

APPENDIX TEXT 1

CHILDREN'S SEDENTARY BEHAVIOR WEARABLE CAMERA ANNOTATION PROTOCOL

Wearable cameras such as the SenseCam can take pictures at fixed intervals or in response to a change in a wearer's environment. Their images provide a rich source of data for researchers interested in sedentary behavior or physical activity. In order to analyze the data, images must first be split and annotated for the activities of interests. Doherty et al. have designed a Wearable Camera Browser for this purpose (<https://github.com/activityMonitoring/wearable-camera-desktop-browser> and <https://github.com/activityMonitoring/oxford-wearable-camera-browser>). This browser enables users to group images into albums that can be annotated with descriptive items. Annotated wearable camera data can then be extracted from the browser's database and used for analysis.

In an effort to ensure reliability and efficiency in this annotation process, this training protocol attempts to provide training information for the ethics, annotation guideline, quality checks and data security. Researchers who plan to use this protocol to create annotated wearable camera data should first be trained. The Wearable Camera Annotation Training Protocol developed by Willetts, Hollowell³⁷ served as a reference for the development of the Children's Sedentary Behaviour Wearable Camera Annotation Protocol by Johanna Hänggi, Efsthathios Christodoulides, Elke Gramespacher, Wolfgang Taube, Aiden Doherty. Comments for improving the protocol are highly appreciated.

TRAINING OVERVIEW

There are five steps for researchers to complete prior to any annotation work:

- A. undertaking the ethics training for handling image data (including data security practices);
- B. learning to use the image browser software;
- C. undertaking the annotation training;
- D. Annotating self-collected data;
- E. Going through cycles of coder training and calibration

A. Ethics Training

Researchers wishing to use image-based data should be aware of related ethical issues. Albeit it is beyond the scope of this protocol to present a detailed account of the ethical issues that arise when images are viewed by a large team of researchers, the decision making of most research ethics boards (e.g., institutional review boards) are based on 3 core principles. These are

- A. respect's autonomy (the right of others to make their own informed decision);
- B. non-maleficence (not doing harm, avoiding personal risks etc.); and
- C. beneficence (doing good or promoting well-being); and
- D. justice (ensuring that the benefits are equitable, including across those who are invited to participate vs those who are not)

At a minimum standard, researchers who wish to annotate wearable camera images should complete a tutorial on research ethics. Study-specific ethics training, which lists out the necessary reading, relevant research ethics courses and important notes, is recommended.

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggi et al.

All researchers must read the two declarations shown below which provide a clear list of journal articles related to research ethics of the wearable camera studies, a link to an online certificated course regarding protecting human research participants, and items of the important ethical considerations.

Data Security Practices

All annotated data are stored on the researcher's local desktop machine, which is secured within access-restricted premises, and need to be backed up periodically to prevent extensive data loss in the event that the machine is damaged. We recommend backing up the camera annotations regularly on a secure server. We also recommend making an archive copy of all participant data that will not be used for annotation.

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

Ethical declaration for researchers viewing and coding wearable camera images

16.06.2020

Study Name:.....

Principal Investigator:.....

Data Set:.....

Please initial box

I confirm I have read and understood the Ethical Framework by Kelly, Marshall ²² for wearable camera research	
I undertake to protect the privacy and confidentiality of all participants and their wearable camera data	
I understand that I must not disclose the content of any images to anyone outside of the research team	
I understand that I cannot make any copies of any images or post copies to any internet sites (Facebook, Twitter, etc)	
I understand that I must not send images over any unsecured internet connection (i.e. e-mail)	
I understand that data will be stored on a secure server (e.g. behind institutional firewall or a password protected external hard drive) and should never be stored on unprotected, mobile devices. (A password encryption will be provided)	
I must not share images with any third-party vendor that was not pre-defined in the approved study protocol	
I understand that any images used in presentations, papers or reports must be approved by the Study Principal Investigator before use	
I understand that I must report images that depict illegal activities (including but not limited to abuse, violence or drug taking) to the Study Principal Investigator	
I understand that the publication (including presentations, posters, papers, reports) of any data related to this study is not allowed without approval of the Study Principal Investigator	

Name of researcher:.....	Signature:.....	Date:.....
Responsible Research Associate:.....	Signature:.....	Date:.....

