted by Kalnikaite et al. [15].

Time to recall was calculated as the time from the point the annotation stopped playing (be it SVG animation, audio, or video) to the end of the participant's verbal explanation, including any time spent thinking. The participants indicated their confidence for each recollection on a scale from 1-5, where 1 is not confident at all, and 5 is very confident.

## 5 RESULTS

In this section, we first present results from a quantitative analysis of time taken, detail, and confidence in recall. After this we present results from a qualitative analysis of recollection behaviour, and data from the questionnaire and interview. In this section we refer to our participants by condition+ID, where V=video, A=audio, G=gesture, and M=mental note.

We used a multivariate ANOVA in SPSS to analyze the recollection data. There were three dependent variables (DVs): Number of Details (Details), Time to Recall (Time) and Confidence, with annotation type as the independent variable (Gesture, Audio, Video, Mental Note). The main analysis for the three DV indicated that there was a significant difference between the groups with  $F(2,48) = 2.241$  ( $p < .035$ ,  $\eta^2 = .296$ ). This indicates that the best linear combination of the three DV is different as a function of group. So, each DV was analyzed separately using a simple BS-ANOVA, followed by Fisher's Least Significant Difference post-hoc test, because we had specific predictions about how the annotation conditions would impact performance.

For number of details, the conditions were not significantly different, with  $F(3,16) = 1.933$  ( $p < .165$ ). Post hoc analysis indicated that there was a significant difference in Detail between Video and Mental Note ( $p < .029$ ), with more detail in Video on average, but no difference between any other two conditions.

For time to recall, the conditions were significantly different with  $F(3,16) = 1.933$  ( $p < .014$ ,  $\eta^2 = .476$ ). Post hoc analysis indicates that there was a significant difference between Gesture and Mental Note ( $p < .012$ ) and between Video and Mental Note ( $p < .009$ ), with Mental Note taking less time on average.

For confidence, the conditions were significantly different with  $F(3,16) = 1.933$  ( $p < .015$ ,  $\eta^2 = .469$ ). As might be expected, confidence was generally lowest in the Mental Note condition. Post hoc analysis shows a significant difference between Gesture and Mental Note ( $p < .004$ ), Audio and Mental Note ( $p < .036$ ), and Video and Mental Note ( $p < .002$ ).

### 5.1 Details Recalled

#### 5.1.1 Use of Annotation

In the Video condition, details recalled were often closely tied to what was visually evident in the video recording. For example, VP1 noted "At that point I had been on the flat street for a while so I was not feeling tired any more...". VP3 said "I was feeling good. It is easy to run around soccer field. I do not have to look for a car or think of what is happening". VP4 stated "I can see I went down. I was tired but better because I was going down so I felt better". Building on these visual elements, participants' descriptions of affective experience provided a broader narrative than the emotional states participants noted in the video annotation. For example, VP4 reported in the video annotation that he was feeling "very heavy", but in his recollection, he stated that "I was getting better ... because I was in the flat (sic) and I was running slow and the elevation did not change." VP1, after reviewing a video annotation commented that it "is interesting to report mellowness I know I was going uphill so I was working harder at that point but it might be the nice neighborhood and the little park that contributed to the mellowness." If video was blurry or jostled, participants relied on the audio description and other cues like annotation location. For example, VP2 recorded video while running such that the video itself was angled toward the road. However, his route was very familiar to him and he recalled his emotions in detail based on annotation location and his statements in the audio.

In the Audio condition, participants' recollections largely mirrored the detail provided in the audio annotations, provided that those annotations were sufficiently detailed. For example, AP3 reported in the audio recording he was thirsty because he did not drink water before running and after listing to the audio recording he stated that "I did not have water before running and after I started running, I was feeling thirsty". When annotations provided little detail, participants relied on memory or other cues (location, weather, time) to explain how they were feeling. For example, AP4 recorded her emotion briefly in all runs. She could remember the emotions that related to aspects of the physical landscape that she knew could affect her emotion, or if an event occurred such as bumping into a friend; otherwise she could not recall the emotion.

