

Social Psychological Bulletin

Psychologia Społeczna

A Categorization of Behaviors Reported in Experience Sampling Studies

Ewa Skimina^a, Dominika Karaś^a, Ewa Topolewska-Siedzik^a, Maria Kłym-Guba^a,

Klaudia Ponikiewska^a, Radosław Rogoza^a, Eldad Davidov^{bc}, Jan Ciecich^{ab}

[a] *Institute of Psychology, Cardinal Stefan Wyszyński University in Warsaw, Warsaw, Poland.* [b] *University Research Priority Program Social Networks, University of Zurich, Zurich, Switzerland.* [c] *Faculty of Management, Economics and Social Sciences, University of Cologne, Cologne, Germany.*

Social Psychological Bulletin, 2020, Vol. 15(2), Article e3029, <https://doi.org/10.32872/spb.3029>

Received: 2019-10-15 • Accepted: 2020-04-03 • Published (VoR): 2020-07-31



Handling Editor: Ewa Gruszczyńska, SWPS University of Social Sciences and Humanities, Warsaw, Poland

Corresponding Author: Jan Ciecich, Institute of Psychology, Cardinal Stefan Wyszyński University in Warsaw, Wóycickiego 1/3 bud. 14; 01-938 Warszawa, Poland. E-mail: j.ciecich@uksw.edu.pl

Supplementary Materials: Data [see [Index of Supplementary Materials](#)]



Abstract

Experience sampling is considered one of the best methods for measuring behavior (Furr, 2009, <https://doi.org/10.1002/per.724>). When used for this purpose, it requires a coding system to transform diversified reports on what people are doing, provided as responses to an open-ended question, into interpretable data. We present a categorization of everyday behaviors that can be used to code responses from experience sampling and diary studies conducted with different groups of participants—from adolescents to elderly people. This categorization was developed and validated on a set of 19,840 responses to an open-ended question about participants' recent activity, provided by 667 persons ranging in age from 12 to 66. As a result of the multistage work, we present a categorization system which forms a hierarchy from three broad categories to 97 narrow ones through middle levels of five, 23, and 63 categories of behaviors. The possible usage of the developed categorization is discussed.

Keywords

experience sampling, diary study, behavior, activities, categorization of behaviors



This is an open access article distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), CC BY 4.0, which permits unrestricted use, distribution, and reproduction, provided the original work is properly cited.

Experience sampling method (ESM), also known as ecological momentary assessment (EMA), is commonly used for studying experience and behavior as they occur. When studying behavior, a coding system needs to be used in order to transform participants' varied responses about their current behavior (in open-ended questions) into interpretable data. Currently, different categorization schemes have been used for this purpose, but none of them are without limitations. They are typically customized to a specific group of people (e.g., adolescents or adults with children), they are rather broad, and they focus on activity only, ignoring situational cues (e.g., location and social partners). Therefore, our objective was to prepare a categorization of behaviors that can be applied in various experience sampling and diary studies, conducted on different age groups. As a result of our work, we present a hierarchical categorization system which can be customized to the researcher's needs: narrow or broader categories can be used depending on the purpose.

Experience Sampling Method

Experience sampling method (ESM) is a self-description based measurement procedure that has three main qualities: the assessment of experience or behavior in natural settings, in real time, and on repeated time occasions (Conner, Feldman Barrett, Tugade, & Tennen, 2007). It enables us to follow fluctuations in experiences and to study the links between the external context (e.g., location) and the content of the mind (e.g., feelings; Hektner, Schmidt, & Csikszentmihalyi, 2007). In an experience sampling study, a participant is asked to answer open- and closed-ended questions multiple times per day for multiple days (typically for a week or two weeks). Reports can be made in response to a random signaling device (signal-contingent sampling), at predetermined times during the day (interval-contingent sampling), or following a particular event (event-contingent sampling; Hektner et al., 2007).

