My Own Private Nightlife: Understanding Youth Personal Spaces from Crowdsourced Video

THANH-TRUNG PHAN*, Idiap Research Institute and EPFL, Switzerland
FLORIAN LABHART, Idiap Research Institute, Switzerland
DANIEL GATICA-PEREZ, Idiap Research Institute and EPFL, Switzerland

Private nightlife environments of young people are likely characterized by their physical attributes, particular ambiance, and activities, but relatively little is known about it from social media studies. For instance, recent work has documented ambiance and physical characteristics of homes using pictures from Airbnb, but questions remain on whether this kind of curated data reliably represents everyday life situations. To describe the physical and ambiance features of homes of youth using manual annotations and machine-extracted features, we used a unique dataset of 301 crowdsourced videos of home environments recorded in-situ by young people on weekend nights. Agreement among five independent annotators was high for most studied variables. Results of the annotation task revealed various patterns of youth home spaces, such as the type of room attended (e.g., living room and bedroom), the number and gender of friends present, and the type of ongoing activities (e.g., watching TV alone; or drinking, chatting and eating in the presence of others.) Then, object and scene visual features of places, extracted via deep learning, were found to correlate with ambiances, while sound features did not. Finally, the results of a regression task for inferring ambiances from those features showed that six of the ambiance categories can be inferred with R^2 in the [0.21, 0.69] range. Our work is novel with regard to the type of data (crowdsourced videos of real homes of young people) and the analytical design (combined use of manual annotation and deep learning to identify relevant cues), and contributes to the understanding of home environments represented through digital media.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing**; *Ubiquitous and mobile computing design and evaluation methods.*

Additional Key Words and Phrases: Youth; Mobile crowdsensing; Ambiance; Nightlife; Home

ACM Reference Format:
Thanh-Trung Phan, Florian Labhart, and Daniel Gatica-Perez. 2019. My Own Private Nightlife: Understanding Youth Personal Spaces from Crowdsourced Video. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 189 (November 2019), 34 pages. https://doi.org/10.1145/3359291

1 INTRODUCTION

The home environment is an important subject of study in several social sciences including psychology and geography, as well as architecture and design, and more recently computing [4], [5]. Private spaces at home include common living spaces in households (living room and kitchen), but

*This is the corresponding author.

Authors' addresses: Thanh-Trung Phan, Idiap Research Institute and EPFL, Switzerland, tphan@idiap.ch; Florian Labhart, Idiap Research Institute, Switzerland, flabhart@idiap.ch; Daniel Gatica-Perez, Idiap Research Institute and EPFL, Switzerland, gatica@idiap.ch.

2573-0142/2019/11-ART189 $15.00
https://doi.org/10.1145/3359291

also individual personal rooms (bedrooms). It is known that young people appropriate their private spaces and manifest aspects of their personal beliefs and traits in this way [52], [53], [54].

One important feature of place (private and otherwise) is ambiance. This is defined as "the character of a place or the quality it seems to have" [24] and used for both indoor environments [75], [68], [60], [16], [12] and outdoor environments [62], [65]. In the context of commercial spaces, ambiance plays an important role in customer behavior related to shopping [45], food choices [10], [82] or hotel experiences [21]. Regarding the ambiance of personal spaces, previous work has shown that ambiance mediates other factors, like gender [69] and personality [30], on choices made on physical and environmental characteristics. Understanding home spaces is a relevant domain that has not been fully studied in social computing. A better understanding of physical and social attributes and ambiance of personal spaces could have various implications for social computing research as a part of an agenda on living spaces and well-being. For example, homes can be reconfigured by their inhabitants with respect to decoration, spatial organization of furniture, light, and music, thus inducing more appropriate ambiances for certain activities and social interactions at home (e.g. a romantic dinner vs. an end-of-year party). Designing systems that both recognize physical and social attributes and support users to reconfigure their home spaces based on their specific goals is a relevant application. This could integrate the many perspectives existing in psychology, architecture, human geography, and public health, with the availability of environmental and mobile sensors and social media, and is a particularly interesting angle to understand and support youth practices. Furthermore, many traditional studies have collected information of personal spaces by using paper-and-pencil questionnaires and interviews. The potential of collecting in-situ information of home spaces (physical and social attributes of the environment) through technological means could add to the existing set of research tools.

Recent work on recognition of indoor ambiance [75], [60] has used still images from online social systems like Foursquare [75] or Airbnb [60]. Social psychologists have also investigated impression formation on home environments [32], [30]. Yet, gaps in the existing body of work emerge as most previous work has been conducted using social media data that either (1) might lack diversity in the representations of private residences [64], [22], as they are naturally focused on outdoor and commercial spaces; or (2) might be beautified, e.g. on Airbnb and similar sites, due to the intrinsic motivations to create and share such images [60]. To investigate home ambiance in a naturalistic setting, a different research direction could use crowdsourcing to collect in-situ videos of the personal environments inhabited by volunteers, which will yield vivid image and sound information, while reducing certain motivations of the video makers (e.g. performative or commercial) that could otherwise affect the generated content. Compared to previous work, our work uses crowdsourced video data of young people in their personal spaces during weekend nights. This provides a new view of youth nightlife activities in the private sphere, which enriches the kind of information that traditional methods in the social sciences can provide, in terms of temporal and scene granularity as well as scale.

The overall aim of the present work is to understand the characteristics of private spaces in youth nightlife in the weekend by investigating physical environment features and ambiances of home. Specifically, using a combination of human annotation and machine learning (computer vision & audio processing), we address the following research questions:

RQ1: Given crowdsourced videos recorded at home spaces by young people at night, what patterns of physical and ambiance attributes of youth home spaces can be revealed by manual coding of videos using external annotators and machine-extracted features?

RQ2: Given machine-extracted features of videos at youth home spaces, can these features infer the perceived ambiance of such spaces?

To address these questions, we use a crowdsourced dataset about nightlife, involving the participation of 241 young people, aged 16 - 25 years old, in the two main Swiss hubs for nightlife (Zurich and Lausanne) [72]. To obtain detailed insights on the locations attended and their related ambiance, participants were requested to record panoramic 10-second video clips of their environment at the start of the night and whenever they changed locations. In total, 841 videos were collected on 10 weekend nights. In this dataset, a significant portion of locations documented were private places [72], which provides a unique snapshot of how weekend nights are experienced by youth in their private environments. We design and implement an annotation task by asking external annotators to watch video clips. To build up this questionnaire, we have adopted several dimensions from related work [29], [31], [74], [60]. As a result, we generated a labeled dataset of 301 video clips in personal environments which contain richer in-situ information than what is often captured in questionnaires or surveys used in previous work, and manually annotated attributes of private spaces for our analysis.

Our paper has the following contributions:

(1) To address RQ1, we use a 301-video dataset of home spaces collected by Swiss young people on weekend nights. Our dataset contains video and audio files. A set of five independent raters annotated all videos with a rich set of questions, including physical attributes, social attributes, and ambiance. The results show that the video dataset can be consistently assessed by external raters, with at least moderate agreement, and in many cases with good or excellent agreement. Detailed analyses of the annotations produce several relevant results. First, we show that activities like eating, drinking, and entertainment (chatting, watching TV, and using digital portable devices) are all popular among young people, but with fluctuations over the night period. Second, we found a substantial number of cases where young people are alone and where home place loudness (chatter and music) is low. For those cases in which people socialize, we observed a same-sex trend between study participants and their companions. Third, we performed a correlation analysis among the ambiance attributes that showed two main opposite dimensions, namely places perceived as large, colorful, comfortable, festive, stylish, and unique; and a second category of places perceived as confined, simple, and boring. Dark and bright ambiances did not show significant correlation with the rest of the ambiance attributes. Finally, we use deep learning models applied on the audio and video tracks to extract automatic features to represent private spaces at the level of objects, scenes, and sounds; the results indicate the feasibility of using deep learning to produce generic semantic descriptions of home environments, although in several cases interpretation remains an issue.

(2) To address RQ2, we find that several of the 1000-object, 365-scene, and 527-sound classes used in our work have a particular correlation with specific ambiances. Finally, we use a machine learning pipeline to automatically infer ambiances of private spaces (as a regression task) using features informative of sounds, objects, and scenes. The results show that object and scene classes can predict six ambiances with R^2 between 0.21 and 0.69: space capacity (*large/spacious vs. cramped/confined*), brightness (*bright/well-lit vs. dark/badly-lit*), *comfortable/cozy*, and *dull/simple*.

The paper is organized as follows. Section 2 discusses related work. Section 3 presents the data collection and annotation process. Section 4 presents the in-depth analysis of private spaces based on the manual annotations. Section 5 presents the approach based on deep learning to extract visual and sound features of videos, examines the correlation between ambiance and the previously extracted cues, and presents experiments on automatic ambiance inference. Section 6 discusses the findings and limitations of our work from the perspectives of social computing. Section 7 concludes the paper.

2 RELATED WORK

Our work is related to a body of work from various disciplines examining issues of urban nightlife and youth; characterization of private spaces; and ambiance modeling. Each of these themes is discussed in the next subsections.

2.1 Urban nightlife and youth

Work in geography has studied the urban night period, often with qualitative methods [90]. The authors in [84] , [33], [14] also studied the dynamics surrounding youth experiences and urban nightlife. There is other work that has investigated the phenomena of human mobility and space usage in urban areas [59], [18], [81], [37]. From the perspective of alcohol consumption and urban youth, researchers investigated pubs and bars [25], house parties [40] and public spaces [23]. Especially, [90], [35], [11], [95] studied alcohol consumption from "pre-loading" (drinking before going out for the night) to excessive drinking with risky consequences. In contrast with these works, our paper aims to understand the characteristics and activities of the nightlife of youth in their home environments based on captured videos of the private spaces, contributed by the study participants in a crowdsensing setting.

2.2 Place characterization and private spaces

Regarding place characterization, the authors of [92] used mobile sensors, i.e. audio signals to infer occupancy, human chatter, music, and noise of places. Meanwhile, the authors of [20] aim to categorize places by using audio signals and images. Chon et al. [19] collected 48,000 place visits from 85 participants in Seoul to study the coverage and scaling properties of place-centric crowdsensing.

As a private space, the home is an environment where many social activities of young people unfold [4], [5]. In geography, Abbott et al. [4] investigated perceptions of young people about home as an idealized social construct and as a private space. Abbott et al. later investigated the social constructs of 'home' and 'neighborhood' as private and public spaces, in the context of leisure activities performed by young adolescents [5]. These studies used standard methods based on recall-based surveys. From a technical perspective, work in ubiquitous computing has developed approaches for place characterization, which use mobile sensors like microphones to extract audio signals through which certain features like human chatter and music can be inferred [92], or a combination of audio signals and still images that capture snapshots of everyday places [19], [20]. This body of work, however, has been largely focused on understanding outdoor spaces, often with goals of automatic place recommendation for urban users. In contrast to this work, we investigate how attributes of the home environment of young people are depicted on videos recording snapshots of weekend nights (a period of intense socialization among youth [4], [5]) using both human observers and machine-generated descriptors of the home environments.

2.3 Home spaces and activities

Home is conceived in different ways, including a physical space (house/apartment), someone's place of origin, or the place where people feel they belong [3]. Regarding the place of origin or where a person feels as belonging to, home is a site of 'shelter' [34] or a 'meaningful' place with multiple experiences through which people feel belonging [6], [87]. Home can not only be a fixed space but also an urban area, e.g., a street in town or a popular area in the city [6]. Home can also be a material place where young people live with their family [13], [77], or a student home or dormitory where students study or live away from their parents [36]. In our research, we aim to understand home as a personal space where young people spend their time alone or with friends on weekend nights.

Home is one of the places where youth spend their leisure time e.g., watching TV, listening to music [77], playing physical games [71], or drinking before going out with their friends at public places in the city [95]. Many people also socialize at their friends' or family's house [35], which emphasizes the importance to understand these practices, as the use of rooms and spaces at home can be influenced by architectural constraints, culture, an individual's daily life [7], or even mental distress [88]. Baillie et al. [9] studied four spaces in the home, including communication, work, leisure (private) and leisure (public) along with their utility to people living there. In our work, we investigate multiple dimensions of home of youth on weekend nights, including physical attributes (e.g. room types, brightness, music), social attributes (people present in the home environment), and ambiance (e.g. festive or fun).

2.4 Ambiance in architecture and psychology

The roles of interior architecture and design on human behavior have been studied in several disciplines, and provide background about the way humans interact with their living spaces. The characteristics of the places where we live, including space quality, interior design, and colors, affect how we feel, and reflect personal and social constructs [28]. Three main factors discussed by [28] influence living spaces: identity claims, thought/feeling regulators, and behavioral residues. Interior ambiance, i.e., "the character of a place or the quality it seems to have" [24] can have specific effects on people's behavior.

In the context of personal spaces, physical and environmental cues can reveal characteristics related to gender [69] and personality [30]. A common method used in psychology [69], [30] involves asking observers to manually rate physical spaces, which is an approach applicable to small-scale studies. Our work uses this methodology, and expands it by using automatic analysis to characterize the content of videos using state-of-art deep learning methods.

Ambiance has also been studied in public and commercial spaces. Quercia et al. [66] presented a crowdsourcing project related to ambiance-related constructs in the outdoor space, which studied how visual cues, color, and texture have effects on London neighborhoods along three perspectives: beautiful, quiet, and happy. In [45], physical and decoration cues had effects on the shopping behaviors of customers, because people's emotions and behaviors can be affected by these places' ambiance. Ambiance cues like color, brightness, and style have an important impact on customer emotions at hotels [21], or on food intake and food choice at restaurants [82]. For instance, [10] showed that decorating the ambiance of a pasta restaurant with a distinctive Italian feeling can make customers order more food.

Specifically for home environments, the Personal Living Space Cue Inventory (PLSCI) [29] describes personal living spaces, including 42 physical attributes and the ambiances of the space along with a checklist of 100 individual items. PLSCI is used by [17], [30] to study various questions in environmental psychology. We also adopt PLSCI for designing the video ambiance questionnaire for our study about home spaces of youth, which is discussed in Section 4.