In the Gesture condition, recollections often included detail not evident in the gesture itself. This and the time taken for recollection suggest that participants really thought about their affective experience and that the gesture prompted their memory. For example, GP1 stated that "I was feeling annoyed because there were some kids on my way. There should be a [gesture] for annoyed." GP4 said, "I felt euphoric because I saw about five people who were swimming and the sun was in the west and the

sea was very beautiful so I felt happy at that time.” GP5 noted, “I was not feeling very well at that day and my legs were sore.” GP3 explained “I think I was more expos[ed] to the sun and there was a lot of people. I was tired because of the crowd and the heat. I can picture myself running around the people so I was tired.”

In the Mental Note condition participants provided very few details about their emotions, tending to focus instead on their running progress and performance. Participants described their general feelings at the start and end of each run, but they encountered difficulty remembering specific emotions at the prescribed points. For example, MP2 stated that “I just started walking because I cannot run more than 20 mins continually but after my run I have beautiful feeling I like that feeling.” MP3 noted that “When I first started I felt stronger than I had for other one. and I remember this one was a harder than the day before and I think because the humidity oh that is right that was on my Saturday morning....”.

### 5.1.2 Detecting Patterns

In all conditions except Mental Note condition, participants could detect patterns in affective experience across runs using the visualization interface. Some saw patterns along the duration of each run. For example, GP4 noted that “...it looks [like] I started bored... then tired, [and] when I am done I was excited”. 15 of 20 participants expressed that their emotions changed based on one or more contextual aspects such as weather, time, location, and physical landscape or surroundings. For example VP3 stated that “when I [annotated] either ‘bored’ or ‘tired’ I seem to be either just finished uphill or [beginning] uphill”. For some, this provided insight into differences between runs. For example, AP2 stated that “...on a muggy day I was feeling pretty low whereas in the sunny or cloudy day I was feeling better”. Others observed that their emotions stayed mostly stable across runs. However, in the Mental Note condition participants could not detect pattern along the duration of each run but they could remember how they felt in the beginning or at the end of their runs. for instance, MP3 stated that “by the time I got to points 5 and 6 I was starting to feel mixed of happy and relieved because I know that I was close to what I need to be. I guess as a very new runner I feel happy when I finish it” and MP4 noted that “typically I started with happy and then by the end of it I would mixed of happy and kind of bored to finish”.

### 5.1.3 Use of Passively Captured Data

Video condition participants did not use other data such as speed, elevation, weather, and heart rate to help them during recall. The plotted location was helpful for them to contextualize the video clips, however.

In the Audio condition, participants generally did not use passively captured data beyond location. AP3 and AP5 did look at heart rate to see how it might be related to a recorded emotion. AP4 used weather and the date of the run to explain some emotions.

All participants in the Gesture condition used location to contextualize each annotation, and 4/5 used weather and time information at least once in their descriptions. Participants did not use other data, except for GP5, who connected his emotion with his speed, stating “when I ran fast I knew I was happy but when I ran slower I knew I was tired”.

Three participants in the Mental Note condition used the location and/or weather when recalling their emotions, but provided few details and expressed low confidence in their recollection. Participants did not use other data for recall of affective experience. MP2 used speed to provide more information about her performance only.

## 5.2 Time to Recall

Recollection for participants in the Video and Audio conditions did not generally involve long pauses or explicit efforts to remember. Instead, the video or audio recording was played and participants reflected on what was stated or shown. Video and audio annotations were generally played once, unless there was an issue with audio quality.

In the Mental Note condition some participants did take a little time to think, but generally provided short observations of their running progress or performance. Any recollections of affective experience were high level descriptions of how they felt before and after the run.