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

Principal Investigator:.....	Signature:.....	Date:.....
------------------------------	-----------------	------------

B. Browser Training

Prior to learning the annotation work, the researchers should download the Doherty Wearable Camera Browser onto their computer and understand the operation. In the training, the researcher should know:

- A. setting up the browser;
- B. downloading participant images from either PC database or the Camera directly;
- C. importing annotation list, choosing the date of interests, splitting and annotating images;
- D. making amendment of image annotation; and
- E. exporting annotated file in the CSV form (depends on the statistical analyses required).

Allowing public access to the browser and instructions, *Documentation of the Researcher Wearable Camera User Guide* is an open-course material available at <https://github.com/activityMonitoring/wearable-camera-desktop-browser/wiki>. Text description and video tutorials are accessible with links in the user guide. For detailed information, please visit the website.

C. Annotation Training

In this protocol, the researcher will find how “events” (i.e., a collection of consecutive images) are split and annotated from 1 of the 3 categories: uncodeable, sedentary type and environment location. Sub-categories are listed with supplementary notes.

The taxonomy of categories in the Children’s Sedentary Behaviour Wearable Camera Annotation Protocol is based on the Taxonomy of Sedentary Behaviour.³ Each item is linked with specified Metabolic Equivalent (MET) value. MET values can be referred to the Youth Compendium of physical activities published in 2018.⁴ Development of this annotation taxonomy is an on-going process and any comments and recommendations are hugely appreciated.

EXTRACTING SEDENTARY BEHAVIOR INFORMATION FROM WEARABLE CAMERA IMAGES: ANNOTATION PROTOCOL

Annotation Pass

An annotation pass is an annotation cycle where a researcher goes through all the images and categorizes only the SB components specific to that task. One annotation pass looks only at one component at a time, which should make annotating the images easier for raters. This allows a rater to look only at certain sub-components at a time and discard other components for the moment. This structure allows images to be annotated with multiple (sub)-components. For example, an image can be annotated for the location in one annotation pass and simultaneously for screen-based behavior in the other annotation pass. The annotation passes are “uncodable pass”, “sedentary non screen type”, “sedentary screen type” and “location. Below, under image annotation each annotation pass and its subcategories are described in detail.

Events

An “event” is a series of images grouped thematically depicting a certain component. Firstly, the manual will run through how to split events accurately.

Five-Image Rule

The definition of the image rule should be based on the population that is studied and the components that are of interest in the study, which defines the sampling rate of image capturing. In our case we studied the SB in children. We chose a short sampling rate (approximately 7 seconds) due to the intermitted and spontaneous physical activity behavior of children.^{5,6}

We chose to annotate events of at least five consecutive images and not image by image. Advantages are that this is less time consuming than image by image annotation and longer episodes of SB won’t be lost if they are interrupted by <5 images.

We defined the 5-image rule as following: Activities will be split into episodic “events”, each containing at least 5 images. The start of an event is the first image in a set of 5 (or more) consecutive images that depict the same component, or where the researcher is almost certain that the same component is occurring across the images. An event ends when an annotation component is no longer visible/happening or interrupted by more than 5 images that show a different component. This means events of SB shorter than 5 images (approximately 35 seconds) are not captured. Events that are interrupted by 4 images (approximately 28 seconds) or less (the annotated component is no longer visible) will not be split if the annotated component is visible after the interrupting images. As a result longer events of SB will not be lost if they are interrupted by an activity shorter than 28 seconds.

Interrupting Images

It is possible that the event does not depict the same component across all images. Events can be interrupted by <5 images showing another component. Accordingly, the researcher should include interrupting images (if <5 in a row) into the annotated event of the dominating component (5-image rule).

Primary Activities

As a participant might do multiple tasks at the same time, the researcher should identify the primary component for the annotation. The primary component is the one which the participant was actively performing. If relaxing takes place simultaneously with another

component, e.g. eating, the more active component is always prioritized over relaxing in the annotation (e.g., if someone is sitting in front of the table while eating, the eating annotation has priority). Therefore “relaxing, sitting, talking, lying around” is only annotated after all other categories of the sub-category can be excluded (See case study 1). If a participant is using several screens at once, the primary component is the one which the participant is actively performing, e.g. if mobile phone use takes place simultaneously with another screenbased SB, e.g. texting in front of the TV, mobile phone use (texting) would be the primary activity.