2.5 Indoor ambiance inference in social computing

Several works have proposed methods to automatically recognize indoor ambiance from social media data. By observing the avatar pictures of Foursquare users, the work in [31] showed that people can identify place ambiance, clientele, and their activities with some degree of reliability. The work in [68] also used 4sq profile pictures to infer place ambiance using aesthetics, colors, emotions, demographics (age and gender), and self-presentation. Although the number of data points used in this work was small (N=49), it showed promise for place ambiance inference. Using data from Foursquare, the work in [74] generated crowdsourced annotations on an image corpus to study 13 ambiance dimensions. This dataset was later used to apply traditional visual features (color, GIST, HOG) and features extracted from a pre-trained CNN for ambiance inference [75]. The work in [12] further examined the problem of ambiance recognition through scene semantics, assuming that there are visual cues within scenes that can be extracted using a scene-centric semantic parser. We also adopt this assumption in our paper for conducting annotation on ambiance by asking raters to watch videos. However, the datasets used in previous work are images from Foursquare places, thus covering restaurants, bars, cafes, etc.. In this paper, we work with substantially different data, namely with videos capturing private spaces during 10 seconds, and through the combination of manual annotation results of ambiance based on the observation of the captured videos, and on semantic video cues extracted from deep learning models.

Airbnb is a social platform for hospitality that shows home environments to possible guests through photos. Ikkala et al. [38] conducted a qualitative research of hospitality exchanges on Airbnb. The study found that hosts on Airbnb have both financial and social reasons. In detail, money plays a role in supporting hosts in their efforts to manage social interaction, select guests consistent with their preferences, and control the volume and type of demand of visitors. In what constitutes the closest work to ours, [60] used a dataset of 1200 Airbnb venues represented by three images of each place to infer ambiances from visual features extracted from deep learning models. This work is an inspiration to our work, with one fundamental difference, namely that the visual data responds to very different motivations: crowdsensing for scientific research in our case, and illustrating home places for monetary purposes on Airbnb. This translates into rather different visual content: on Airbnb, images are curated to appear as appealing as possible to viewers; in our work, the videos produced by youth on weekend nights are unfiltered (except for reasons of sensitive situations and privacy) and non-beautified (as the study participants are sharing this data for research only and not for performative purposes as is often the case on Instagram and other social media).

To the best of our knowledge, our work extends the current understanding of private nightlife settings with respect to physical attributes at homes, activities of young people, and ambiances, building upon previous work in the CSCW, social computing, and ubicomp literature [77], [71], [67], [9], [66], [38]. In Table 1, we summarize the most closely related work and distinguish what we contribute to this domain.

3 DATA COLLECTION

Our work uses data from the Youth@Night project [72], which aimed at studying young people nightlife behavior in Switzerland using a smartphone application [72], [48]. This section provides an overview of the study design, the data collection procedure, and the specific data we use in this work.

3.1 Study design

3.1.1 Study context. Participants were recruited in Zurich and Lausanne, two of the four largest Swiss cities [72],[47] and the two main hubs of nightlife activities [58], [61]. They were approached by small groups of research assistants on the street between 8 PM and midnight in September 2014. In order to obtain a representative sample of nightlife goers, participants were recruited in popular areas (e.g., nightlife districts, public parks, streets), pro-rata of the area popularity at the city level. Quotas of people to recruit per area were determined using geo-localized venue data from Foursquare [73], and were validated with local experts (social workers and police). Eligibility criteria for participation were being aged between 16 and 25, owning an Android phone, having been out in the city at least once in the past month, and have consumed alcohol at least once in the

Work	Goal	Data	Tasks	Finding
[5] (2001)	Observe young people's favourite places and associated leisure activities at home and neighbourhood.	256 completed questionnaires and 58 interviews (28 girls, 30 boys from secondary school)	Quantitative and qualitative analyses	Young people want their homes to be friendly, spacious, modern, and quiet, they hangout with friends at home and friends' homes.
[92] (2014)	Infer the ambiance of business places from audio recordings	150 audio traces of indoor business and external surveys	Regression task for inferring the level of occupancy, human chatter, music, and noise levels using audio features	Classification performance of ambiance at 79% accuracy
[68] (2015)	Determine which visual cues of profile pictures can predict places' ambiances	Ambiance surveys of 49 places, with 250 annotations on 25 profile pictures on each place.	Regression task for predicting place ambiance using profile pictures' features	Predict ambiance based on faces at 78%
[74] (2015)	Investigate which types of social media images best convey indoor ambiance	50K images from 300 places on Foursquare, and 13 ambiance labels.	Interannotator agreement (ICC) analysis ans correlation analysis.	All 13 dimensions have ICC>0.5
[75] (2016)	Infer impressions of place ambiance, using generic and deep learning features	45,000 Foursquare images from 300 popular places in six cities	Regression task for inferring ambiance using machined-extracted features	Inferring place ambiance is feasible with a maximum R^2 of 0.53
[12] (2017)	Examine correlation of visual cues with ambiance of Foursquare images to automatically infer place ambiance	50K Foursquare images and 20K scene centric image dataset	Regression task for inferring ambiance using deep learning features	Ten of the ambiances can be inferred using scene objects and demographic attributes
[60] (2018)	Predict ambiance from pictures of listings on Airbnb	1200 Airbnb listings and crowdsourced annotations of images	Regression task for inferring ambiance using deep learning features	Ambiance can be inferred with R^2 up to 0.42
Our work	Describe youth personal spaces by means of crowd-sourced videos recorded in-situ. Labels are different than all above work except [60]. Infer ambiance at youth personal spaces from physical attributes	301 videos recorded in participants' home space on weekend nights. Manual annotations of the 301 videos by 5 independent annotators and CNN-based extraction of visual and audio descriptors.	Descriptive and correlation analyses of ambiances and physical features of home spaces. Regression task to infer ambiance using machined-extracted features	Living room, bed room, kitchen, and dining room are all represented at home on weekend nights. Top activities include drinking, chatting, watching TV, and eating. Home ambiance was often described as quiet and simply decorated. Regression for ambiance inference achieved R^2 between $0.21 - 0.69$.

Table 1. Comparison between previous work and our work.

past month (legal drinking in Switzerland, as in many other European countries, is 16 for beer and wine). The study protocol was approved by the ethical review boards of Vaud and Zurich cantons, and authorization to recruit on the street was obtained from the local authorities.

3.1.2 Data collection. The study took place on Friday and Saturday nights between September and December 2014. Participants were required to download and install Youth@Night applications. The survey logger application allowed participants to document, in real time, various aspects of their night, such as the locations attended (e.g. home, park, bar/pub), the type of drinks consumed, and 10-second video clips of their environment from 8 PM to 4 AM. Meanwhile, the sensor logger application, a background running app without any user interaction, collected many types of sensors and log data, such as GPS coordinates, accelerometer, and battery status [72],[47]. In this work, we will only use data from the survey logger application.

Questionnaires and sensor datasets were automatically uploaded to a back-end server when participants' smartphones had access to Wifi. Whenever the data was successfully uploaded, it was removed from the device. The participants could choose to manually upload data in case there was a problem with the automatic upload. At the end of the study, participants were paid 100 CHF if they documented at least 10 weekend nights. Participants completing less than 10 evenings with a minimum of three nights were paid on a pro-rata basis.

After the app-based data collection fieldwork, 40 qualitative interviews were conducted with study participants and focused on their experiences with the smartphone application, their experiences of nights out, and the ways in which mobile technologies shape contemporary nightlife [85], [86].

3.2 Measures

3.2.1 Video clips of environments. The survey logger application contained different questionnaires and media to capture participants nightlife behaviors, the locations attended, and the characteristics of their surrounding environment (see [48] for an overview of the different kinds and sequences of questionnaires). Participants were instructed to document any weekend night, including those during which they did not drink or did not go out, in order to have an overall representation of the different activities and events taking place on weekends. In the present work, we use the short video clips collected with the application at specific times of the night: whenever participants had their first drink (alcoholic or non-alcoholic) after 8 PM, and whenever they had a new drink (alcoholic or non-alcoholic) in a new location, they were required to indicate the type of location they attended (e.g. bar/pub, parks, home) and to record a 10-second video clip, which captured a panorama of their environment by slowly turning from left to right in landscape format. Participants thus recorded videos in varied environments, including pubs, clubs, public parks, means of transportation, and homes [72]. In case they were not able to record video (e.g., forbidden, felt uncomfortable), participants were told to skip the task and specify the reasons for it. Overall, participants recorded videos in 68% for the cases, while reasons for not recording were mostly because they did not it feel it as appropriate or safe [48]. Each video file was stamped with its time of submission. In total, 843 videos were collected from 204 participants on 646 participant-nights.

3.2.2 Annotation of home environments. After the fieldwork, we designed an annotation task to get qualitative information on the type of location, ambiance, physical attributes, and people shown on the 843 video clips recorded by the participants. Five independent annotators were hired and trained to watch the entire corpus of videos and answer 17 single and multiple choice questions on the type of location, the ambiance of the place, and characteristics of the social and physical environment. The exact questions and response options are presented in Section 4.

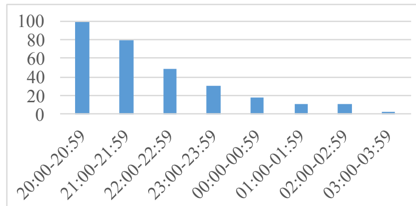


Fig. 1. Number of videos of private places (N=301) per hour.

3.2.3 Identification of home environments. Based on the annotators' answers to the question *"In what kind of place was the video clip taken?"*, places were considered as *"homes"* in case all five annotators agreed on this label. In total, this procedure retained 301 videos representing home places. In these environments, participants recorded videos in 64% of the cases. Reasons for not recording a video were: *"I was asked by someone not to do it"* (36%), *"it is not appropriate"* (25%), *"I don't feel safe"* (24%) and *"other"* (21%). Given that participants recorded videos of the environment whenever they had their first drink after 8 PM or moved to another location, (i.e., change of home, or come back home, in the present case), the 301 videos illustrate home environments throughout the night, although with a larger proportion of those taken early in the night if the participants did not change location. Figure 1 shows the number of videos per hour. Because of the small number of observations per hour after midnight, environments documented after midnight will be aggregated in the rest of the analyses.

Due to privacy requirements requested by the Ethical Review Boards that reviewed and approved the project, we cannot make this dataset publicly available.

4 PHYSICAL/SOCIAL ATTRIBUTES AND AMBIANCE OF HOME SPACES (RQ1)

In this section, we investigate how main patterns of physical attributes and ambiance can be extracted from videos recorded in private spaces using external annotators. In the following subsections, we first explain the measure (i.e. exact questions and response options), investigate the consistency of annotations across the five annotators using Intraclass Correlation analyses, and provide descriptive results. The Intraclass Correlation Coefficient (ICC) is a standard measure of reliability of raters [79]. As recommended by [44], we used ICC(2,k) which is used for a fixed set of k judges rating each target (N=301) and reflects the absolute agreement. Following the guidelines from from Koo and Li [44], ICC scores below 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 are indicative of poor, moderate, good, and excellent reliability, respectively. ICC(2,k) can only be computed on numerical variables, not on categorical ones, so there are few cases in this section that does not show ICC. We summarize results of ICC scores for each possible question in Table 2.

4.1 Overall representation of the space

In the annotation task, after carefully watching each video, several times if required, annotators were asked to indicate *"How well does the video capture the physical space (i.e. space layout, background scene, furniture, decoration, etc.)?"* with five single-item response options. *"[1]not well at all"*, *"[2]not so well""*, *"[3]regular"*, *"[4]well"*, and *"[5]very well"*. Results showed a good level of agreement on this question (ICC = 0.83). Figure 2 shows the histogram of all individual responses of all annotators to this question (301 x 5 =1505). As seen in Figure 2, most of the videos were rated as providing a *"regular"* representation of the space. The mean of this variable is 2.91 (SD=0.86), which is sightly lower than 3 ("regular"). In some cases, participants avoided recording directly physical spaces that

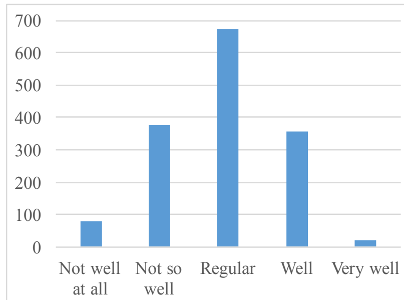


Fig. 2. How well physical spaces are captured in videos. The y-axis represents the total number of annotations.

Physical attributes at homes	ICC(2,k)	mean	std	skew
Physical space (i.e. space layout, scene, decoration, etc)	0.83	2.91	0.86	-0.17
Amount of light	0.87	2.90	0.69	-0.30
How loud is the music	0.95	1.44	0.81	1.70
Level of overall chatter	0.94	1.71	0.94	0.80
Level of occupancy	0.97	1.82	1.09	1.12

Table 2. ICC of physical attributes at homes based on N(video)=301, N(raters)=5, with scale (1-5).

could contain people. For instance, some participants recorded the ceiling or floor while panning the camera.

4.2 Physical and Social Attributes

4.2.1 Room of the home. In the annotation task, the room type within the home was labeled using the question: *"Where in the home was the video taken?"* and the following single-item response options: *"[a]living room"*, *"[b]dining room"*, *"[c]kitchen"*, *"[d]bedroom"*, *"[e]corridor"*, *"[f]terrace/balcony"*, *"[g]other"*, and *"[h]impossible to say"*.

Figure 3 shows the frequency with which individual annotators identified specific rooms of the homes in the 301 videos. Living room, bedroom, kitchen, and dining room are the most attended spaces within homes. This result echoes previous work that using traditional methods reported that living rooms and bedrooms are the most used places in small and large homes by occupants [43], and extends this previous finding by showing that for the specific case of young people on weekend nights, kitchens and dining rooms are also frequently used indoor spaces. As mentioned previously, a few videos avoid capturing directly the physical spaces by turning the camera to the ceiling and floor. This is one of reasons why *"Impossible to say"* appears in Figure 3.