Recollection was a more effortful process for participants in the Gesture condition overall. Unexpected emotions tended to generate particularly thoughtful recall. For example, GP3 usually did not feel good at the beginning of her run, so she took time to recollect her experience, saying “Ooh now I can remember, I remember when I started off, I was surprised that I was feeling as good as I was because I thought I should be tired because of a long run in Prince Edward Island but I actually felt ok so that is why I was mellow, yes I do remember that.” GP1 normally feels good at the beginning of her run, so she took time to consider why she had made a “bored” gesture in one instance, then recalled “I did not sleep very well the day before so I am just feeling down I am not really bored in the sense of being bored. I am more depressed”. At other times participants could not recollect. For example, GP2 stated that “I usually start my run bored. I do not really remember this one”. GP5 noted, “I was sick at that day. I do not know why I was happy after I was tired”.

When Gesture condition participants recorded the same gesture annotation consecutively, they generally took time to think about and discuss the first instance, and spent less time thinking about and discussing subsequent instances, instead indicating that they felt the same way. This contrasts with the Audio and Video condition participants, who consecutively played each annotation and provided an explanation for each.

## 5.3 Recall and the Act of Annotation

All Video condition participants stated that video annotation helped them to recall many aspects of their running experience. VP1 stated “Capturing the video particularly helped me remember my emotional states, the thoughts I had while recording, and the particular details of the day (weather, sleep the night before, other factors contributing to mood or exhaustion).” VP3 said “I think I was able to give more details about my particular feelings and experiences during the run than I would have been able to if I had just noted, for example, ‘ran 4k yesterday’”.

Most (4 of 5) Audio condition participants felt that audio annotation helped them to recall their running experience. Despite largely mirroring audio description details in their recollections, participants also felt the act of annotation serves as a marker for recall, and the recording itself (the sound of one’s voice, ambient noise) serves to trigger memory. AP1 stated “... I think that actually having the recording of my voice triggers those memories, however just the act of recording creates an experience to base memories around. So even if I did not get to hear the recordings again, I have a vivid memory of doing the recording at certain places during the run because I did an activity that was outside the norm of my normal run. I remember making the recording itself.” AP2 said “My voice reminds me of how exhausted, how tired I was, etc., during that moment in time.”, while AP5 noted “...it helps me remember the environment I was in... For example, the car who never stopped for me in my third run”.

4 of 5 Gesture condition participants indicated that gestural annotations supported recall of their running experience.



Participants cited the act of annotation itself as an event that could be remembered, leading to recall of more detail about the run at that moment. GP3 stated “It helped with the internal visualization of where I was in the city at the moment that the emotion was recorded. It allowed me to picture in my mind where I was at that moment, and allowed me to recall how I felt while I looked at whatever scenery was around me at that time.” GP2 said “I think recording these instances makes each individual moment more unique. I think memories need to stick out in some way for them to be easily recalled. I don't find running or exercise particularly exciting or fun and if I wasn't recording my emotions at the time, or tracking my location using a GPS I would not be able to recall any specifics about the run other than how I feel in general about running or very generic details like my route or an extreme temperature/weather.”

Mental Note condition participants felt that the reminders to make a mental note while running served to increase mindfulness and to remember details about their run, but not necessarily their affective experience. MP2 stated “It kind of keeps me alert at the time of running and my mind remains active and remembers what I was thinking at that moment.” MP1 said “I remember where I was when I thought about certain things. For example, when I was close to my work place I started to think about work”.

## 5.4 Annotation Techniques

### 5.4.1 Participant Feedback

Three Video condition participants stopped running to make a video annotation. Of these, two felt it wasn't possible to use the interface without stopping, while the third wanted to ensure a good quality image. Another participant stopped in the first run but ran while annotating after that because he felt stopping distracted from his run. All participants felt the video aided in recall.