Images With No Label

This annotation protocol aims to assess components of SB in children. Therefore images showing physical activity behavior will not be labeled. This means that a component of a SB event can be surrounded by images that will not be annotated as they show active behavior components. Figure 1 illustrates how events should be split and which behavior should be annotated later.

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

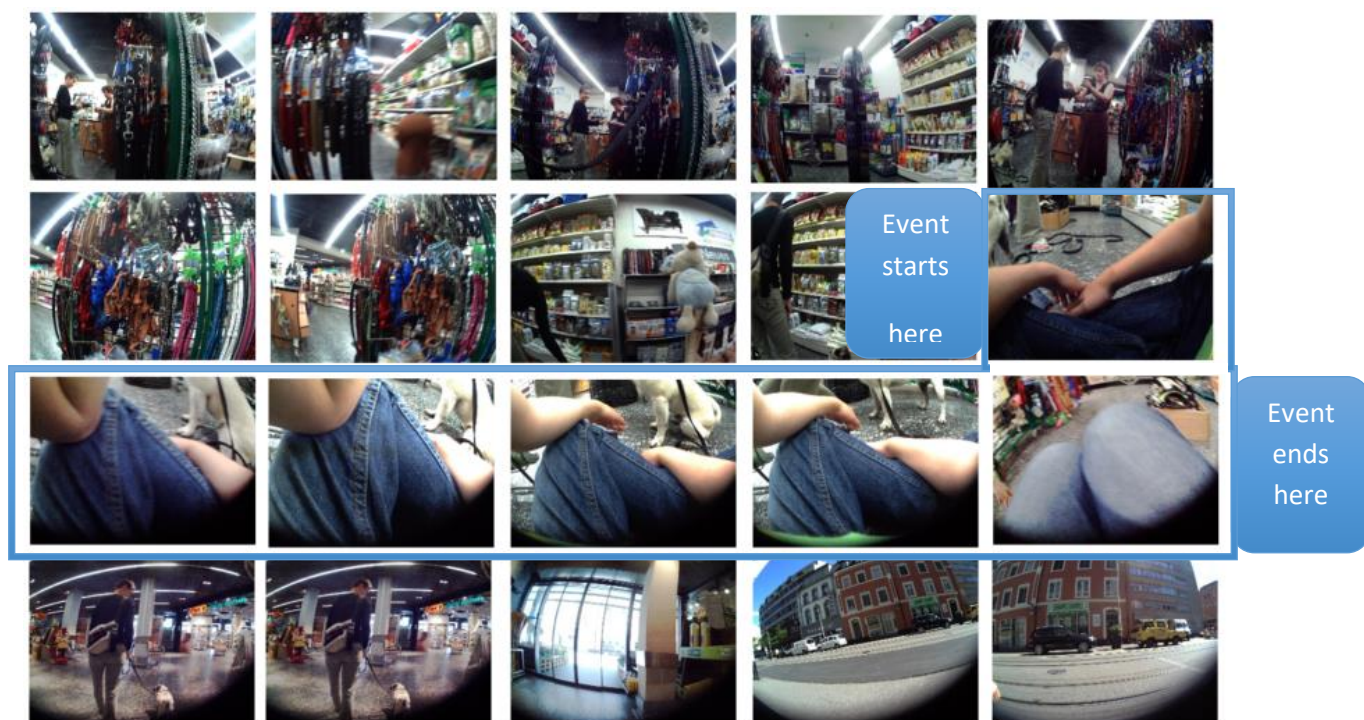


Figure 1. An example of events split.

The first 9 images show the participant moving around in a shop. This does not show a sedentary non screen type component, therefore images are not annotated. As the next 6 images show the participant sitting. This is a sedentary non screen type component and will be annotated as “relaxing/sitting/talking”. In the last 5 images the participant is standing up and walking away. As this is not a sedentary type component but an active behavior, the last 5 images will not be annotated.

R = Relaxing/sitting/talking

N= No label

1) Images: NNNNNNNN|NNRRR|NNNNNN

2) Split image: NNNNNNNN|RRRR|NNNNNN

Event 1-
Relaxing/sitting/

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

Image Annotation

After uploading photos and the import annotation list, the researcher should annotate events beginning with the lowest tier in the list. Table 1 summarizes the general rules of the annotation list for looking for the appropriate categories.