4.2.2 Brightness. The annotators were asked to answer a single choice question *"Describe the amount of light in the place"* with five choices *"[1]It is very dark"*, *"[2]It is quite dark"*, *"[3]Normal"*, *"[4]It has a good lighting"* to *"[5]Is is very bright"*. The ICC(2,k) of brightness is high (0.87). The brightness variable has a mean slightly below the middle of the scale (2.9, SD=0.86). Figure 4 shows the histogram of annotated brightness, brightness per hour (8:00-8:59 PM, etc.), and brightness per hour expressed as a percentage within that timeslot, respectively. The percentage of darkness (*quite dark* and *very dark*) increases from 18% (8PM) to 35% (0-3AM) in Figure 4c. Conceptual work in geography [78] has recently discussed how individuals at home in the dark might be more willing to open themselves to others, and how adjusting the darkness of the home environment can be

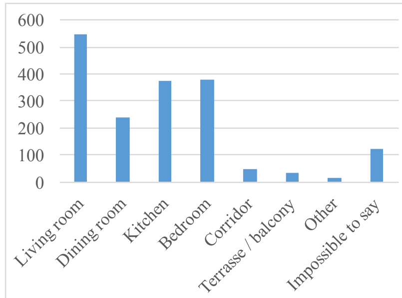
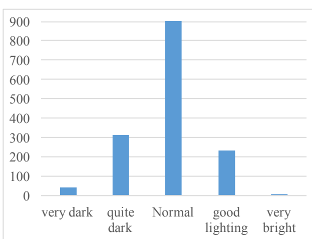
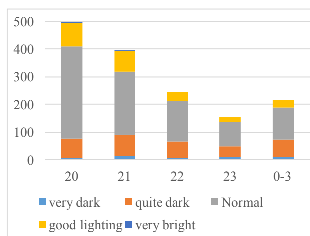


Fig. 3. Types of spaces at homes captured in videos. The y-axis represents the total number of annotations.

a)


b)


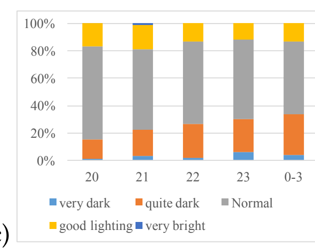
c)


Fig. 4. (a) Brightness, (b) Brightness per hour (8:00-8:59 PM, etc), (c) Brightness per hour expressed as a percentage within that timeslot. The order of levels of brightness (very dark, quite dark, normal, etc.) is left-to-right in graph a, and top-to-bottom in graphs b and c. The x-axis on graph b and c is hour on Friday and Saturday nights from 20:00 to 3:00. The y-axis on graphs a and b represents the total number of annotations, while the y-axis on graph c represents the percentage normalized on each hour.

empowering. Our annotations suggest that as the weekend night goes on, young people at home indeed tend to be in conditions of lower illumination. As a reminder, note that given the season of the year when the data was collected (mid September through December), it was past sunset time at the beginning of each recorded night (8PM).

4.2.3 Music Loudness. Regarding music loudness at home places, the annotators were asked to answer *"Describe how loud is the music in the place"* with five choices *"[1]No music"*, *"[2]Low"*, *"[3]Medium"*, *"[4]Loud"*, and *"[5]Very loud"*. The ICC(2,k) of music loudness is excellent (0.95). The mean value (1.44) is low (SD = 0.81). The skew is large (1.70) showing that the distribution has a tail. Figure 5 also shows the corresponding temporal trends. Overall, the present results on music and brightness levels at home are consistent with recent ethnographic research showing that young people tune their home by turning off lights and choosing slow paced music when they spend time drinking with their friends at night [94]. We found that no music was played in most of the recorded environments (frequency: 76%; see Figure 5a). When music was played, the loudness level was quite low throughout the night (Figure 5c), suggesting that the cohort of young people are relatively quiet in their private nightlife.

4.2.4 Chatter Loudness. The annotators were asked to describe the level of chatter loudness at home space by answering the question *"Describe how loud is the chatter in the place"* with five single choices *"[1]No chatter"*, *"[2]Low"*, *"[3]Medium"*, *"[4]Loud"*, and *"[5]Very loud"*. Similarly to music loudness, the ICC agreement for chatting loudness is very high (0.94). The mean value is low (1.71, SD = 0.94). Figure 6a-c shows that there is not much loud talking in the recorded videos. Relative to

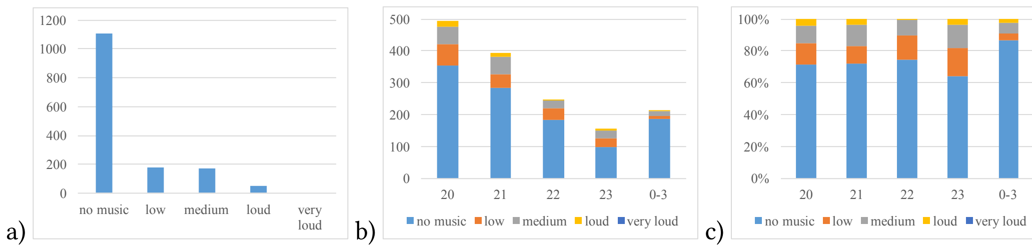


Fig. 5. (a) Music loudness, (b) Music loudness per hour, (c) Music loudness per hour expressed as a percentage within that timeslot. The x-axis on graph b and c is hour on Friday and Saturday nights from 20:00 to 3:00. The y-axis on graphs a and b represents the total number of annotations, while the y-axis on graph c represents percentage normalized on each hour.

each hourly slot, medium and loud chatting slightly increase from 8 PM to 11 PM (Figure 6c). This result is clearly connected to the results obtained for the occupancy of the physical space discussed next.

4.2.5 Occupancy. Annotators were also required to describe the level of occupancy of the place by using the following single choice question *"Describe the level of occupancy of the place based on what you hear or see"* with five choices *"[1]Empty"*, *"[2]There are few people for this space"*, *"[3]It's half empty/half full"*, *"[4]It's well attended, but there could still be more people"* to *"[5]It's highly crowded/packed"*. Annotator agreement of occupancy level was excellent (ICC = 0.97). The mean of level occupancy of the place is 1.82 (SD=1.09). While we anticipated that most young people meet with others at home on weekend nights, Figure 6d shows that empty is the most common category. Figure 6f also shows that young people slightly reduce gathering together from 8 PM to 10 PM; then, gathering increases at 11 PM, and decreases again after midnight.

4.2.6 Number of people present. As a complement to occupancy, we asked the annotators *"How many people appear on the video (in addition to the phone holder)"* with six choices *"[1]0 (the person seems to be alone)"*, *"[2]1"*, *"[3]2-4"*, *"[4]5-10"*, *"[5]More than 10"* to *"[6]Impossible to say"*. Figure 6g shows that around 40% of videos are labeled as containing no people, which is consistent with the labeling of occupancy.

4.2.7 Gender of people present in videos. Among the people present in the videos, we examined their gender ratio by asking one single choice question: *"What is the gender ratio of the relatives, friends, or acquaintances appearing in the video?"* with 6 response options *"[1]Women only"*, *"[2]Mostly women"*, *"[3]Half-half"*, *"[4]Mostly men"*, *"[5]Men only"* to *"[6]Impossible to say"*. Figure 7a shows that "men only" is the most common situation, followed by "women only" and "half-half". The total number of situations with "men only" and "mostly men" is higher than those with "women only" and "mostly women", suggesting that men appeared more often in the videos than women. Surprisingly, the 301 videos were fairly evenly distributed per gender, with 144 videos recorded by 52 male participants and 157 videos recorded by 50 female participants. Figures 7b and 7c show the gender repartition of the people present in the videos recorded by male and female participants, respectively. Male participants mostly tend to spend their nights at home with other male friends and less so with women, while no clear preference could be observed for female participants. As a point of reference, work on a sample of 377 students [93] showed that young females tend to hang out at home with friends more than males do.

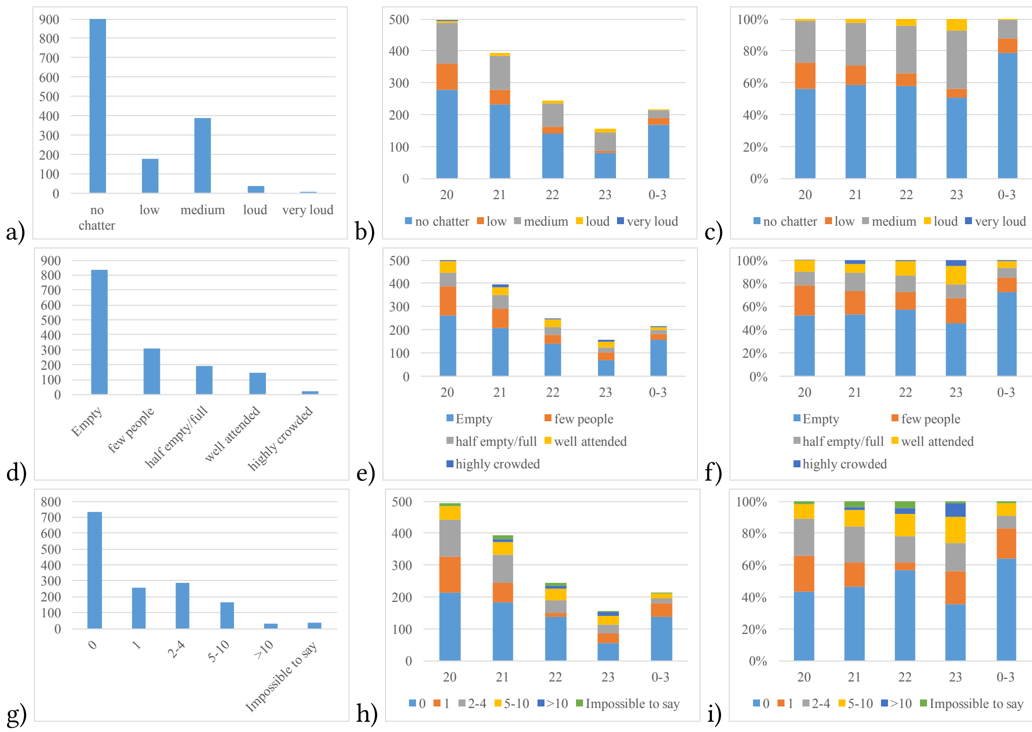


Fig. 6. Annotation of (a-c) chatter level, (d-f) occupancy level, (g-i) the number of people in the videos. The left column shows the overall trend, the middle column the trend per hour, and the right column the relative percentage for each timeslot. The order of values of all legends is left-to-right, top-to-bottom in all graphs. The x-axis on graphs b-c, e-f, h-i is hour on Friday and Saturday nights from 20:00 to 3:00. The y-axis on graphs a-b, d-e, g-h represents the total number of annotations while the y-axis on graph c, f, i represents percentage normalized on each hour.

4.2.8 Activities of people. In order to assess the activities of young people at their home spaces, we asked annotators to indicate *"What things are people doing in the video?"* with 14 multiple choices items shown in Figure 8a. Results showed that activities are quite diverse, with drinking, chatting, watching TV, using smartphone/tablet/computer, and eating as the five most common activities. As seen in Figure 8c, these main activities are roughly constant from 8 PM to midnight. Drinking, as the most commonly annotated activity, takes 15-25 percent in relative terms across all hourly slots. The prevalence of this activity at home is not surprising given that participants were requested to document the environment when they had their first alcoholic or non-alcoholic drink there. Nevertheless, this finding also echoes to previous research from Valentine et al. [89] showing that 73% of young people report having consumed alcohol at their homes and 64% at their friends' houses over the last year. Yet, our analysis brings a finer grained description of temporal trends. In addition, we also examine activities of young people depending on the level of occupancy and type of space at homes, as shown in Figure 9. In Figure 9a, when the place is empty (i.e. only the person recording the video is present), the most commonly annotated activities are watching TV, using a computer/tablet/smartphone and, to some extent, drinking. Conversely, in the presence of other people, the commonly annotated activities are chatting, drinking, and eating, whose proportions

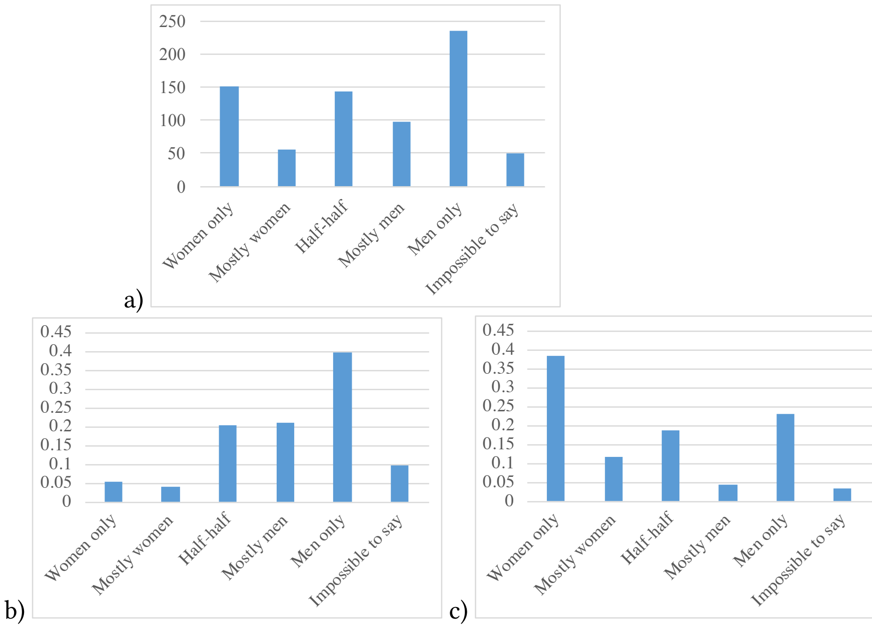


Fig. 7. (a) Frequency of gender of people appearing in 301 videos recorded by 102 male and female participants, (b) percentage of gender of people appearing in 144 videos recorded by 52 male participants, and (c) percentage of gender of people appearing in 157 videos recorded by 50 female participants. The y-axis on graph a represents the total number of annotations, while the y-axis on graphs b and c represents percentage normalized on each possible value on the x-axis.

increase along with levels of occupancy. It might also be noticed that playing board games was the most frequently reported in "half empty/full" homes, and some dancing was reported in highly crowed homes. Figure 9b shows that there are four places at home spaces that co-occur with specific activities: terrace/balcony/corridor; kitchen/dining room; living room; and bedroom. In related CSCW work, Baillie et al. [9] study leisure (private) and leisure (public) places in terms of their utility to inhabitants of a house. We complement this by showing that chatting and drinking occur more (in distributional terms) in leisure public areas within homes (terrace/balcony/corridor, kitchen/dining room, living room), while activities like using computer/tablet/smartphone and watching TV occur around 60% in private leisure spaces (bedroom).