ESM has unique qualities that make it, according to Furr (2009), one of the best methods for studying human behavior. It combines the ecological validity of assessment in natural settings with the nonintrusive nature of the diary method (Hektner et al., 2007). As the experience is reported in real time, the risk of recall bias is reduced. By virtue of its intensive repeated measures design, ESM can be used to study between- and within-person variability (Conner et al., 2007). In the era of smartphone apps' popularity, an experience sampling study can be conducted on participants' own mobile devices, which reduces the cost of its implementation and also the burden for participants.

ESM can be employed to address questions regarding relations between behavior and psychological variables. For example, Larson, Richards, Sims, and Dworkin (2001) used ESM to analyze the time budgets of young urban African Americans. Fleeson (2007) investigated whether situations are associated with the manifestation of the Big Five traits in everyday behavior, and Sun, Harris, and Vazire (2019) addressed relations between well-being and the quantity and quality of social interactions. ESM can also

be applied to study intraindividual change and processes—it allows researchers to track fluctuations in experience. It is a useful procedure to answer questions about how and why individuals change over time. Researchers have used ESM for studying intraindividual dynamics of various variables, including fatigue and mood (Hegarty, Treharne, Stebbings, & Conner, 2016), happiness (Mueller et al., 2019), loneliness (van Roekel et al., 2014), impulsivity (Sperry, Lynam, Walsh, Horton, & Kwapil, 2016), motivational conflict, well-being, self-control, and mindfulness (Grund, Grunschel, Bruhn, & Fries, 2015). ESM has gained popularity in different fields, such as education research (Zirkel, Garcia, & Murphy, 2015) or clinical research (Trull & Ebner-Priemer, 2009).

Using ESM for Studying Behavior

One of the most popular questions asked in experience sampling studies is about current or very recent behavior (activity). A typical question would be: “What is the main thing you are doing?” or “What were you doing before receiving the signal?” (Hektner et al., 2007). This question can be either closed- or open-ended. If it is closed-ended, participants are presented with a list of activities and they are instructed to choose the one that best describes their behavior. The advantage of this response format is there is no need for coding, whereas the limitations include: having a small number of behavioral categories (in order to make this task easy for participants), which may result in participants having difficulties identifying a well-suited category, and not being able to detect careless responding. The open-ended format overcomes these limitations. When a participant is asked to describe his or her activity, the measurement is more accurate. Moreover, a researcher can identify careless responses and remove them from the dataset. However, open-ended responses require being coded before they can be included in statistical analysis.

Researchers use different categorization systems for this purpose. Some of them are customized to the needs of the study. For instance, when a study aims to measure adolescents’ engagement in after-school program activities, the categorization is likely to only measure free time activities and to distinguish between structured and unstructured activity (e.g., Bohnert, Richards, Kohl, & Randall, 2009; Shernoff & Vandell, 2007). Interest in physical activity will likely use a categorization that differentiates sedentary behaviors from active ones (e.g., Snippe et al., 2016).

There are also more comprehensive categorizations used in the literature. They do not stress any specific aspect of behavior and can be used for coding any possible response to questions about activities provided at any time. Some examples are presented in Table 1. Each categorization is dedicated to one age group: adolescents or adults. Narrow categories are organized into broader ones.

Table 1*The Comparison of Categorizations of Behaviors Reported in Experience Sampling Studies*