Appendix Table 1. A Summary of the General Rules for Identifying Events Under Each Category

Categories	Rules
Uncodable	When the images are blocked and/or dark or the images show that the participant took off the camera.
Type	The events after the participants had entered the working place and before they left that.
Environment	The events after the participants had left their destination (e.g., homes, working place, churches, restaurants, shopping malls) for a certain activity and before they entered into another one.
Not annotated	The events that are not sedentary or did not fall into the above categories.

The next section is the descriptions of annotation passes and sub-categories under these major categories (annotation passes). Supplementary notes are provided for a better understanding of appropriate annotation.

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

Uncodeable

Appendix Table 2. Categorization of Uncodeable Image

	Lowest tier items	Description and visual cues
Annotation pass 1: Uncodable	Camera taken off (MET value: undefinable)	A series of images in which the position of the photo remains the same in relation to fixed objects. Lighting commonly changes without movement of the images. There is an exceptional case for this annotation: if the researcher can clearly identify the person captured in the images to be the participant per se over the course when the camera was taken off, the researcher should annotate the observed events that the participant was doing
	Image dark/blurred/obscured (including camera being blocked) (MET value: undefinable)	Any set of images where the visibility of the image is impaired and therefore the coder cannot be sure of what is happening in the series of images in question. It has to be carefully considered whether the images could be codable for a certain annotation category, e.g. activity type cannot be identified but the location could still be codable. In this case, images should not be annotated as uncodable. To determine the annotation of darkened images, a rule of thumb is that a researcher cannot be certain about what is happening out of context (i.e., when looking at those images separately without referencing the previous and the following events).

Sedentary Type

Table 3. Categorization of Sedentary Type

	Sub-category 1	Lowest tier items	Description and visual cues
Annotation pass 2: Non screen sedentary	Non screen sedentary	Reading, memorizing	Reading, memorizing: Reading a book, newspaper, magazine. Because it is sometimes hard to distinguish between reading and writing the following rule of thumb should be considered: The activity should always be annotated as reading except it is obvious that the child is writing. If a paper with writing on it or a pen is lying on the table but the participant is not holding the pen, the event will be annotated as reading.
		Writing	Hand holding a pen in hand.
		Eating/Drinking	Presence of a plate, glass, sandwich, ice cream etc. eating activity does not have to be directly visible.
		Playing music	Presence of an instrument or music stand
		Spiritual	In church praying etc.
		Household	Cooking, preparing meals, setting the table, taking things out of the fridge, chores, loading the dishwasher, cleaning the window etc.
		Playing quietly	Includes e.g. board games, parlor games, puzzles, cards, children's games, panini stickers. Visual cue: If it takes place in the child's room or if a toy or a friend is present.
		Handicraft	Creating something with your hands, e.g. painting, sewing, crochet, tinkering, crafting etc.

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

		Relaxing/Sitting/Talking/Lying down	Relaxing is annotated if you don't see any other activity of the same sub-category 2. The background should not change too much during the relaxing activity. Relaxing is annotated independently of posture and location, however if participants are walking around while talking the event cannot be annotated as relaxing. If the participant is standing in a group talking, the event will be annotated as relaxing, as long as the location does not change and the participant does not move around. Waiting for a tram or bus is also annotated as relaxing. Screen time will never be annotated as relaxing. Holding something in their hands during sitting, relaxing, talking would be coded as relaxing and not as playing quiet games, if it is obvious that the participant is not playing/doing anything.
		Personal care	For activities typically performed in the bathroom, such as, grooming, washing hands, shaving, brushing teeth and putting on make-up, but also includes activities such as (un-) dressing, or (un-) packing a bag or suitcase.
	Screen time	TV	Watching TV is always annotated when a screen is visible even when the participant is not looking in the direction of the TV or screen.
Annotation pass 3: Screen time		Computer	
		Gaming console	
		Mobile phone	Using a mobile phone for gaming, texting, reading, watching videos etc.
		Tablet	Using a tablet for gaming, texting, reading, watching videos etc.
		iPod/MP3 player	Using an iPod, MP3 player for listening to music, gaming etc.
		Cinema	
		Other	