4.2.9 Reactions of people around in videos. To conclude our research on physical and social attributes at home spaces, we examined reactions of people around in videos by asking five annotators to answer a single choice question *"Can you see or hear one or more persons reacting to or being aware of the video being recorded?"* with two answers *"[1]Yes"* and *"[2]No"*. If the previous question gets answered *"Yes"*, we will ask five questions listed in Figure 10b with three single choices *"[1]Yes"*, *"[2]No"*, *"[3]Not sure"*

We are interested in how people in videos react to video recording in home spaces. As we mentioned, many videos did not get recorded by design, as participants were told not to do it if not appropriate. Regarding the 301 recorded videos at home spaces, in 25% of cases did people in the video react to the camera (shown in Figure 10a). Two of the main reactions were *having fun while the video is recorded* and *asking about or commenting on the purpose of the video*. It is important to note that participants in the study were explicitly instructed to record video only when it was

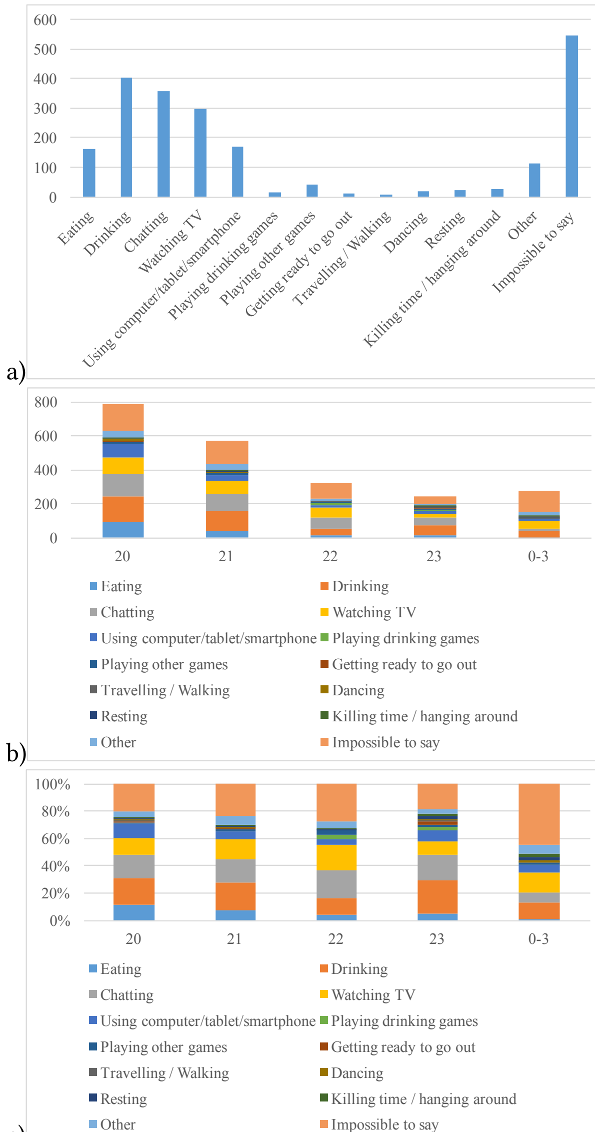


Fig. 8. Frequency of occurrence of (a) activities (b) activities per hour, and (c) percentage of activities within each timeslot. The order of activities (eating, drinking, chatting, watching TV, etc.) is left-to-right in graph a, and top-to-bottom in graphs b and c. The x-axis on graph b and c is hour on Friday and Saturday nights from 20:00 to 3:00. The y-axis on graphs a and b is the total number of annotations while the y-axis on graph c is percentage normalized on each hour.

socially acceptable and agreed and they were free to avoid recording [72]. The video dataset used here was recorded with such guidelines. There are just a few cases showing that people in the video were not comfortable about being recorded or to hide their face.

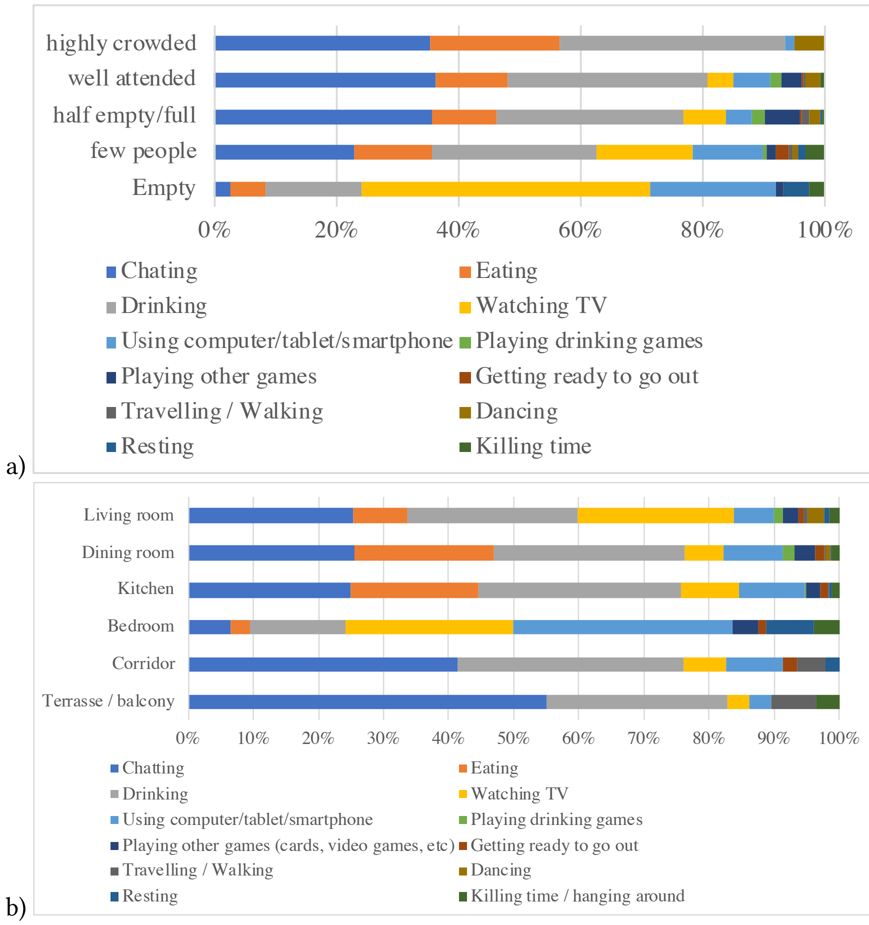


Fig. 9. Percentage of activity based on a) level of occupancy, and b) type of space at homes. The order of activities (chatting, eating, drinking, watching TV, etc.) is left-to-right, top-to-bottom in graphs a and b.

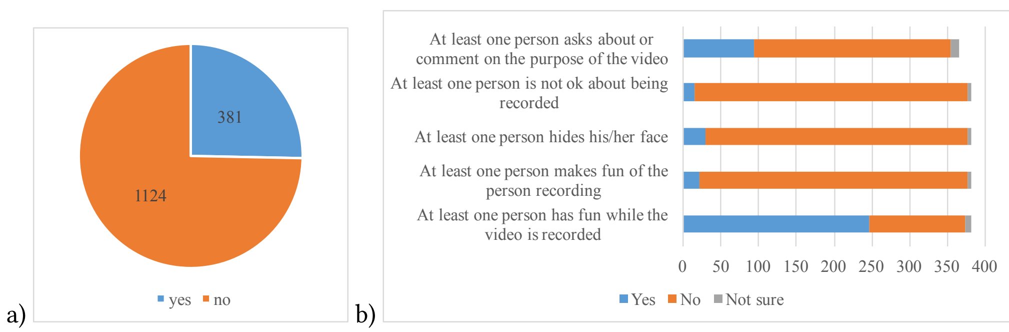


Fig. 10. (a) Yes/No reactions of people in videos. (b) Description of reacrions to the videos.

Ambiances	ICC(2,k)	min	max	mean	std	skew
Large, spacious	0.81	2.0	6.4	3.9	1.3	-0.09
Dark, badly-lit	0.83	1.4	7.0	3.8	1.6	-0.26
Colorful, decorated	0.66	1.8	6.2	4.0	1.3	-0.40
Cramped, confined	0.67	1.4	6.2	3.7	1.6	-0.23
Bright, well-lit	0.81	1.0	6.2	3.7	1.5	-0.18
Comfortable, cozy	0.61	1.8	5.6	3.7	1.2	-0.33
Dull, simple	0.63	2.0	6.2	4.1	1.3	-0.40
Festive, fun	0.27	1.8	4.8	2.9	1.4	0.00
Sophisticated, stylish	0.65	1.0	6.0	2.8	1.6	0.52
Off-the-beaten-path, unique	0.35	1.4	6.2	2.8	1.5	0.18
Serious, boring	0.21	1.8	5.2	3.5	1.5	-0.42

Table 3. ICC of ambiance categories at homes based on N(videos)=301, N(raters)=5 with scale (1-7).

4.3 Ambiance attributes

To assess the ambiance of home environments , we used a modified version of the Personal Living Space Cue Inventory (PLSCI) [29]. This instrument was originally designed to describe personal living spaces, e.g. rooms in family households, dormitories, or residential places. In our case, we augmented the PLSCI with ambiance attributes from previous work [31], [74], [60], [67]. As a result, we obtained a list of 11 ambiance word groups (e.g. *large/spacious, cramped/confined*; all items are listed in Table 3). A Likert scale, used in previous ambiance work but also as a reliable methodology to annotate image aesthetics [80] was used in our work. Annotators had to rate each ambiance by indicating, on a 7-point Likert scale ranging from *"[1]strongly disagree"* to *"[7]strongly agree"*, the degree to which they agreed with each of the ambiance attributes.

As seen in Table 3, moderate-to-good agreement levels were found for 8 out of the 11 ambiance characteristics (ICC greater than 0.5), but 3 items, namely *festive/fun*, *serious/boring*, and *off-the-beaten-path/unique* had ICC under 0.5. Attributes relating to physical characteristics of the place (*large/spacious*, *cramped/confined*) and its brightness (*dark/badly-lit*, *bright/well-lit*) have the highest agreement ranked as good (between 0.75 and 0.9). This indicates that the ambiances relating to physical attributes are easier to rate than attributes relating to the annotators' judgments on more subjective variables (*Serious, boring, Festive, fun* and *Off-the-beaten-path/unique*). This result is in concordance with the work of Nguyen et al. [60] on Airbnb personal homes, in that annotation on ambiance requires observers to make abstract impressions, which makes consistent annotation challenging for variables like *festive/fun*, *serious/boring* and *Off-the- beaten-path/unique*. Regardings descriptive statistics, the highest mean values are obtained for *dull/simple* (4.12), *colorful/decorated* (4.02), *large/spacious* (3.89), and *dark/badly-lit* (3.82).

4.3.1 Ambiance Correlation. Table 4 displays the Pearson correlation between the annotated ambiances for all home places (N=301). In Table 4, we only show correlation above 0.20 and p-value <0.001. From this analysis, we can identify opposing pairs, e.g. *large/spacious* vs. *cramped/confined*, and *dark/badly-lit* vs. *bright/well-lit*. but also observe other effects. All characteristics are associated with some others, with clearly identifiable patterns. First, characteristics related to brightness, namely *dark/badly-lit* and *bright/well-lit*, are uncorrelated to all other ambiance characteristics, suggesting that variations in lightings are independent of the general perceived ambiance. Second, characteristics of *serious/boring*, *cramped/confined*, and *dull/simple* were all grouped together (i.e., positive correlations between all three characteristics), while characteristics of *large/spacious*,

Ambiance attributes	[a]	[b]	[c]	[d]	[e]	[f]	[g]	[h]	[i]	[j]	[k]
[a] Large, spacious	-	*	*	-0.92	*	0.42	-0.38	0.23	0.66	0.26	-0.29
[b] Dark, badly-lit		-	*	*	-0.94	*	*	*	*	*	*
[c] Colorful, decorated			-	*	*	0.56	-0.72	0.55	0.28	0.54	-0.56
[d] Cramped, confined				-	*	-0.41	0.35	*	-0.66	-0.21	0.24
[e] Bright, well-lit					-	*	*	*	*	*	*
[f] Comfortable, cozy						-	-0.61	0.48	0.49	0.45	-0.50
[g] Dull, simple							-	-0.67	-0.54	-0.67	0.72
[h] Festive, fun								-	0.29	0.54	-0.68
[i] Sophisticated, stylish									-	0.41	-0.31
[j] Off-the-beaten-path, unique										-	-0.64
[k] Serious, boring											-

Table 4. Pearson correlation of ambiance (based on N(video)=301 having p-value <0.001). Entries marked with (*) correspond to correlation <0.20 and p-value >0.001.

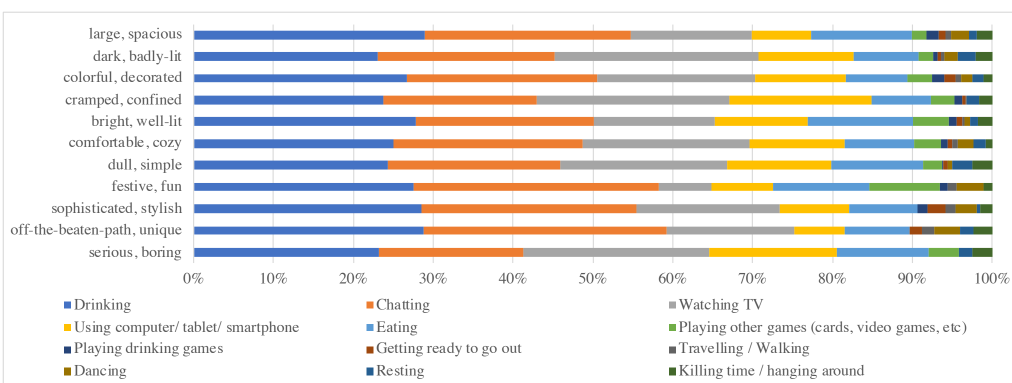


Fig. 11. Percentage of all activities co-occur with all ambiances. The order of activities (left-right and top-bottom on the legend), e.g., drinking, chatting, watching TV, using computer/tablet/smartphone, etc. are plotted from the left (0%) to the right (100%) on the stacked bar.

colorful/decorated, *comfortable/cozy*, *sophisticated/stylish*, *off-the-beaten-path/unique*, and *festive/fun* were also grouped together.

4.3.2 Co-occurrence of Ambiance and Activities. Figure 11 shows the relative distribution of activities for the different types of ambiances. For the figure, each ambiance was binarized, such that each place is associated to a given ambiance only if the average rating over all annotators is above the mean scale (4.0). Overall, 'Drinking', 'Chatting', 'Watching TV', 'Using computer/ tablet/ smartphone' and 'Eating' were the most prevalent activities, independently of the ambiance, although subtle variations can be observed. For example, chatting was more prevalent in unique, large, and sophisticated places, while the use of electronic devices seemed more prevalent in serious, dull, and confined places. The only ambiance that seemed largely different from the others is festive, fun, which showed a lower proportion of watching TV and using electronic devices than the other ambiances.