All Audio condition participants felt audio annotations were easy to make while running. However, three participants felt embarrassed or self-conscious when recording their emotion around other people. For example, AP1 noted that “I did not record my emotion immediately [when notified] because there was a person coming at me ... so I said wait until we passed each other so I do not look crazy. Participants also stated that they felt challenged during recall if the annotation was short or if there was a lot of environmental noise.

All Gesture condition participants stated that that making gesture annotations was easy and did not interfere with their run. However, three participants said that they needed to look at the interface to make sure the gesture was drawn properly, which sometimes made them slow down. Two participants indicated that the annotations, by only recording an emotion, made it challenging to differentiate specific instances. As GP1 stated, “the good thing [is] the application is easy to use and fast but it does not capture what's actually going on”.

Three participants mentioned that they did not normally have the phone in their hands but they got used to it in our study whereas others did not mention any difficulty of holding the phone while their running. All participants liked the idea of revisiting how they were feeling and why, and enjoyed recalling the various factors that affect their running experience. For example, GP1 stated that “I would use this interface if it is integrated with what I have used”. Providing the emotions in the visualization made participants aware about what positively and negatively affected their run and would enhance the way they run. In addition, the visualization helped all participants regardless of the condition to reflect on their enjoyment during the run and what can influence their enjoyment and performance during runs. Moreover, most participants (17 out of 20) felt that using the

visualization long term might help them make better choices and have more fulfilling runs. For example, AP5 stated that “presenting the emotion with heart rate and speed on the interface can help me to make better choice about the run if I use it for long term”.

### 5.4.2 Use of Techniques

One key difference between the Gestural annotation condition and the Audio and Video annotation conditions was the consistency with which the interfaces were used. Due to the constrained nature of the gestural annotation, participants in this condition captured their emotional states with high consistency across and within participants. Given the freeform nature of the audio and video annotation interfaces, we see some variation in annotation behaviour and content in these conditions, however in general participants recorded *how* they were feeling and often also *why*.

Three Audio condition participants recorded quite detailed annotations. For instance, AP5 said “still running... feeling very tired and my legs are cramping”. AP2 stated, “I feel a bit self-conscious running in a crowded street and a bit scared because there is [a] fire engine outside the medical building...”. AP3 noted, “I am on Quinpool street and it is cloudy but I am still bored”. AP4 recorded very brief annotations, for example “happy but tired”. AP1's annotations were mixed in terms of detail, varying from “I feel a little better but still sad” to “I am feeling happy and the cramp is gone. I do not feel tired but I feel I am running kind of slowly”.

All Video condition participants recorded verbal annotations along with the video. For example, VP1 recorded “...South Street: I feel very good, very excited and a little mellowness. I am a little tired. It is very humid today. It is getting hard to breathe but it is nice, clear and sunny”. VP2: “I am running down the hill feeling very good and running is nice and cool day today”. VP3 stated “I am feeling good today. Some clouds came by so it is not hot right now.

Wishing I [would] run a little fast[er] but that is not unusual”. VP4 reported that “I am feeling a little heavy, maybe I am running [faster] than my capacity to complete half an hour”. VP5 recorded “I am not tired right now. I think I can run more but I feel thirsty, my mouth is dry, but I feel good, it is very sunny...”.

## 6 DISCUSSION

Using the visualization interfaces made our participant runners aware about what positively and negatively affected their running experience. Over time, such an interface could enhance the way they run, and lead to better choices about when and where to run. Participants in the participatory design sessions and the controlled study liked the idea of revisiting how they were feeling and why, and enjoyed recalling the various factors that affected their running experience over a span of time.

Analysis of collected data shows that gestural annotations prompted runners to remember details of captured emotion that were not evident in the gesture itself. The gesture animations triggered recollection, especially when the emotion expressed in the gesture was out of the ordinary. Gestural annotations seemed to encourage thinking, about when they made that gesture and why they felt that way, while the act of making the gesture itself acted as a marker around which to frame recall. Confidence in the details recalled for gestural annotation was also very high (as high as video). This came at a cost, however: the average time taken to recall details was higher with gesture annotations compared to mental notes, and while the time was comparable to that for audio and video, in those conditions participants spent more time playing the annotation vs. trying to remember.