Level of comparison	Csikszentmihalyi & Larson (1984)	Larson et al. (2001)	Robinson & Godbey (1999)	Kubey & Csikszentmihalyi (1990)
Dedicated group	Adolescents	Adolescents	Adults	Adults
Broad categories	Productive (P)	Productive (P)	Non-free time (N)	Work (W)
	Leisure (L)	Leisure (L)	Free time (F)	Home (H)
	Maintenance (M)	Maintenance (M)		Public or others' homes (P)
Narrow categories	classwork (P); studying (P); job (P); sports and games (L); watching TV (L); listening to music (L); art and hobbies (L); reading (non-school; L); thinking (L); other leisure (L); eating (M); personal care (M); transportation (M); chores, errands and other (M); rest and napping (M)	classwork (P); homework (P); extracurricular activities (P); working for pay (P); religious activities (P); TV viewing (L); music listening (L); creative activities (L); talking (L); playing (L); playing games (L); playing sports (L); public leisure (L); idling (L); eating (M); transportation (M); resting (M); chores and errands (M); personal maintenance (M)	paid work (N), household work (N), child care (N), obtaining goods (N), personal needs (N), education (N), organizational (F), entertainment/social (F), recreation (F), communications (F), physical affection (F; added by Graham, 2008)	working (W), doing other things at work (W); watching TV (H); cooking (H); cleaning (H); eating (H); snacking, drinking, smoking (H); reading (H); talking (H); grooming (H); hobbies, repairing, sewing, gardening (H); other chores (H); idling, resting (H); other, miscellaneous (H); leisure and other activities (P); shopping (P); transportation (P)

Why Is Another Categorization Needed?

The categorizations mentioned above are some examples of strategies that can be used in a large number of studies because of their comprehensive character. Nevertheless, they all have some limitations. First of all, each categorization is dedicated to a predefined age group—for instance, children, adolescents, or young adults with children. What is more, they differ in their priority areas—some of them distinguish between more categories of leisure behaviors, whereas others distinguish between more categories of household activities. This can be considered an advantage, as researchers can select a categorization that suits their study's needs best, but this also makes them less universally applicable.

Another feature shared by popular categorization methods is that they focus on activity only and not on situational cues, such as location and social partners. This is a limitation, because information about social partners especially, but also about location, may be crucial for interpreting a person's behavior. Let us take watching a film with a romantic partner as an example: For some respondents, the main activity will be watching a film but for others it can be spending time with a partner. We assume that the fact that a respondent includes information on his or her social partner in the description of activity indicates that this situational cue is important and, therefore, should be considered while coding responses. Experience sampling forms often contain

separate questions about social partners and locations, therefore the responses to all of these questions can be combined to describe the reported behavior more accurately. However, these two approaches do not lead to the same result. Reporting that another person was present while a participant was engaged in an activity is not the same as mentioning that person's presence in the description of activity. When a behavior is coded as watching a film and a participant reports that their partner was present, at least three different scenarios are possible: (a) the partner is engaging in an activity other than watching the film, (b) the participant and their partner are watching the film together and the participant is primarily concentrated on the film, or (c) the participant and their partner are watching the film together and the participant is primarily concentrating on spending time with their partner. Hence, we assume that a participant would not mention their partner's presence in the description of behavior if this fact were not important to them at that moment. We believe that taking this information into account while coding open-ended responses leads to more accurate representations of activities. After all, trying to represent reported activities as accurately as possible is one of the main preconditions of the success of research conducted using ESM. As we explained above, ESM can be used to study behavior in natural settings and in real time, a feature which makes this method quite exceptional. Behavior itself is, however, reported by participants and not observed by researchers, thus an accurate categorization of the behavior descriptions is crucial.

Current Study

Our aim was to prepare a categorization of momentary behaviors reported in ESM that can be considered more universal than other existing categorizations. We wanted to make it suitable for studies on different age groups and with different objectives.

In order to achieve our goal, we analyzed a pool of responses to a question about a participant's recent behavior from experience sampling studies conducted on different groups of participants—from teenagers to older adults. We aimed to develop a number of categories large enough to differentiate them well, but small enough to be simple to use. We grouped them into broader categories. In the next step, we developed a larger set of narrow categories which can be seen as a lower level of the hierarchy of categories. We had no preliminary assumptions or anticipations about the number of categories or about the structure of categorization.