4.4 Automatic extraction of audio and visual descriptors of home environments

4.4.1 Video Preprocessing. We extract visual and audio descriptors of places from the 301 10-second video clips using deep learning. Table 5 summarizes the outcomes of the learning models presented below. Following the recommendation in [42] to extract at least 8 frames per second using uniform sampling, we extract a total of 29K frames. Meanwhile, we also extract 301 audio files for all videos by using command line FFmpeg [26].

4.4.2 Object Parser. To obtain an object-level description for each video, we used a deep learning model to extract the probability of object appearance in each frame. We applied the Inception-v3 model [83] trained on the ImageNet Large Visual Recognition Challenge. This model classifies entire images into 1000 classes (e.g. dishwasher, refrigerator, etc.) where the output for each image at the last layer is the probability distribution over all object classes (i.e., the sum of the scores over the 1000 classes is 1.0). The work in [83] presented the fraction of test images for which the correct class label is not among the top five labels identified by the algorithm, namely "top-5 error rate", reported to be 3.46%. As a result, for each frame, we have a 1000-dimensional vector with each element as a probability. Then, we aggregate them at the video-level over all frames to include all the existing objects by computing, for each class, the maximum probability over the set of video frames.

4.4.3 Scene Parser. To obtain a scene-level description for each video, we extract 365 place classes (e.g. kitchen, living room, etc) using Resnet18 [98] trained on the Place-365 database [97] for each frame. The semantic categories of the place classes are defined by their function, e.g., dressing room for dressing, locker room for storing, etc. As explained online, the database is meant to be used for "high-level visual understanding tasks, such as scene context, object recognition, and action and event prediction." The output of the last layer is a 365-dimensional vector in which the sum of all element values is 1. In order to represent the scene of the full video, we aggregate vectors over all frames of each video by computing the average for each class.

4.4.4 Sound Parser. To get a scene-level representation of the sounds present in a video, we extract 527 audio classes using Vggish trained on the Audio Set dataset of generic audio events, which has 1.7 million human-labeled 10-second YouTube video soundtracks [27]. The output of the last layer is the probability of each individual sound detected by the model.

Figure 13a, b, c shows the top 30 descriptions extracted for 1000 objects, 365 places, and 527 sound classes, respectively. Overall, most of the identified top objects (e.g. TV, closet, sliding door, etc.), places (e.g. dorm, closet, etc.) and sound (e.g. speech, music, etc.) clearly correspond to home environments. This said, a few unexpected results are worth commenting. First, the first place obtained by category "jail cell" in Figure 13b seems strange. However, manual inspection of these images shows that studios with shelves or small rooms can indeed be mistaken with jail cells.

In order to illustrate the kind of content of the Y@N video dataset, we plot four pictures in Figure 12, with the first row as examples of good recognition, and the second row as examples of partly incorrect recognition. For privacy reason, the original Y@N video content cannot be shown. Second, only two sounds (music and speech) are often identified in the audio tracks, while no other sounds seem to be typical to home contexts on weekend nights.

In summary, this section answers RQ1 (consistency of annotation and the main findings from the annotation results and machine-extracted features). In each section of physical/social attributes and ambiances, we present measures, ICC, and main findings . The ICC(2,k) shows that ambiance and physical/social attributes at home (e.g., presentation of home spaces, brightness, music loudness, chatter loudness) can be consistently annotated by external observers. The results also reveal that living room, dining room, kitchen, and bedroom are common places at home where nightlife

Fig. 12. Illustration of similar content to Y@N videos. We use example photos from Pixabay [2] with Pixabay Licence [1] (instead of original examples from Y@N) for privacy reasons: image a by JayMantri [41], image b by JamesDeMers [39], image c by RonPorter [70], and image d by viganhajdari [91]. Images (a-b) are examples of good recognition. They contain the top-5 detected CNN features from 1000-object classes (for image a: ('quilt, comforter, comfort, puff', 'studio couch, day bed', 'wardrobe, closet, press', 'sliding door', 'four poster'); for image b: ('table lamp', 'studio couch, day bed', 'china cabinet, china closet', 'lampshade, lamp shade', 'four poster'); and 365-place classes (for image a: ('youth hostel', 'dorm room', 'bedroom', 'berth', 'hotel room'); for image b: ('living room', 'television room', 'waiting room', 'home theater', 'beauty salon'). Images (c-d) are examples of partly incorrect recognition. They contain the top-5 detected CNN features from 1000-object classes (for image c: ('crate', 'safe', 'chest', 'cradle', 'carton'); for image d: ('china cabinet, china closet', 'toyshop', 'thimble', 'bookcase', 'medicine chest, medicine cabinet', 'tobacco shop, tobacconist shop, tobacconist'); and 365-place classes (for image c: ('jail cell', 'burial chamber', 'dorm room', 'bedchamber', 'stable'); for image d: ('bookstore', 'gas station', 'toyshop','library/indoor', 'storage room'). The top-5 detected objects and scenes of images a-b are more relevant, while the recognized content of images c-d is more irrelevant.

activities like eating/drinking, entertainment (watching TV or using mobile devices) and chatting happen. Young people at home weekend nights seem to be mindful about the loudness of music and level of chatter. In addition, we found a surprisingly large proportion of videos with no people other than the volunteer, engaged in relatively quiet activities. Although the number of videos contains people do not take a large portion, they describe the gender ratio and their activities as well as their reactions to our participants. Moreover, although there are still unexpected results of extracted objects and scenes, many identified CNN-extracted classes from objects, scenes, and sounds are relevant to home environments. To our knowledge, this analysis of nightlife activities at home, which was enabled by the crowdsensing experience, has not been previously reported.

Feature Classes	Frame Level (28K frames)	Video Level (301 videos)
1000 classes	Probability distribution over 1000 object classes (Sum of 1000 classes is 1)	Class-specific aggregate for each video: maximum probability over the set of frames for each class. Purpose: obtain a representation of the objects present in the video.
365 scene categories	Probability distribution over the 365 scene classes (Sum of 365 classes is 1)	Class-specific aggregate for each video: average probability over the set of frames for each class. Purpose: obtain a representation of the most likely scene in the video.
527 sounds	Not available	Probability distribution over the 527 sound classes

Table 5. Visual and sound extracted features for the video dataset.

5 MACHINE-EXTRACTED FEATURES AND AMBIANCE RECOGNITION OF HOME SPACES (RQ2)

This section describes how machine analysis of the audio-visual tracks of videos can be used to characterize and enrich the understanding of youth home spaces on weekend nights.

5.1 Correlation between Machine-extracted Features and Ambiance

This section aims to identify what machine-learning extracted features (1000-object classes, 365-place classes, and 527-sound classes) are correlated with the 11 ambiances categories assessed by the annotators.

5.1.1 Correlation between ambiances and object classes. Correlation results with ambiance are shown in Table 6. Only the largest correlations are shown, (i.e. those higher or equal to 0.25 and with p-value < 0.001). Places described as *comfortable/cozy* have couches and beds present in the videos, while *festive/fun* places were positively correlated with eating places and movie places. These results were confirmed by manual inspection of the videos. We also noted that, in a few cases, participants recorded the TV program they were watching as part of their home space videos. This might explain why dark ambiances are correlated with objects like cinema, but also with seemingly random objects like car mirror or grey fox. This is a known limitation of using CNN models trained on datasets which are not specifically designed for home environments [83]. This could make some unexpected objects recognized and associated. Interestingly, object category *"restaurant, eating house, eating place, eatery"* has a positive association with *festive/fun* ambiance, while has a negative correlation with *dull/simple*, and *serious/boring* ambiance.

5.1.2 Correlation between ambiances and scene classes. Correlations between the 365-scene classes and the 11 ambiances categories are shown in Table 7. Overall, the results show similar associations to those identified in Table 6. For example, a bedroom and a living room associate positively with comfortable ambiance. A dining hall and dining room are positively linked to *large/spacious* ambiance, while pantry or closet do with *cramped/confined* and *dull/simple* ambiance. Results also show that dark and bright ambiances are correlated, negatively and positively, with a large number of scene classes. As mentioned above, participants have sometimes recorded videos of TV programs in dark places at homes, which made the model recognize some places types erroneously, i.e., the places depicted on the TV shows; recall that watching TV was a very popular activity (Figure 8).

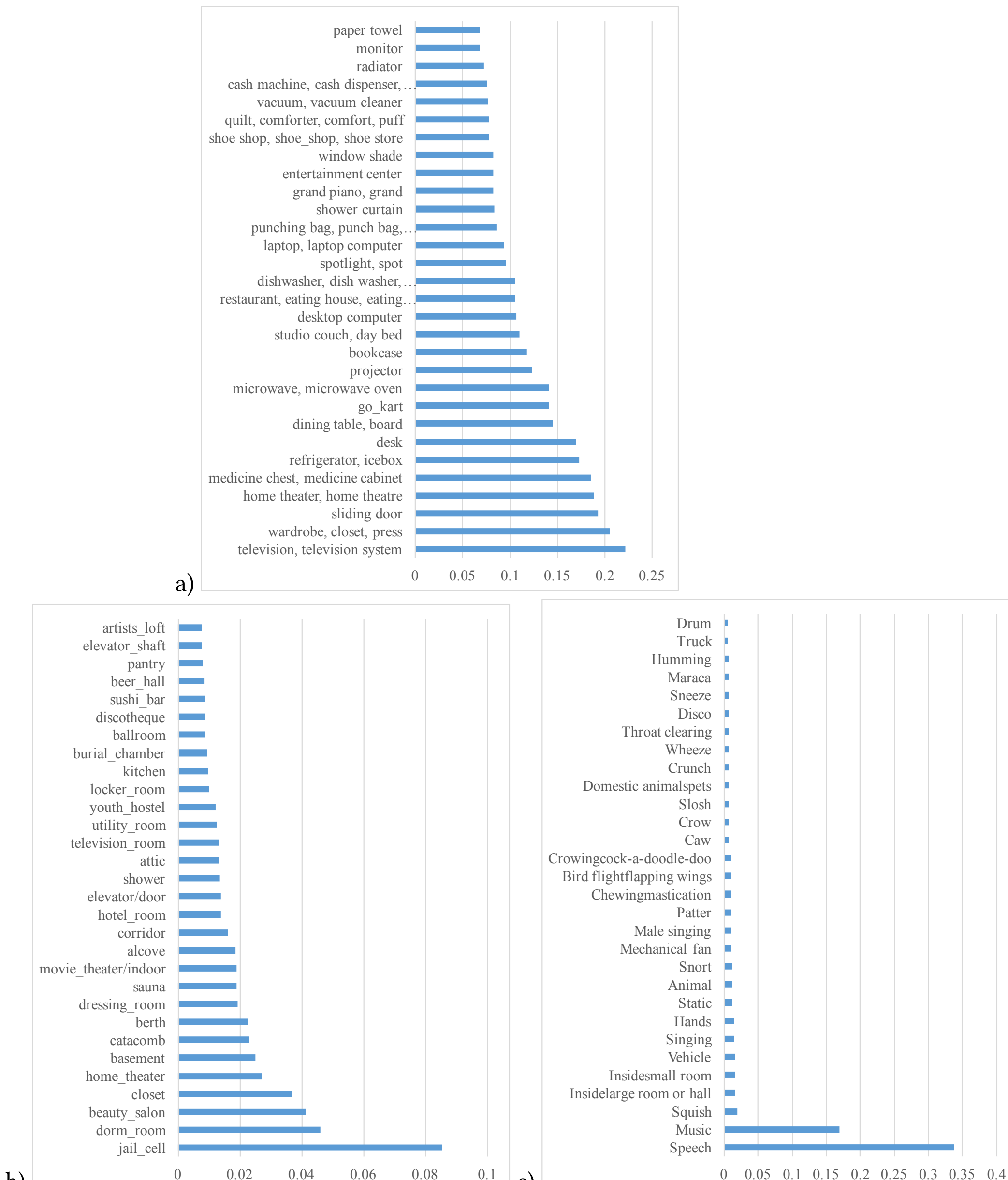


Fig. 13. Top 30 features of (a) 1000 object classes, (b) 365 place classes, and (c) 527 sound classes.

5.1.3 Correlation between ambiances and sound classes. Finally, we examine the correlation between the 527 sound features with ambiances. Only three correlations were above 0.2 with p-value <0.001. In particular, *festive/fun* ambiance are positively associated with *chewing/mastication* (r = 0.21), which might be explained as the same ambiance categories were associated to people eating (see Tables 6 and 7). Also, the correlation of *female singing* (0.20) and *techno* (0.20) (both music-related sounds) with *off-the-beaten-path/unique* ambiance could help explain why that ambiance has

Ambiance Features	1000-object classes
[a] Large, spacious	medicine chest, medicine cabinet(-0.33), refrigerator, icebox(-0.31)
[b] Dark, badly-lit	cinema, movie theater, movie theatre, movie house, picture palace(0.37) , grey fox, gray fox, Urocyon cinereoargenteus(0.32) , suspension bridge(0.30) , hyena, hyaena(0.28) , wing(0.25), badger(0.27) , miniature pinscher(0.26), jack-o-lantern(0.26), desktop computer(0.26), car mirror(0.25) whiptail, whiptail lizard(-0.28), microwave, microwave oven(-0.28)
[c] Colorful, decorated	tobacco shop, tobacconist shop, tobacconist(0.25) restaurant, eating house, eating place, eatery(-0.25)
[d]Cramped, confined	medicine chest, medicine cabinet(0.33), refrigerator, icebox(0.31) grand piano, grand(-0.27)
[e]Bright, well-lit	whiptail, whiptail lizard(0.27), microwave, microwave oven(0.25), dishwasher, dish washer, dishwashing machine(0.25) cinema, movie theater, movie theatre, movie house, picture palace(-0.35), grey fox, gray fox, Urocyon cinereoargenteus(-0.28), suspension bridge(-0.27), badger(-0.26), jack-o-lantern(-0.26), hyena, hyaena(-0.25), theater curtain, theatre curtain(-0.25), wing(-0.25)
[f]Comfortable, cozy	studio couch, day bed(0.29) dishwasher, dish washer, dishwashing machine(-0.28)
[g]Dull, simple	restaurant, eating house, eating place, eatery(-0.29)
[h]Festive, fun	restaurant, eating house, eating place, eatery(0.26), cinema, movie theater, movie theatre, movie house, picture palace(0.25)
[i]Sophisticated, stylish	*
[j]Off-the-beaten-path, unique	lumbermill, sawmill(0.31), dam, dike, dyke(0.27)
[k]Serious, boring	restaurant, eating house, eating place, eatery(-0.26)

Table 6. Pearson correlation between ambiance and 1000-object classes limited to classes with Pearson correlation score >=0.25 and p-value <0.001. Negative and positive correlation values are ranked in descending order by absolute correlation value and are shown in red and blue, respectively. Entries marked with (*) correspond to p-value >0.001 and are not discussed.

positive correlations in home environments with features like bedchamber (0.42), throne room (0.37), or living room (0.33) in Table 7.