Participants in the Gesture condition were restricted to five running related emotions (tired, mellow, euphoric, exhilarated,

and bored). This sometimes did not allow them to capture what they felt, and so they chose the closest emotion, which may have made it more difficult to recollect affective experience. It is possible that more gestures may promote greater differentiation and better recall, however at the expense of needing to learn a larger gesture vocabulary.

Unlike audio and gesture, video annotations allowed participants to refer to visually evident details and discuss why and in what way they affected their running experience. These results agreed with Eldridge et al. [9], which indicate that video has valuable cues that increase recall. Our participants used visual cues not only to explain why they were feeling the way they stated in the video, but to connect that emotion to a larger narrative that included how they were feeling leading up to, or after, the moment of annotation. This could sometimes diverge from the stated emotion (e.g., I stated I was tired, but I was starting to feel better).

Runners who made a mental note encountered difficulty remembering their emotional experience because there was no sign of the emotion type on the map. Even if some participants could remember or infer how they felt based on the location, their confidence in recall was very low. In addition, runners tended to remember their running behaviour more than their emotional state. For example, MP5 noted that “I was thinking about my physical body what is going on while I am running, while, MP1 stated “I decided that since I was running slower than normal I should take that route instead”.

In terms of capturing emotion while running, touch-based gesture is convenient since it does not interfere with a run like video recording can, and it is more private than audio recording. With repeated use, it is likely that the need to look at the screen while performing a gesture would be reduced. Moreover, Bin Hannan et al.[5] provided an evidence that one might imbue more than just a discrete state selection when making these gestures.

Annotation modality impacted annotation consistency: participants who used gestural annotation were consistent by design, whereas participants who used video or audio annotation were less consistent due to the freeform nature of capture.

All runners in each of the four condition groups suggested some improvements in the visualization to make it more effective for emotional recall. These include changing the view of the map to a satellite view or topographical map, and showing the passively captured data like speed, heart rate, and elevation as continuous values in a chart next to each run. Participants also suggested hybrid capture approaches, using video or snapshots to supplement gestural or audio annotations.

## 6.1 Limitations

Perhaps the most notable limitation of this study is sample size: with five participants in each condition, there is a potential impact of individual differences on results. While we partially address this through mixed quantitative and qualitative analysis, a study with more runners would help solidify our findings.

Despite positive indications given by the quality of recollections and stated confidence in the Gesture condition, we could not objectively assess the accuracy of recall from gestural annotation in our study. In future work, we may ask participants to explain their reason for each gesture annotation after making it, and use this to assess a subsequent interpretation of the annotation.

Participants in the gestural condition were given a distinct set of emotional states to select from, while in the other conditions participants were free to express or make note of their emotional state in a less constrained manner, and were not made aware of the discrete states that gesture participants were limited to. While we acknowledge this as a form of confound in that we do not

isolate gesture vs. the other modalities without also changing the quality of annotation, this is a deliberate choice for a naturalistic comparison. Runners are unlikely to annotate with single words given video or audio. Additionally, the set of emotional states represented by gestures in our study was chosen through iterative tests to be easily memorized and parsimonious—representative of common emotional states that were indeed expressed many times by our participants in other conditions. An interesting area for future work would be to consider using a gestural language to communicate affective experience in more nuanced and/or descriptive ways.

Every method used was restricted by the features of the input device. While this gave a certain consistency in experience across conditions, each modality could be more optimally supported. A phone is less than ideal for video recording during activities such as running, for example. Future work that is primarily concerned with the process of mobile annotation itself might compare more optimal configurations for each annotation technique.