We describe the process of developing a categorization based on reducing a pool of responses to a much smaller number of categories (Steps 1–4). Then we used additional datasets to develop a more detailed set of categories and to validate the categorization system (Steps 5 and 6). All studies described in this paper were carried out in accordance with the recommendations of the Commission of Ethics and Bioethics at Cardinal Stefan Wyszyński University in Warsaw with written informed consent from all subjects in accordance with the Declaration of Helsinki. Parents of participants under the age of 16

provided written informed consent as well. The studies were carried out as part of the project that received ethical approval from the Commission of Ethics and Bioethics at Cardinal Stefan Wyszyński University in Warsaw (registration number: KEiB – 20/2015; date of the decision: October 1, 2015).

Method

Procedure

Our aim was to find a (semi)-universal categorization of behaviors by reducing a pool of responses to a question about recent behavior asked in experience sampling studies. In order to achieve this goal and overcome the limitations of other categorizations used in the literature, we applied an extensive and complex research plan consisting of six steps performed in six studies on 19,840 behavioral descriptions in total. The research and analysis plan contained a preliminary analysis, a development of the first version of the categorization, a modification, and a final validation of the proposed categorization. Table 2 presents the overview and sequence of our work.

Table 2

The Steps of the Categorization Process

Step	Goal	Description/Action
1.	<i>Preliminary semantic analysis</i>	<ul style="list-style-type: none"> • Removing answers identical or almost identical (differing only in tense or gender) to any other responses • Result: reduction of Pool 1 of 5,435 responses to 1,355
2.	<i>Further reduction of the pool of responses</i>	<ul style="list-style-type: none"> • Analyzing the meaning of the 1,355 responses and narrowing of the list by putting together similar responses (e.g., “Playing with my child,” “Playing with my daughter,” “Playing with my son”) • Result: 730 non-duplicative behaviors
3.	<i>First version of the categorization</i>	<ul style="list-style-type: none"> • Random division of the whole list of 730 behaviors into two parts • Working independently in two separate teams to create two propositions of the categorization • Comparison of the results and discussion on producing one version of a categorization • Result: a list of 23 categories and a proposition of a three-level structure of the categorization
4.	<i>Verification and modification</i>	<ul style="list-style-type: none"> • Verification of the proposed categorization on a different pool of behaviors (performed individually by five researchers) • Working on the new data set (Pool 2 of 336 non-duplicative behaviors) • Discussion of the results • Modification of the categorization by removing the confusing categories and introducing some small changes • Result: the final categorization including 23 categories (presented in the Results section)

Step	Goal	Description/Action
5.	<i>Final verification</i>	<ul style="list-style-type: none"> Collecting the third pool of responses in order to verify the modified version of the categorization (Pool 3 of 196 non-duplicative behaviors) Coding these responses by five judges separately and discussion of the results Using 13,873 responses from Study 6 to assess reliability and validity of the 23 categories. Two new judges participated in the assessment of reliability
6.	<i>Development of a set of narrow categories (lower level of the hierarchy of categories)</i>	<ul style="list-style-type: none"> Selection of 9,592 behaviors described by participants as autonomous Exclusion of 97 unreliable responses Selection of 3,914 behavioral acts most strongly related to value states Reduction of the 3,914 responses to a pool of 646 by following actions described in Steps 1 and 2 Grouping 646 descriptions of behaviors by two judges with the aim to reduce this pool to about 100 categories Result: 98 narrow categories of behaviors

Participants and Material

We used four pools of responses to an open-ended question about participants' recent behavior (see Table 3) drawn from six studies that are described below conducted on Polish samples. The total number of responses included in analyses was 19,840 (5,435 responses in the first pool; 336 responses in the second pool; 196 responses in the third pool; and 13,873 in the fourth pool) provided by 667 participants.

Table 3

The Four Pools of Descriptions of Behaviors Used in the Analyses

Pool	Responses (<i>n</i>)	Study from which responses were taken	Step of our work in which the pool was used ^a
1	5,435	Study 1, Study 2, Study 3	Step 1, Step 2, Step 3
2	336	Study 4	Step 4
3	196	Study 5	Step 5
4	13,873	Study 6	Step 5, Step 6

^aSteps are described in Table 2.