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

Environment, Location

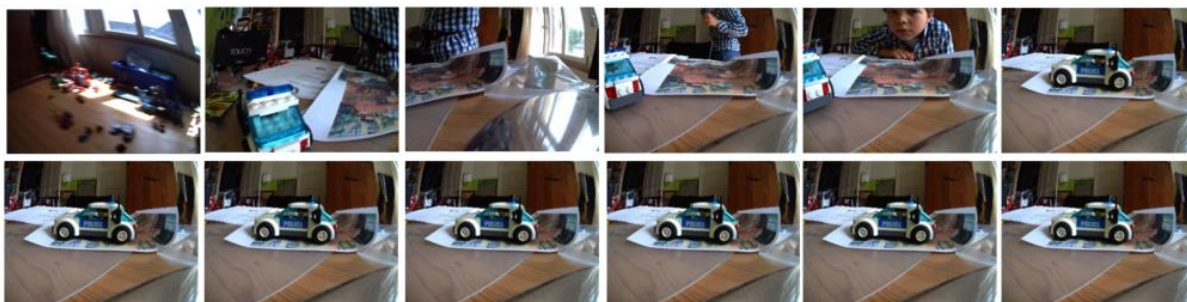
Table 4. Categorization of Location

	Sub-category 1	Lowest tier items	Description and visual cues
Annotation pass 4: Location	Outdoor	Nature	E.g. forest, mountains, while hiking
		Urban green space	E.g. parks, garden, private gardens
		Grey space street	Street, place, walking through the forest on a paved street is annotated as “grey space street”
		Grey space playground	Playground, Skatepark, Schoolyard etc.
		Other/mixed	E.g. at a train station, in a garden restaurant or in a half-open building with no doors/walls
	Indoor ^a	Home	Includes balcony, terrace
		School	
		Day care	
		Shops	Supermarkets, shopping centers etc.
		Sport facility	Gym, community facility, sport facility
		Other	

^aBeing in a means of transportation will not be categorized in the location tier, as it would be annotated in an additional tier, the transportation tier that was not of interest in our study.

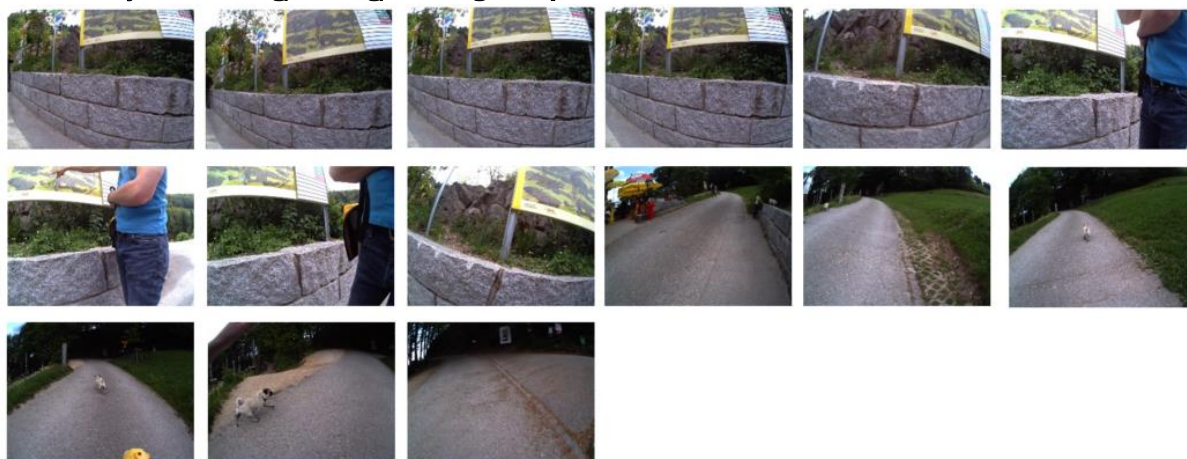
I. CASE STUDIES FOR ANNOTATION

Case Study 1. Camera Taken Off



The first 5 images show the participant, which indicates that the camera was taken off. As the participant is visible in the first 5 images, the images can still be annotated and would be annotated by the researcher as “playing quietly”. The last 7 images show all the same image, indicating that the camera has been taken off and the participant left into a different room. Therefore these images have to be annotated as “camera taken off”, because it can’t be determined what the participant is doing.