Given the limited timeframe of our study, we could only explore how participants see and use patterns in annotation of emotion in a rudimentary way. While we hypothesize that gesture has advantages over audio and video when visualizing trends, more research is required to determine how best to accomplish this. A longer field study would help determine the utility of visualizing such trends, and to explore whether passively captured data becomes more relevant for understanding emotion over the long term.

## 7 CONCLUSION

We presented the results of a comparative study exploring whether gestural annotations of felt emotion presented on a map-based visualization can support recall of one’s affective experience during recreational runs. We compared gestural annotations with audio and video notes and a “mental note” baseline for this purpose. We also explored how passive captured data cues like weather, time, speed, elevation and heart rate could be used to supplement annotation. Gestural annotation promoted recall of affective experience more effectively than the baseline condition, as measured by confidence in recall and detail provided. Gestural annotation was also comparable to video and audio annotation in terms of time, confidence and detail. Participants in the Gesture condition expressed that the act of annotation was itself an anchor for recall. While participants required time to think about gestural annotations, their recollections involved detail not embedded in the gesture itself. Audio annotation supported recall primarily through the runner’s spoken annotation, but background noise was sometimes used. Participants felt self-conscious annotating how they were feeling in public, however. Video annotation yielded the most detail, much directly related to visual cues in the video, however using video annotations required runners to stop during their runs. The used of passively captured data (weather, time, location, heart rate, elevation, and speed) varied between runners, with location being an important cue for many annotations, and weather information an indicator that helped some participants explain how they felt.

## ACKNOWLEDGMENTS

We thank the reviewers for their insights and recommendations; the paper is much improved as a result. We also gratefully acknowledge support from King Khalid University and the Saudi Arabian Cultural Bureau in Canada.