5.2 Ambiance Inference

This section presents the investigation of whether and how the ambiance of home places can be automatically inferred using machine-extracted features.

5.2.1 Inference task, method, and performance evaluation. The goal is to infer (in the regression sense) the ambiance of home spaces as perceived by external observers. This inference task uses the aggregated annotations of ambiance discussed in previous sections and is run on the video, which is aggregated as described in the previous section. Random Forest (RF) [15] is used as a regression model in our inference task. By using RF, multiple decision trees are built up to form various classification outputs. In this experiment, we set parameters ntrees = 500 as recommended by [50]. We ensure that the train and test set take 80% and 20%, respectively. We also apply 5-fold cross validation for training phase. After obtaining RF trained models, we quantify the performance by

Ambiance Features	365-place classes
[a] Large, spacious	lobby(0.45) , living room(0.35) , restaurant patio(0.30) , dining room(0.28) , dining hall(0.26), waiting room(0.25) closet(-0.30), pantry(-0.26), clean room(-0.26), shower(-0.25)
[b] Dark, badly-lit	catacomb(0.58) , movie theater/indoor(0.52) , barn door(0.51) , alley(0.49), stage/indoor(0.48), ruin(0.47) , orchestra pit(0.46), auditorium(0.45), arena/performance(0.43) , castle(0.43), elevator shaft(0.43), grotto(0.43), mosque/outdoor(0.42), skyscraper(0.42), tower(0.41), , house(0.41) , courtyard(0.41), aquarium(0.41) , cockpit(0.41), downtown(0.39), music studio(0.38), mausoleum(0.38), tree house(0.37) , fountain(0.37), forest path(0.37), water tower(0.36), palace(0.36), temple/asia(0.36), hotel/outdoor(0.36), motel(0.35), office building(0.35), cottage(0.35), volcano(0.35), pagoda(0.35), plaza(0.35) , mansion(0.34) , throne room(0.34), viaduct(0.33) , canal/urban(0.33) , oast house(0.32), arch(0.32), building facade(0.32), church/outdoor(0.31) , aqueduct(0.31), oilrig(0.30), schoolhouse(0.30), waterfall(0.30), amphitheater(0.30) , cemetery(0.29),tree farm(0.29) , lock chamber(0.29), mountain(0.29) , creek(0.28), landing deck(0.28), formal garden(0.27), diner/outdoor(0.27), forest road(0.27) , village(0.27) , home theater(0.27), chalet(0.27), amusement park(0.27), burial chamber(0.27), harbor(0.26), hardware store(0.26), embassy(0.26) , bridge(0.26), parking lot(0.26), campsite(0.26), kasbah(0.26), windmill(0.26), jail cell(0.25), medina(0.25) laundromat(-0.40), kinder garden classroom(-0.35), art studio(-0.35) , pantry(-0.34), clean room(-0.33), nursery(-0.33), beauty salon(-0.32), hunting lodge/outdoor(0.32), playroom(-0.32), art school(-0.31), utility room(-0.30), art gallery(-0.28), storage room(-0.28), veterinarians office(-0.28), department store(-0.28), bathroom(-0.27), classroom(-0.26), office cubicles(-0.26), garage/indoor(-0.26), pet shop(-0.25), reception(-0.25), artists loft(-0.25)
[c] Colorful, decorated	bazaar/outdoor(0.30), throne room(0.28), bazaar/indoor(0.27), bedchamber(0.27), lobby(0.25)
[d]Cramped, confined	closet(0.30), pantry(0.29),clean room(0.27) living room(-0.38), dining room(-0.25), lobby(-0.41), restaurant patio(-0.26), waiting room(-0.29)
[e]Bright, well-lit	laundromat(0.41), clean room(0.36), kinder garden classroom(0.33), art studio(0.33), nursery(0.32), utility room(0.32), pantry(0.30), beauty salon(0.30), art gallery(0.29), playroom(0.29), veterinarians office(0.28), bathroom(0.27), biology laboratory(0.27), artists loft(0.27), department store(0.27), art school(0.27), physics laboratory(0.26), office cubicles(0.26), dressing room(0.25), garage/indoor(0.25) catacomb(-0.55), barn door(-0.48), movie theater/indoor(-0.48), stage/indoor(-0.46), alley(-0.45), auditorium(-0.44), orchestra pit(-0.43), ruin(-0.42), grotto(-0.41), elevator shaft(-0.40), arena/performance(-0.40), aquarium(-0.39), cockpit(-0.39), skyscraper(-0.38), castle(-0.38), music studio(-0.38), tower(-0.37), mosque/outdoor(-0.37), house(-0.36), courtyard(-0.35), throne room(-0.34), forest path(-0.34), mausoleum(-0.34), volcano(-0.34), tree house(-0.34), downtown(-0.34), fountain(-0.33), water tower(-0.33), temple/asia(-0.33), pagoda(-0.32), hotel/outdoor(-0.32), palace(-0.32), motel(-0.32), cottage(-0.31), plaza(-0.31), arch(-0.31), office building(-0.30), oast house(-0.30), canal/urban(-0.29), mansion(-0.29), waterfall(-0.28), aqueduct(-0.28), mountain(-0.28), cemetery(-0.28), viaduct(-0.28), hunting lodge/outdoor(-0.28), building facade(-0.28), oil rig(-0.27), burial chamber(-0.27), schoolhouse(-0.27), amphitheater(-0.27), landing deck(-0.26), church/outdoor(-0.26), lock chamber(-0.26), amusement park(-0.26), campsite(-0.26), tree farm(-0.26), forest road(-0.26), creek(-0.25), canyon(-0.25), home theater(-0.25)
[f]Comfortable, cozy	living room(0.35), bedroom(0.28), hotel room(0.26) pantry(-0.36), laundromat(-0.31), bedchamber(-0.29), clean room(-0.28)
[g]Dull, simple	alcove(0.26), closet(0.26) lobby(-0.30), throne room(-0.28), living room(-0.25)
[h]Festive, fun	discotheque(0.29), auditorium(0.28), stage/indoor(0.26)
[i]Sophisticated, stylish	lobby(0.37), roof garden(0.31), restaurant patio(0.30), living room(0.28) closet(-0.26)
[j]Off-the-beaten-path, unique	bedchamber(0.42), throne room(0.37), living room(0.33), bazaar/outdoor(0.32), bazaar/indoor(0.30), market/indoor(0.29), diner/outdoor(0.28), lobby(0.28), sandbox(0.28), junkyard(0.27), stable(0.26), pavilion(0.25)
[k]Serious, boring	*

Table 7. Pearson correlation between ambiance and 365-scene classes limited to classes with Pearson correlation score >=0.25 and p-value <0.001. Negative and positive correlation values are ranked in descending order by absolute correlation value and are shown in red and blue, respectively. Entry marked with (*) corresponds to p-value >0.001 and Pearson correlation score <0.25.

Ambiance Features	527 sound classes
[a] Large, spacious	*
[b] Dark, badly-lit	*
[c] Colorful, decorated	*
[d]Cramped, confined	*
[e]Bright, well-lit	*
[f]Comfortable, cozy	*
[g]Dull, simple	*
[h]Festive, fun	Chewing mastication(0.21)
[i]Sophisticated, stylish	*
[j]Off-the-beaten-path, unique	Female singing(0.20), Techno(0.20)
[k]Serious, boring	*

Table 8. Pearson correlation between ambiances and 527-sound classes with Pearson score >=0.20 and p-value <0.001. Entries marked with (*) correspond to p-value >0.001 and are not discussed.

Feature Groups	127 sound classes		1000 object classes		365 scene classes	
	r	R^2	r	R^2	r	R^2
[a] Large, spacious	0.07	0.005	0.47	**0.23**	0.52	**0.27**
[b] Dark, badly-lit	0.08	0.01	0.66	**0.43**	0.83	**0.69**
[c] Colorful, decorated	-0.13	0.02	0.24	0.06	0.31	0.10
[d] Cramped, confined	0.03	0.001	0.44	0.19	0.56	**0.31**
[e] Bright, well-lit	0.02	0.0005	0.67	**0.44**	0.79	**0.63**
[f] Comfortable, cozy	-0.03	0.0007	0.36	0.13	0.46	**0.21**
[g] Dull, simple	0.002	0.000006	0.48	**0.23**	0.44	0.19
[h] Festive, fun (*)	0.17	0.03	0.12	0.01	0.31	0.09
[i] Sophisticated, stylish	0.04	0.001	0.24	0.06	0.25	0.06
[j] Off-the-beaten-path, unique (*)	-0.09	0.008	0.12	0.02	0.28	0.08
[k] Serious, boring (*)	-0.02	0.0005	0.32	0.11	0.37	0.14

Table 9. Inference results including Pearson's correlation coefficient (r), coefficient of determination (R^2). All R^2 with score >= 0.20 are shown in bold font. Rows marked with (*) correspond to ambiance categories that did not reach sufficient annotator agreement (ICC).

using Pearson's correlation coefficient (r), and the coefficient of determination (R^2). In the context of our RF model, R^2 measures how much variance in ambiance is explained by the RF model.

5.2.2 Experiment and results. We randomly divide the 301 videos into two subsets: 80% (241 videos) for training and 20% (60 videos) for testing. We apply RF on 241 videos for training with 5-fold cross validation. The evaluation of RF model is shown in Table 9. We observe that the audio features are not capable of improving over a simple prediction of the mean score ($R^2 \sim 0$). In contrast, using 1000 object classes can infer certain ambiances of home spaces with $R^2 > 0.2$, namely *large/spacious, dark/badly-lit, bright/well-lit, and dull/simple*. The highest R^2 obtained is 0.44 for *bright/well-lit*. Meanwhile, the rest of ambiance categories cannot be inferred by the object representation. Recall that three of these ambiance categories (*festive/fun, serious/boring, off-the-beaten-path/unique*) had not reach sufficient ICC agreement (Table 3), but we decided to include the results for purposes of completeness. Regarding the 365-scene classes, five of the eleven ambiance variables (*large/spacious, dark/badly-lit, bright/well-lit, cramped/confined, and comfortable/cozy*) are predicted by using 365-scene classes with $R^2 > 0.2$ ($R^2 = 0.69$ for dark/badly-lit). In Section 5, we discussed the correlation of ambiances and scenes. Clearly, certain scenes can predict those ambiances related to space capacity (*large/spacious vs cramped/confined*), and brightness (*bright/well-lit vs dark/badly-lit*). For *comfortable/cozy* ambiance, Section 5 also showed that living room with couch, and bedroom with bed, have positive correlation. Interestingly, two of the ambiance variables (*colorful/decorated* and *sophisticated/stylish*) could not be inferred by any of the visual representations, regardless of the fact that they achieved good inter-annotator reliability, (0.66 and 0.65, respectively, see Table 3).

In summary, we use RF to train a regression model and use R^2 as the main measure to evaluate which features can predict the ambiance of a home space. Our findings show that six of the ambiance categories can be inferred with R^2 in the $[0.21, 0.69]$ range (four with object-based features, and five with scene-level features), and with higher R^2 values when a scene deep network is used. More specifically, space capacity (*large/spacious vs cramped/confined*), brightness (*bright/well-lit vs dark/badly-lit*), *comfortable/cozy*, and *dull/simple* can be predicted by object-level and scene-level description. In contrast, audio features were not effective at inferring ambiance.

6 DISCUSSION AND IMPLICATIONS

Table 10 summarizes our main findings for RQ1 and RQ2. We now discuss the results and some of their implications.

6.1 A Unique video Dataset of Home Spaces

In terms of data source, collecting data on home environments via crowdsourced videos is novel in comparison to previous work using social media sources. This includes research on Foursquare, which showed that users underreported home presence by checking into homes considerably less frequently than into other places, given the logic of such social network [55], [22]; and also includes recent work on Airbnb, which is known to feature photos of homes that are taken with the explicit purpose of attracting possible guests, in some cases taken by professional photographers [60]. Our study used 301 ten-second video clips of young people's home spaces on weekend nights. To our knowledge, this is a unique dataset of real-life home environments that cannot be compared to any other publicly available dataset, in the sense that participants' showed their home spaces simply as they are (with no artistic filters or advertising intentions) on their weekend nights.

From the total set of 843 videos collected in the study, slightly less than one third were consistently identified as representing homes by the five annotators. Yet, this does not mean that only one-third of the nights were spent at home, but rather it can be seen as a consequence of the study design, which requested participants to provide only one video per night if they did not change location during the night. Given that about half of all drinks (non-alcoholic or alcoholic) in the Youth@Night dataset were documented in homes [48], this result suggests that participants were less likely to change locations when starting the night at home than when going out [46], highlighting the relevance to research and understand what happens in this usually hidden or hard-to-reach kind of environment. In addition, the levels of inter-annotator agreement for most of the physical attributes at homes were globally good to excellent. This result echoes previous work in psychology [30] that found that personal environments elicit similar impressions from independent observers, while adding the novel angle of using short video as stimuli (rather than photos). This result also indicates that, despite being relatively short, 10-second videos are long enough to provide adequate cues of the physical and social environment, the ongoing activities, and the ambiance.

6.2 Home as a Nightlife Space

As mentioned above, about one third of the Y@N videos were recorded in homes, and participants were less likely to change locations when starting the night at home than when going out. This highlights the need to understand this particular environment. Qualitative feedback from the participants at the end of the fieldwork echoed previous research that has found that homes can serve both as 'prequel night out spaces', where young people meet, dress up, and get ready for the night out, as well as a standalone nightlife space where they hang out with friends or have parties [51]. For one participant, home was his main nightlife destination: "Now that I study in Lausanne and live here, when I go out it's really to other people's place or at my place. Which still does not prevent me from going out [to pubs and clubs] now and then". Another participant mainly conceived home as the starting point of the night: "Well, when I go out, I prefer drinking before going out, well, not before going out but, let's say we meet with friends and we go to someone's place to drink or just eat and we drink something, or in a park during the summer, yeah, let's say I start drinking [in a residential neighborhood] and then we move on and continue the party downtown". Finally, several participants considered the home as an alternative to commercial nightlife venues: "For me, there are two types of nights out: the dancing ones, when we go to clubs and the point is to dance [...] and then there are the quiet ones, when we just sit, at someone's place or in a bar, and we talk and that's it" or "There are different kinds of nights out. Sometimes, people want to go out to meet others and that's it, it all depends on the mood we are in that night. It's true that sometimes we enjoy staying with friends and have big parties in homes, or go out in the city, but as a small group."