Study 1

The first study was conducted on first year middle school pupils (12–13 years old) and on first year high school pupils (15–16 years old). The number of responses was 754, provided by 70 participants ($M_{\text{age}} = 14.62$, $SD_{\text{age}} = 1.51$, 59% females, no data available for two participants). They were requested to download and run a mobile app to fill out an experience sampling form five times a day while doing something important to them. The first question they were asked was about the chosen behavior ("What were you doing before using the app?"). The study was conducted on participants' own mobile devices. These 754 responses by 70 participants were included in the first pool of behaviors.

Study 2

The second study consisted of a sample of 151 adults (no data available for 41 participants¹, age range 20–60, $M_{\text{age}} = 28.90$, $SD_{\text{age}} = 8.59$, 69% female) who participated in an experience sampling study conducted on their own mobile devices. They were prompted at random occasions within a given time frame to answer the following open-ended question: “What have you been doing for the past 15 minutes?” There were 3,774 responses to this question provided by 151 participants included in the first pool of behaviors.

Study 3

The procedure of the third study was the same as in Study 2: Participants were prompted to answer an open-ended question about their activity during the past 15 minutes five times a day. They responded using their own mobile devices. The sample consisted of 907 responses provided by 31 adults aged from 19 to 41, $M_{\text{age}} = 28.88$, $SD_{\text{age}} = 5.29$, 88% females (no data available for seven participants). The 907 responses were included in the first pool of behaviors.

Study 4

The procedure was the same as in Studies 2 and 3. The sample consisted of 2,256 responses provided by 42 participants aged from 12 to 66, $M_{\text{age}} = 29.37$, $SD_{\text{age}} = 10.82$, 79.4% females (no data available for 11 participants). From this pool of responses, we selected 336 responses describing different behaviors to constitute the second pool.

Study 5

In Study 5, 138 persons who had also participated in Study 2 ($M_{\text{age}} = 29.86$, $SD_{\text{age}} = 9.29$, 70% females, no data available for 46 participants) filled out a short diary at the end of each day of participation in the study. They chose three activities that were most important to them during a day and wrote them down. This resulted in 1,499 responses; from which we selected 196 descriptions of different behaviors for the third pool.

Study 6

In Study 6, we collected 13,873 responses to the ESM form, provided by 374 participants aged 17 to 53 ($M = 23.72$, $SD = 4.67$), all Caucasian, 79% female. Each participant responded to 19–49 ESM forms with a mean of 37.1 forms per person (a 76% response rate)².

1) The questions about sex and age were not included in the ESM form, but asked in a questionnaire. The large number of missing data on sex and age of respondents is a result of missing questionnaires or errors in the identification codes assigned to participants and can be deemed as random.

2) The same dataset was used in other statistical analyses, presented in three papers: (1) Skimina et al. (2018), (2) Skimina et al. (2019), (3) Skimina and Ciecuch (2020).

The data collected in this study were used to validate the categorization system (Step 5) and to develop a set of narrow categories (Step 6). For validation, we used not only the descriptions of behaviors, but also the participants' responses to other questions included in the ESM form.

The ESM form started with an open-ended question about participants' recent behavior: "What have you been doing for the past 15 minutes? (Please describe only one, the most important activity)". It was followed by a question about perceived autonomy in that behavior. Participants had to choose between two options: "This activity was imposed by another person or by the circumstances" or "This activity was my choice—I could either do or not do it." Participants then indicated who they interacted with by answering the following question: "Who participated with you in this activity?" The possible answers were: (a) nobody, (b) partner/spouse, (c) other family member, (d) friend, (e) colleague, (f) stranger. In four subsequent questions, participants assessed their emotional states during the past 15 minutes using the following dimensions on a 7-pt scale: (a) apathy versus enthusiasm, (b) worry versus cheerfulness, (c) anger versus calmness, (d) arousal versus restraint. After that, they responded to nine questions measuring value-states (momentary importance of values) selected from Schwartz et al.'s theory (2012; for applying ESM to measuring values see Skimina et al., 2018). Each question started with "When you were engaging in this activity, how important was it to you to...?" The endings corresponding to particular values are presented in Table 4.