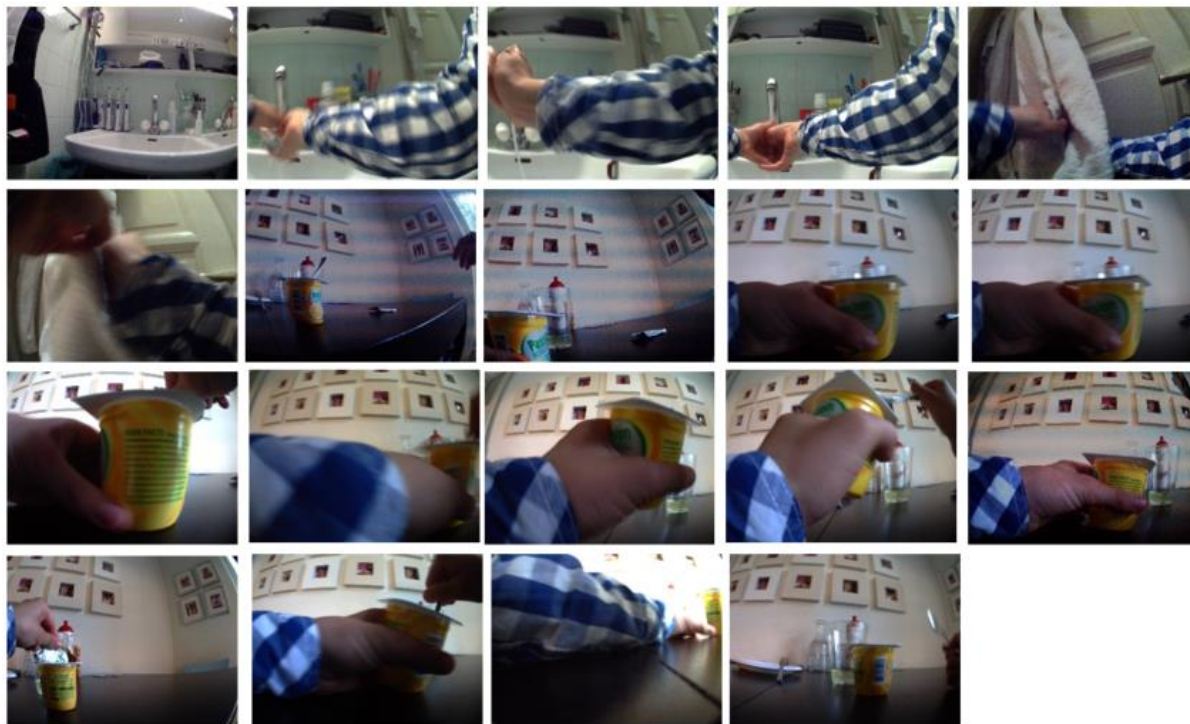
Case Study 2. Relaxing/Sitting/Talking Independent of Posture



The first 9 images show the participant assumedly sitting or standing in the same place in front of a poster. As we are interested in types of SB independent of posture and activity intensity, which could be annotated in different sur categories, the first 9 images would be annotated as “relaxing/sitting/talking”. “Relaxing/sitting/talking” will be annotated if the participant stays in the same place, which means the background is only slightly changing.

Appendix
Sedentary Behavior in Children by Wearable Cameras: Development of an Annotation Protocol
Hänggl et al.

Case Study 3. Personal Care and Eating



The first 5 images show the participant washing hands and would be annotated as “personal care”. The following images show a table and yoghurt, and sometimes the hands of the participants. Although the eating activity is not shown in the first 2 images, these images would be included in the “eating/drinking” event.

Case Study 4. Reading and Relaxing



The first 5 image show the participant in a relaxing position while holding a book and reading. As reading is the more active component, this event is annotated as “reading”. In the next 5 images, the participant is in a similar position but the book is not visible. Because the book is not visible in the last 5 images, this event should be annotated as “relaxing/sitting/talking”.

D. Annotating Self-Collected Data

To get familiar with the annotation protocol researchers should first annotate image data of a day that they collected by themselves. This process helps to understand how the different annotation categories are used and how events can be split. If open questions occur regarding annotation rules, these should be discussed with a research expert who already went through annotation training.