In order to better represent home environments, the annotation task developed for this study revealed detailed attributes of physical and social environment, including the types of rooms attended, levels of brightness, loudness and occupancy, the number and gender of people, and the ongoing activities. Altogether, this information provides a comprehensive picture of young people's nightlife environments. Specifically, we examined co-occurrence between activities and levels of occupancy and types of spaces at home. The authors of [7], [88], [9] studied usage of domestic spaces that were used in daily life activities, and specific psychological states (e.g., mental stress). In our research, home spaces were analyzed from the perspective of activities of young people on weekend nights. Through physical and social attributes, we have insights of activities in the context of Swiss young people (16-25 year-old), who present differences to other populations, e.g. in the US, where legal drinking age and norms about the use of the public space differ from those in Europe. We found that young people spent weekend night time watching TV, listening to music [77], and playing games. Previous findings about pre-drinking before going out in [95] or drinking at friends' or family's homes [35] were also partly shown in our work.

One particular instrument of the annotation task, the ambiance scale, aimed to capture the different dimensions of this construct. Dimensions related to the physical space (e.g. large, spacious), which could be rated rather objectively by the annotators, showed a high degree of agreement among them. Dimensions relating to the personal evaluation of the annotators (e.g. off-the-beaten-path) were indeed more subjective and showed a lower degree of agreement among the annotators. From the correlation analysis, three main groups of ambiances were identified: positively perceived characteristics (large, colorful, festive, stylish, and unique), negatively perceived characteristics (cramped, simple, and boring) and independent characteristics (dark and bright). In addition, the main types of ongoing activities were consistent across ambiance categories (drinking, chatting, watching TV, and computer device use), and small variations were found, e.g. less TV watching for the fun/festive ambiance.

While the aim of the annotation task was to describe home spaces from the perspective of human annotators, the aim of the machine learning task was to observe home spaces through automatic-extracted features using CNNs models, [83], [98], [27]. Thus, without using external annotation of physical and social attributes, the latter task was able to automatically describe home spaces by observing the probability distribution of visual and audio labels. Correlation results between automatically extracted features based on the image frames of the videos and ambiance labels provided promising results for the visual cues (i.e. objects or scenes) from the videos. Yet, results also showed that the existing classes are made to recognize all kinds of objects or situations, even some that are not supposed to be in homes, such as jail cell, car parts, etc. Future research is clearly needed for the development of a specialized dictionary of classes focused on home environments.

Regarding automatic-extracted features based on the soundtrack of the videos, however, only two of the sounds dominated the dataset (speech and music), and thus only a few associations were found with ambiance features. These might be related to the way audio was recorded, but also because homes at night are generally quiet or because not enough information was found in the sound measure in [27].

6.3 Feasibility of Ambiance Inference

We examined the use of machine-extracted features, i.e., 527-sound, 100-object, 365-scene features, for automatic inference of ambiance. As a result, *large, dark, bright, confined, comfortable*, and *simple* ambiances could be inferred by using object and scene classes. These ambiances could be also be perceived by people while unique, festive, and boring ambiances could be ambiguous when being annotated by humans. Comparing these inference results to those reported in [60] on Airbnb home photos, our results corroborate that ambiances closer to physical attributes reach better recognition performance, although the performance we obtained is lower than that obtained on Airbnb data for three variables (large, comfortable, and simple), similar for one variable (confined); and higher for two variables (bright, dark). Note that in addition to the datasets being different, the specific CNN models and the CNN outputs used as features are different too (last convolutional layer in [60] vs. final output equal to the number of objects or scenes in our work). Note also that we made this choice in order to interpret the CNN-derived features in the correlation analysis in Section 5.1. For future work, we believe that regression performance could be improved by CNN adaptation, i.e., by fine-tuning the last CNN layers to the ambiance target class as demonstrated in other visual tasks [56]. Home ambiance recognizers built around short duration mobile videos could be advantageous as they might in general contain more information than still images, and used in future applications, as discussed in the next subsection.

6.4 Implications for CSCW research

We conclude this section by discussing some of the implications of our work for CSCW and social computing research.

Understanding youth practices at home from mobile crowdsourced data. Using crowdsourced personal videos as input, we showed that a mixed methodology combining manual annotation and automatically extracted features enabled an in-depth study of youth personal spaces on weekend nights with respect to physical attributes, activities, and social attributes, including joint patterns of activities and places. Crowdsourced visual datasets like the one used here complement another common source of data used in CSCW research, namely social media like Instagram. While early research showed that the home environment was infrequently reported or talked about by users [55], [22], future research could investigate whether certain sub-communities specifically depict nightlife in private spaces, and what specific practices are promoted or enacted around this theme, including ephemerality, self-representation, and sociality. This investigation would require the use of mixed methods of inquiry, combining machine analyses with user interviews and surveys. Furthermore, given that the concept of nightlife is broad and encompasses both the private and public spheres, a second promising line of future work could investigate the interplay between private and public spaces in urban nightlife, and how this is expressed digitally both in crowdsourced campaigns and social media. For instance, recent qualitative work showed that several participants in the Youth@Night campaign coordinated nightlife activities via Whatsapp [85]. This research could benefit from previous CSCW literature on coordination of action and social participation.

Applications of home ambiance recognition. Our work on recognition of ambiance at home also has potential implications for future CSCW work. First, it is evident from our study that state-of-art deep visual learning systems, while useful, still generate erroneous visual descriptors. We believe that it is important to make these limitations explicit to inform other CSCW researchers who plan to use deep learning as a toolbox for their future work. At the same time, in a fast-moving domain, it is not unreasonable to expect progress that could mitigate some of the current limitations, and thus to anticipate that the shown recognition performance will be improved (e.g., Facebook has published results on deep learning models trained on 1 billion Instagram images) [57]. With this, one could envision applications in home supporting systems. Homes are reconfigurable spaces, in which certain elements can be readily changed (decoration, spatial organization of furniture, light, and music). A system able to recognize ambiance could also make recommendations of suitable ambiances at home for specific activities, e.g. to promote socialization. This kind of work would require human-centered approaches to design such prototypes, integrating perspectives of privacy, ethics, and transparency, all of whom are active topics of investigation in CSCW and social computing [99], [8], [96], [49].

Future work on human factor in home research. In CSCW research, human factors play important roles requiring interdisciplinary researchers (social psychologists, sociologists, anthropologists, computer scientists) to find appropriate methods for an individual or group to adopt technology into their daily activities. In this case, technology plays a supporting role while the human factor plays a central one. In our work, we focused on home environments and inferring home ambiance from videos which contributes to CSCW applications, while adding to themes that are relevant to CSCW. Besides physical and social attributes, emotional states and nightlife behaviors and their links to ambiances could need the expertise of other researchers [63], [76]. In future work, youth practices at home and ambiances, technologists could collaborate with specialists in interior decoration art, or psychologists, to build systems to support people to link their home ambiances to their current emotions as well as their behaviors. Beyond building this technology, users would increase their self-awareness about their home ambiances and their own behaviors to promote positive changes and share them with others.

7 CONCLUSION

In this paper, we presented an original study of the characteristics of night personal spaces, including manual coding of places, machine extraction of acoustic and visual description of places, and inference of ambiance of homes of young people in the weekend night setting. We conclude by revisiting the research questions posed at the beginning of the paper.

RQ1: *Given crowdsourced videos recorded at home spaces by young people at night, what patterns of physical and ambiance attributes of youth home spaces can be revealed by manual coding of videos using external annotators and machine-extracted features?* By describing measures, discussing ICC, and showing results, we sequentially analyzed the problem from physical/social attributes (home spaces, brightness, loudness, human presence, activities) to ambiances. We observed co-occurrence between activities and spaces at homes as well as ambiances. Then, we showed that ambiances could be grouped into two clusters: "unlike" characteristics with *serious/boring*, *cramped/confined*, *dull/simple*, and "like" characteristics with *Large/spacious*, *colorful/decorated*, *comfortable/cozy*, *sophisticated/stylish*, *off-the-beaten-path/unique*, *festive/fun*. Finally, we used state-of-the-art pre-trained deep learning models to extract automatic features to represent videos, namely objects, scenes, and sounds. Most machine-extracted classes relevantly characterize home environments, but there were some unexpected features.

RQ2: *What do machine-extracted features of videos reveal about physical attributes of youth home spaces? Can these machine-extracted features infer the perceived ambiance of such spaces?* Correlations between ambiance and automatic features potentially show the feasibility of using machine-extracted features to automatically describe home spaces, although there are certain limitations. Regarding the inference task, ambiances like space capacity (*large/spacious* vs. *cramped/confined*), brightness (*bright/well-lit* vs. *dark/badly-lit*), *comfortable/cozy*, and *dull/simple* can be inferred for private spaces in the weekend nights by using 1000 object classes and 365 scene classes. The total number of videos (N=301) could be a limitation for mo0del training in the automatic inference experiments. However, our results show that six of the ambiance categories can be inferred with R^2 in the $[0.21, 0.69]$ range, and with higher R^2 values when a scene deep network is used.

RQ	Factors	Message
RQ1 - Physical and Social Attributes	Home spaces	The most attended type of room is the living room; followed by bedroom; kitchen/dining room were also frequently attended rooms at night
	Brightness	It tends to reduce from early night to late night
	Music loudness	Videos contained no music on 76% of all situations
	Chatter loudness	Home are mostly quiet with slight increase from 8PM to 11PM
	Occupancy and number of people present	Around 60% of videos contained people gathering from 8 PM to 11 PM; then reducing after 11 PM
	Gender	A gender-matching pattern is evident: female participants tend to gather more with other women, and conversely for male participants. Mixed groups, however, also occur.
	Activities	Drinking, chatting, watching TV, using smartphones/computer, and eating are the most popular activities of young people on weekend nights.
RQ1 - Ambiances	Agreement on ambiance	8 of the 11 ambiance variables have ICCs above 0.5.
	Correlation between ambiances	Place ambiances are grouped on two main opposite dimensions, namely places seen as large, colorful, comfortable, festive, stylish, unique; versus places seen as confined, simple, boring. Dark and bright ambiances do not have a significant correlation with the rest of ambiances.
RQ1 - Machine-extracted Features	Automatic description	1000-object, 365-scene, 527-sound auto-extracted features can express ambiances but with a certain level of noise, because labels of these classes for CNN models are not specifically designed for homes.
RQ2 - Ambiance Regression	Correlation between ambiance and automatic descriptions	Although there are some limitations on the labels of CNNs model, automatic-extracted features have reasonable correlation with ambiances.
	Regression Performances	Six of the ambiance variables (large, dark, bright, confined, comfortable, simple) can be inferred by using object and scene features with coefficient of determination above 0.2. For the other five variables (including three with low ICC). regression performance is low.

Table 10. Summary of findings related to our two RQs.

ACKNOWLEDGMENTS

This work has been funded by the Swiss National Science Foundation through the Dusk2Dawn Sinergia project, and a Swiss Government Excellence Scholarship.

REFERENCES

[1] [n.d.]. Pixabay Licence. Retrieved Aug 19, 2019 from https://pixabay.com/service/license/
[2] [n.d.]. Pixabay Website. Retrieved Aug 19, 2019 from https://pixabay.com/
[3] 2019. Definition of home in English. Retrieved June 22, 2019 from https://dictionary.cambridge.org/dictionary/english/home#cald4-1-1-2
[4] Joan Abbott-Chapman and Margaret Robertson. 1999. Home as a private space: some adolescent constructs. *Journal of youth studies* 2, 1 (1999), 23–43.
[5] Joan Abbott-Chapman and Margaret Robertson. 2001. Youth, leisure and home: Space, place and identity. *Loisir et société/Society and Leisure* 24, 2 (2001), 485–506.
[6] Akile Ahmet. 2013. Home sites: the location (s) of 'home'for young men. *Urban Studies* 50, 3 (2013), 621–634.

[7] Saeid Alitajer and Ghazaleh Molavi Nojoumi. 2016. Privacy at home: Analysis of behavioral patterns in the spatial configuration of traditional and modern houses in the city of Hamedan based on the notion of space syntax. *Frontiers of Architectural Research* 5, 3 (2016), 341–352.
[8] Karla Badillo-Urquiola, Yaxing Yao, Oshrat Ayalon, Bart Knijnenurg, Xinru Page, Eran Toch, Yang Wang, and Pamela J Wisniewski. 2018. Privacy in Context: Critically Engaging with Theory to Guide Privacy Research and Design. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 425–431.
[9] Lynne Baillie and David Benyon. 2008. Place and technology in the home. *Computer Supported Cooperative Work (CSCW)* 17, 2-3 (2008), 227–256.
[10] Rick Bell, Herbert L Meiselman, Barry J Pierson, and William G Reeve. 1994. Effects of adding an Italian theme to a restaurant on the perceived ethnicity, acceptability, and selection of foods. *Appetite* 22, 1 (1994), 11–24.
[11] Mark A Bellis and Karen Hughes. 2011. Getting drunk safely? Night-life policy in the UK and its public health consequences. *Drug and alcohol review* 30, 5 (2011), 536–545.
[12] Yassir Benkhedda, Darshan Santani, and Daniel Gatica-Perez. 2017. Venues in social media: Examining ambiance perception through scene semantics. In *Proceedings of the 2017 ACM on Multimedia Conference*. ACM, 1416–1424.
[13] Alison Blunt. 2005. Cultural geography: cultural geographies of home. *Progress in human geography* 29, 4 (2005), 505–515.
[14] Danah Boyd. 2014. *It's complicated: The social lives of networked teens*. Yale University Press.
[15] Leo Breiman. 2001. Random forests. *Machine learning* 45, 1 (2001), 5–32.
[16] Gulcan Can, Yassir Benkhedda, and Daniel Gatica-Perez. 2018. Ambiance in Social Media Venues: Visual Cue Interpretation by Machines and Crowds. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2363–2372.
[17] Dana R Carney, John T Jost, Samuel D Gosling, and Jeff Potter. 2008. The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology* 29, 6 (2008), 807–840.
[18] Paul Chatterton and Robert Hollands. 2002. Theorising urban playscapes: producing, regulating and consuming youthful nightlife city spaces. *Urban studies* 39, 1 (2002), 95–116.
[19] Yohan Chon, Nicholas D Lane, Yunjong Kim, Feng Zhao, and Hojung Cha. 2013. Understanding the coverage and scalability of place-centric crowdsensing. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 3–12.
[20] Yohan Chon, Nicholas D Lane, Fan Li, Hojung Cha, and Feng Zhao. 2012. Automatically characterizing places with opportunistic crowdsensing using smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 481–490.
[21] Cary C Countryman and SooCheong Jang. 2006. The effects of atmospheric elements on customer impression: the case of hotel lobbies. *International Journal of Contemporary Hospitality Management* 18, 7 (2006), 534–545.
[22] Henriette Cramer, Mattias Rost, and Lars Erik Holmquist. 2011. Performing a check-in: emerging practices, norms and'conflicts' in location-sharing using foursquare. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services*. ACM, 57–66.
[23] Jakob Demant and Sara Landolt. 2014. Youth drinking in public places: The production of drinking spaces in and outside nightlife areas. *Urban Studies* 51, 1 (2014), 170–184.
[24] Cambridge Dictionary. 2018. Definition of ambiance in English. Retrieved November 11, 2018 from https://dictionary.cambridge.org/dictionary/english/ambience
[25] Cameron Duff. 2012. Accounting for context: Exploring the role of objects and spaces in the consumption of alcohol and other drugs. *Social & Cultural Geography* 13, 2 (2012), 145–159.
[26] Ffmpeg. 2010. FFMPEG. http://www.ffmpeg.org
[27] Jort F Gemmeke, Daniel PW Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R Channing Moore, Manoj Plakal, and Marvin Ritter. 2017. Audio set: An ontology and human-labeled dataset for audio events. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*. IEEE, 776–780.
[28] SD Gosling, Robert Gifford, and Lindsay J McCunn. 2013. The selection, creation, and perception of interior spaces: An environmental psychology approach. (2013).
[29] Samuel D Gosling, Kenneth H Craik, Nicholas R Martin, and Michelle R Pryor. 2005. The personal living space cue inventory: An analysis and evaluation. *Environment and Behavior* 37, 5 (2005), 683–705.
[30] Samuel D Gosling, Sei Jin Ko, Thomas Mannarelli, and Margaret E Morris. 2002. A room with a cue: Personality judgments based on offices and bedrooms. *Journal of personality and social psychology* 82, 3 (2002), 379.
[31] Lindsay T Graham and Samuel D Gosling. 2011. Can the ambiance of a place be determined by the user profiles of the people who visit it?. In *ICWSM*.
[32] Lindsay T Graham, Samuel D Gosling, and Christopher K Travis. 2015. The psychology of home environments: A call for research on residential space. *Perspectives on Psychological Science* 10, 3 (2015), 346–356.