## REFERENCES

- [1] J. Alvina, J. Malloch, and W.E. Mackay, "Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices", Proceedings of UIST 2016, 583–593, 2016.
- [2] L. Barkhuus, J. Maitland, I. Anderson, S. Sherwood, M. Hall, and M. Chalmers, "Shakra: Sharing and motivating awareness of everyday activity". In Proceedings of Ubicomp, 2006.
- [3] S. J. Biddle, "Emotion, mood and physical activity". Physical activity and psychological well-being, 63,2000..
- [4] N. Bin Hannan, F. Alqahtani and D. Reilly, "JogChalking: Capturing and Visualizing Affective Experience for Recreational Runners". In Proceedings of the 2016 ACM Conference Companion Publication on Designing Interactive Systems (pp. 129-132). ACM. 2016.
- [5] N. Bin Hannan, K. Tearo, J. Malloch and D. Reilly, "Once More, With Feeling: "Expressing Emotional Intensity in Touchscreen Gestures", Limassol, Cyprus, In IUI 2017.
- [6] S. Consolvo, K. Everitt, I. Smith, and J.A. Landay, "Design Requirements for Technologies That Encourage Physical Activity". Proceedings of CHI 2006, 457–466,2006,
- [7] S. Consolvo, D.W. McDonald, T. Toscos, M.Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. Smith, and J.A. Landay, "Activity Sensing in the Wild: A Field Trial of Ubifit Garden". Proceedings of CHI 2008, 1797–1806, 2008.
- [8] M. Czerwinski, D.W. Gage, J. Gemmell, C.C. Marshall, M.A. Pérez-Quñones, M.M. Skeels, and T. Catarci, "Digital Memories in an Era of Ubiquitous Computing and Abundant Storage". Commun. ACM. 49,44–50, (Jan. 2006).
- [9] M. Eldridge, M. Lamming, and M. Flynn, "Does a video diary help recall?", People and Computers, 257-257, (1992).
- [10] A. Fessl, V. Rivera-Pelayo, V. Pammer, and S. Braun, "Mood Tracking in Virtual Meetings". 21st Century Learning for 21st Century Skills (Sep. 2012), 377–382, 2012.
- [11] J. Hailpern, N. Jitkoff, A. Warr, K. Karahalios, R. Sesek, and N. Shkrob, "YouPivot: Improving Recall with Contextual Search", Proceedings of CHI 2011, 1521–1530, 2011.
- [12] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, "SenseCam: A Retrospective Memory Aid". Proceedings of Ubicomp 2006, 177–193, 2006.
- [13] D. Huang, M. Tory, B.A. Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury, "Personal Visualization and Personal Visual Analytics". IEEE TVCG. 21, 3 (Mar. 2015), 420–433. 2015.
- [14] Y. Huang, Y. Tang, and Y. Wang, "Emotion Map: A Location-based Mobile Social System for Improving Emotion Awareness and Regulation". Proceedings of CSCW 2015, 130–142, 2015.
- [15] V. Kalnikaite, A. Sellen, S. Whittaker and D. Kirk, "Now Let Me See Where I Was: Understanding How Lifelogs Mediate Memory". Proceedings of CHI 2010, 2045–2054,2010.
- [16] V. Kalnikaitė, and S. Whittaker, "A saunter down memory lane: Digital reflection on personal mementos", International Journal of Human-Computer Studies. 69,298–310, 5 (May 2011).
- [17] I. Li, A.K. Dey and J. Forlizzi, "Using Context to Reveal Factors That Affect Physical Activity". ACM Trans. Comput.-Hum. Interact. 19, 7:1–7:21, 1 (May 2012).
- [18] J.J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H.B. Strub, "Fish'N'Steps: Encouraging Physical Activity with an Interactive Computer Game". Proceedings of Ubicomp 2006, 261–278, 2006.
- [19] Mapbox Studio: /mapbox-studio/. Accessed: 2017-02-01.
- [20] A. Matassa, and A. Rapp, "Map: Making: Designing a Mobile Application for Enhancing Memories' Retrieval". Proceedings of MobileHCI 2015, 994–1001, 2015.
- [21] A. Mathur, A. Majumder, S. Datta, S. Menon, S. Malhotra, and A. Dahiya, "LifeView: A Lifelog Visualization Tool for Supporting Sentimental Recall and Sharing". Proceedings of OzCHI 2012, 371–380, 2012.
- [22] McDuff, D., Karlson, A., Kapoor, A., Roseway, A. and Czerwinski, M. "AffectAura: An Intelligent System for Emotional Memory". Proceedings of CHI 2012, 849–858, 2012.
- [23] E. Niforatos, and E. Karapanos, "EmoSnaps: a mobile application for emotion recall from facial expressions", Personal and Ubiquitous Computing. 19,425–444, 2 (Feb. 2015).
- [24] R. d. Oliveira, and N. Oliver, "TripleBeat: Enhancing Exercise Performance with Persuasion". Proceedings of MobileHCI 2008, 255–264, 2008.
- [25] A.J. Sellen, A. Fogg, M. Aitken, S. Hodges, C. Rother, and K. Wood, "Do Life-logging Technologies Support Memory for the Past?: An Experimental Study Using Sensecam", Proceedings of CHI 2007, 81–90, 2007.
- [26] M.M. Stevens, G.D. Abowd, K.N. Truong, and F. Vollmer, "Getting into the Living Memory Box: Family Archives & Holistic Design". Personal Ubiquitous Comput. 7, 210–216, 3–4 (Jul. 2003).
- [27] A. Thudt, D. Baur, S. Huron, and S. Carpendale, "Visual Mementos: Reflecting Memories with Personal Data". IEEE TVCG. 22, 369–378,1 (Jan. 2016).
- [28] J. Tholander, and S. Nylander, "Snot, sweat, pain, mud, and snow: Performance and experience in the use of sports watches". In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 2913-2922). ACM, 2015, April.
- [29] B. Wienke and D. Jekauc, "A Qualitative Analysis of Emotional Facilitators in Exercise". Frontiers in Psychology, 7,2016.