E. Going Through Cycles of Coder Training and Calibration **Evaluating the Validity of the Annotation**

In the context of annotating wearable camera images, annotation validity concerns the extent to which the researcher is accurately annotating what is actually occurring in the images. As we cannot be certain about what was actually occurring at the time the image was taken, we use an expert's annotation as the proxy of the real activities. An expert's judgment is still an opinion (albeit a good one) of a person's SB type, or environment, we consider the level of agreement between a new researcher's annotation and an expert's annotation a measure of content and face validity.³⁸

Evaluating the Reliability of a Researcher

In the context of annotating wearable camera images, annotation reliability refers to:

- inter-rater agreement: the extent to which they agree with other researchers annotating the same image. Using these definitions, you should see that validity is only a special case of inter-observer agreement (except that the "other researcher" is an expert).

In order to reach a high inter-rater reliability we suggest a coder training cycle by using nominal group technique.⁸ Figure 2: Coder training and calibration cycle explains the coder training and calibration cycle. Researchers should start with blinded coding of a test sample. Inter-rater reliability agreement should be assessed and a Cohen's kappa 0.80 or higher should be reached. If Cohen's kappa is below 0.80, a confusion matrix between 2 annotating researchers can uncover disagreements in annotations. These disagreements should be discussed within the group of researchers and a consensus should be found on how annotation certain cases. After a consensus was found blinded coding should be repeated until an inter-rater agreement of a Cohen's Kappa of 0.80 or above is reached.

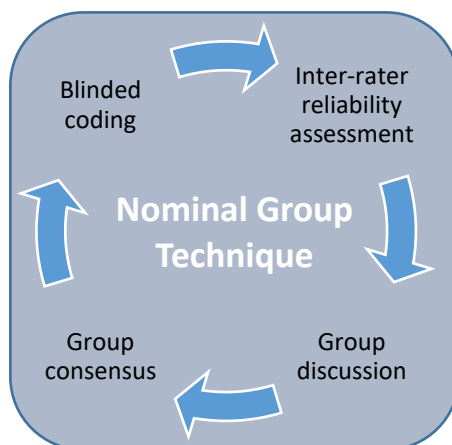


Figure 2: Coder training and calibration cycle

Guidelines for Inter-Rater Agreement Test

Inter-Rater Agreement tests (IRR) should be performed for validity and reliability tests using a statistical software program (such as R, SPSS, STATA). Training requires that a researcher to annotate a minimum of three sample sets of images pre-annotated by an expert with the Cohen's kappa score in the inter-rater agreement test of 0.80 or higher. Attempters who do not pass the training test must be re-trained and re-tested.

APPENDIX REFERENCES

1. Willetts M, Hollowell S, Aslett L, Holmes C, Doherty A. Statistical machine learning of sleep and physical activity phenotypes from sensor data in 96,220 UK Biobank participants. *Sci Rep*. 2018;8(1):7961. <https://doi.org/10.1038/s41598-018-26174-1>.
2. Kelly P, Marshall SJ, Badland H, et al. An ethical framework for automated, wearable cameras in health behavior research. *Am J Prev Med*. 2013;44(3):314–319. <https://doi.org/10.1016/j.amepre.2012.11.006>.
3. Chastin SFM, Schwarz U, Skelton DA. Development of a consensus taxonomy of sedentary behaviors (SIT): report of Delphi round 1. *PloS One*. 2013;8(12):e82313. <https://doi.org/10.1371/journal.pone.0082313>.
4. Butte NF, Watson KB, Ridley K, et al. A youth compendium of physical activities: activity codes and metabolic intensities. *Med Sci Sports Exerc*. 2018;50(2):246–256. <https://doi.org/10.1249/mss.0000000000001430>.
5. Bailey RC, Olson J, Pepper SL, Porszasz J, Barstow TJ, Cooper DM. The level and tempo of children's physical activities: an observational study. *Med Sci Sport Exerc*. 1995;27(7):1033–1041. <https://doi.org/10.1249/00005768-199507000-00012>.
6. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. *Prev Med*. 2007;44(2):143–147. <https://doi.org/10.1016/j.ypmed.2006.10.004>.
7. Kelly P, Fitzsimons C, Baker G. Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *Int J Behav Nutr Phys Act*. 2016;13:32. <https://doi.org/10.1186/s12966-016-0351-4>.
8. Potter M, Gordon S, Hamer P. The Nominal Group Technique: A useful consensus methodology in physiotherapy research. *NZ Journal of Physiotherapy*. 2004;32(2):70–75.