[33] Tali Hatuka and Eran Toch. 2017. Being visible in public space: The normalisation of asymmetrical visibility. *Urban Studies* 54, 4 (2017), 984–998.

[34] Dolores Hayden. 2004. *Building suburbia: Green fields and urban growth, 1820-2000*. Vintage.

[35] Sarah L Holloway, Mark Jayne, and Gill Valentine. 2008. 'Sainsbury's is my local': English alcohol policy, domestic drinking practices and the meaning of home. *Transactions of the Institute of British Geographers* 33, 4 (2008), 532–547.

[36] Mark Holton. 2016. Living together in student accommodation: performances, boundaries and homemaking. *Area* 48, 1 (2016), 57–63.

[37] Phil Hubbard. 2008. Regulating the social impacts of studentification: a Loughborough case study. *Environment and Planning A* 40, 2 (2008), 323–341.

[38] Tapio Ikkala and Airi Lampinen. 2015. Monetizing network hospitality: Hospitality and sociability in the context of Airbnb. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. ACM, 1033–1044.

[39] JamesDeMers. [n.d.]. An example of living room on Pixabay. Retrieved Aug 19, 2019 from https://pixabay.com/photos/farmhouse-bedroom-old-515619/

[40] Margaretha Järvinen and Jeanette Østergaard. 2009. Governing adolescent drinking. *Youth & society* 40, 3 (2009), 377–402.

[41] JayMantri. [n.d.]. An example of bedroom on Pixabay. Retrieved Aug 19, 2019 from https://pixabay.com/photos/bedroom-sleeping-bed-furniture-405920/

[42] Hina Keval and M Angela Sasse. 2008. to catch a thief–you need at least 8 frames per second: the impact of frame rates on user performance in a CCTV detection task. In *Proceedings of the 16th ACM international conference on Multimedia*. ACM, 941–944.

[43] Iman Khajehzadeh and Brenda Vale. 2015. How do people use large houses. In *Living and Learning: Research for a Better Built Environment: Proceeding of the 49th International Conference of the Architectural Science Association 2015*. 153–162.

[44] Terry K Koo and Mae Y Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* 15, 2 (2016), 155–163.

[45] Philip Kotler. 1973. Atmospherics as a marketing tool. *Journal of retailing* 49, 4 (1973), 48–64.

[46] Florian Labhart, Kathryn Graham, Samantha Wells, and Emmanuel Kuntsche. 2013. Drinking before going to licensed premises: An event-level analysis of predrinking, alcohol consumption, and adverse outcomes. *Alcoholism: Clinical and Experimental Research* 37, 2 (2013), 284–291.

[47] Florian Labhart, Darshan Santani, Jasmine Truong, Flavio Tarsetti, Olivier Bornet, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2017. Development of the Geographical Proportional-to-size Street-Intercept Sampling (GPSIS) method for recruiting urban nightlife-goers in an entire city. *International Journal of Social Research Methodology* 20, 6 (2017), 721–736.

[48] Florian Labhart, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez, and Emmanuel Kuntsche. 2019. Capturing drinking and nightlife behaviours and their social and physical context with a smartphone application - investigation of users' experience and reactivity. *Addiction Research and Theory* (2019). http://infoscience.epfl.ch/record/263840

[49] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 102.

[50] Andy Liaw, Matthew Wiener, et al. 2002. Classification and regression by randomForest. *R news* 2, 3 (2002), 18–22.

[51] Siân Lincoln. 2012. *Youth culture and private space*. Springer.

[52] Siân Lincoln. 2014. "I've Stamped My Personality All Over It" The Meaning of Objects in Teenage Bedroom Space. *Space and Culture* 17, 3 (2014), 266–279.

[53] Siân Lincoln. 2015. 'My Bedroom Is Me': Young People, Private Space, Consumption and the Family Home. In *Intimacies, Critical Consumption and Diverse Economies*. Springer, 87–106.

[54] Siân Lincoln and Brady Robards. 2016. Being strategic and taking control: Bedrooms, social network sites and the narratives of growing up. *new media & society* 18, 6 (2016), 927–943.

[55] Janne Lindqvist, Justin Cranshaw, Jason Wiese, Jason Hong, and John Zimmerman. 2011. I'm the mayor of my house: examining why people use foursquare-a social-driven location sharing application. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2409–2418.

[56] Mingsheng Long, Yue Cao, Jianmin Wang, and Michael I Jordan. 2015. Learning transferable features with deep adaptation networks. *arXiv preprint arXiv:1502.02791* (2015).

[57] Dhruv Mahajan, Ross Girshick, Vignesh Ramanathan, Kaiming He, Manohar Paluri, Yixuan Li, Ashwin Bharambe, and Laurens van der Maaten. 2018. Exploring the limits of weakly supervised pretraining. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 181–196.

[58] Denise Marquard. 2014. Zurich, the Party Town. http://bit.ly/28M8nbk Accessed: July 15, 2016.

[59] Hugh Matthews, Melanie Limb, and Barry Percy-Smith. 1998. Changing worlds: the microgeographies of young teenagers. *Tijdschrift voor economische en sociale geografie* 89, 2 (1998), 193–202.

[60] Laurent Son Nguyen, Salvador Ruiz-Correa, Marianne Schmid Mast, and Daniel Gatica-Perez. 2018. Check out this place: Inferring ambiance from airbnb photos. *IEEE Transactions on Multimedia* 20, 6 (2018), 1499–1511.

[61] City of Lausanne. 2010. Drinking in Public Space in City of Lausanne: A Challenge for Police. http://bit.ly/28M8X8X Accessed: July 15, 2016.

[62] Vicente Ordonez and Tamara L Berg. 2014. Learning high-level judgments of urban perception. In *European Conference on Computer Vision*. Springer, 494–510.

[63] Stephen A Petrill, Alison Pike, Tom Price, and Robert Plomin. 2004. Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. *Intelligence* 32, 5 (2004), 445–460.

[64] Thanh-Trung Phan and Daniel Gatica-Perez. 2017. Healthy# fondue# dinner: analysis and inference of food and drink consumption patterns on instagram. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. ACM, 327–338.

[65] Lorenzo Porzi, Samuel Rota Bulò, Bruno Lepri, and Elisa Ricci. 2015. Predicting and understanding urban perception with convolutional neural networks. In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 139–148.

[66] Daniele Quercia, Neil Keith O'Hare, and Henriette Cramer. 2014. Aesthetic capital: what makes London look beautiful, quiet, and happy?. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 945–955.

[67] Miriam Redi, Luca Maria Aiello, Rossano Schifanella, and Daniele Quercia. 2018. The Spirit of the City: Using Social Media to Capture Neighborhood Ambiance. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 144.

[68] Miriam Redi, Daniele Quercia, Lindsay T Graham, and Samuel D Gosling. 2015. Like partying? your face says it all. predicting the ambiance of places with profile pictures. *arXiv preprint arXiv:1505.07522* (2015).

[69] Harriet L Rheingold and Kaye V Cook. 1975. The contents of boys' and girls' rooms as an index of parents' behavior. *Child development* (1975), 459–463.

[70] RonPorter. [n.d.]. An example of bedroom on Pixabay. Retrieved Aug 19, 2019 from https://pixabay.com/photos/living-room-great-room-670240/

[71] Allison Sall and Rebecca E Grinter. 2007. Let's get physical! In, out and around the gaming circle of physical gaming at home. *Computer Supported Cooperative Work (CSCW)* 16, 1-2 (2007), 199–229.

[72] Darshan Santani, Joan-Isaac Biel, Florian Labhart, Jasmine Truong, Sara Landolt, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2016. The night is young: urban crowdsourcing of nightlife patterns. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 427–438.

[73] Darshan Santani and Daniel Gatica-Perez. 2013. Speaking swiss: languages and venues in foursquare. In *Proceedings of the 21st ACM international conference on Multimedia*. ACM, 501–504.

[74] Darshan Santani and Daniel Gatica-Perez. 2015. Loud and trendy: Crowdsourcing impressions of social ambiance in popular indoor urban places. In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 211–220.

[75] Darshan Santani, Rui Hu, and Daniel Gatica-Perez. 2016. InnerView: Learning place ambiance from social media images. In *Proceedings of the 2016 ACM on Multimedia Conference*. ACM, 451–455.

[76] Klaus R Scherer, Marcel R Zentner, et al. 2001. Emotional effects of music: Production rules. *Music and emotion: Theory and research* 361, 2001 (2001), 392.

[77] Diane J Schiano, Ame Elliott, and Victoria Bellotti. 2007. A look at Tokyo youth at leisure: Towards the design of new media to support leisure outings. *Computer Supported Cooperative Work (CSCW)* 16, 1-2 (2007), 45–73.

[78] Robert Shaw. 2015. Controlling darkness: self, dark and the domestic night. *cultural geographies* 22, 4 (2015), 585–600.

[79] Patrick E Shrout and Joseph L Fleiss. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin* 86, 2 (1979), 420.

[80] Ernestasia Siahaan, Alan Hanjalic, and Judith Redi. 2016. A reliable methodology to collect ground truth data of image aesthetic appeal. *IEEE Transactions on Multimedia* 18, 7 (2016), 1338–1350.

[81] Tracey Skelton and Gill Valentine. 2005. Cool places. In *Cool Places*. Routledge, 11–42.

[82] Nanette Stroebele and John M De Castro. 2004. Effect of ambience on food intake and food choice. *Nutrition* 20, 9 (2004), 821–838.

[83] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2818–2826.

[84] Eran Toch and Inbal Levi. 2013. Locality and privacy in people-nearby applications. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 539–548.
[85] Jasmine Truong. 2018. Attending to others: how digital technologies direct young people's nightlife. *Geographica Helvetica* 73, 2 (2018), 193–201.
[86] Jasmine Truong. 2018. Collapsing contexts: social networking technologies in young people's nightlife. *Children's Geographies* 16, 3 (2018), 266–278.
[87] Yi-Fu Tuan. 1971. Geography, phenomenology, and the study of human nature. *Canadian Geographer/Le Géographe canadien* 15, 3 (1971), 181–192.
[88] Ian Tucker. 2010. Everyday spaces of mental distress: the spatial habituation of home. *Environment and Planning D: Society and Space* 28, 3 (2010), 526–538.
[89] Gill Valentine, Sarah Holloway, Charlotte Knell, and Mark Jayne. 2008. Drinking places: Young people and cultures of alcohol consumption in rural environments. *Journal of Rural Studies* 24, 1 (2008), 28–40.
[90] Ilse Van Liempt, Irina Van Aalst, and Tim Schwanen. 2015. Introduction: Geographies of the urban night.
[91] viganhajdari. [n.d.]. An example of living room on Pixabay. Retrieved Aug 19, 2019 from https://pixabay.com/photos/interior-decoration-home-design-1519595/
[92] He Wang, Dimitrios Lymberopoulos, and Jie Liu. 2014. Local business ambience characterization through mobile audio sensing. In *Proceedings of the 23rd international conference on World wide web*. ACM, 293–304.
[93] Clifton Edwin Watts. 2004. Predicting Free Time Activity Involvement of Adolescents: The Influence of Adolescent Motivation, Adolescent Initiative, and Perceptions of Parenting. (2004).
[94] Samantha Wilkinson. 2017. Drinking in the dark: shedding light on young people's alcohol consumption experiences. *Social & Cultural Geography* 18, 6 (2017), 739–757.
[95] Samantha Wilkinson and Catherine Wilkinson. 2018. Night-Life and Young People's Atmospheric Mobilities. *Mobile Culture Studies. The Journal* 3, 2017 (2018), 77–96.
[96] Richmond Y Wong, Deirdre K Mulligan, Ellen Van Wyk, James Pierce, and John Chuang. 2017. Eliciting values reflections by engaging privacy futures using design workbooks. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 111.
[97] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2017. Places: A 10 million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017).
[98] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2018. Places: A 10 million image database for scene recognition. *IEEE transactions on pattern analysis and machine intelligence* 40, 6 (2018), 1452–1464.
[99] Jonathan Zong and J Nathan Matias. 2018. Automated Debriefing: Interface for Large-Scale Research Ethics. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 21–24.

Received April 2019; revised June 2019; accepted August 2019