Table 4

Items Measuring Momentary Importance of Values (Value-States)

When you were engaging in this activity, how important was it to you to ...	Value measured	Value definition in terms of motivational goal (Schwartz et al., 2012)
be better at something than others are	Achievement	Success according to social standards
get some advantage for yourself	Power-Resources	Power through control of material and social resources
avoid danger	Security-Personal	Safety in one's immediate environment
do what someone else expected	Conformity-Rules/ Conformity-Interpersonal	Compliance with rules, laws, and formal obligations/Avoidance of upsetting or harming other people
help people you care about	Benevolence-Caring	Devotion to the welfare of ingroup members
help someone you did not know	Universalism-Concern	Commitment to equality, justice, and protection for all people
understand something or form an opinion on your own	Self-Direction-Thought	Freedom to cultivate one's own ideas and abilities
experience something new or exciting	Stimulation	Excitement, novelty, and change
enjoy yourself	Hedonism	Pleasure and sensuous gratification

Note. The table was reprinted from Skimina et al. (2018).

Participants responded on a scale from 1 (*not important at all*) to 4 (*very important*).

The ESM study was conducted on participants' own mobile devices. It lasted for seven days. Seven prompts were scheduled to show up randomly each day between 9.30 a.m. and 9.30 p.m. with a minimum of 60 minutes between two prompts. After each prompt, the ESM form was available for 45 minutes.

Results

Development of the Main Categorization

The process of categorization development included six steps. Steps 1–4 refer to the development of the main categorization. Step 5 refers to validation of this categorization. Step 6 refers to the development of a set of narrow categories of behaviors that can be perceived as a lower level of the categorization hierarchy.

To develop the main categorization, we worked on a pool of 5,435 descriptions of behavioral acts. The first and the second step of our work (described in Table 4) resulted in reducing the response pool from 5,435 to 730 non-duplicative behaviors (as *duplicative* we mean identical or differing only in tense or gender). In the third step, we obtained 23 initial categories: Interactions with strangers, Interactions with friends, Time spent with animals, Time spent with family, Other interactions (with undefined people), Health and beauty, Indoor entertainment, Outdoor entertainment, Traditions and customs (e.g., preparing for Christmas Eve), Religion, School/Work, DIY/Manual activities, Errands, Transportation, Housework, Physical activity, Physiology (e.g., eating, sleeping, going to the bathroom), Hobbies, Shopping, Waiting, Unusual (activities done rarely, e.g., moving into a new apartment), Others, Unclassifiable.

The fourth step was to verify the initial 23 categorizations on a different pool of behaviors, consisting of 336 non-duplicative behaviors (Pool 2). Five persons who participated in the previous work coded each of 336 responses using the developed categorization (we worked separately). We calculated the compatibility rate for the assessment of each behavior. The judges' rating agreement was 100%—which means assignment to the same category—in the case of 181 behaviors (54% of responses). Five or four judges assigned 256 responses (76% of total) to the same category (agreement rate of at least 80%). The mean agreement rate for all 336 behaviors was 84.2%. Disagreement seems to focus on certain categories, which turned out to be ambiguous. Thirteen of the 23 categories had high compatibility: Waiting, Housework, Physiology, Indoor entertainment, Outdoor entertainment, School/Work, Religion, Time spent with family, Transportation, Health and beauty, Interactions with friends, Time spent with animals, and Shopping. The ambiguous categories that turned out to be troublesome for judges were: Hobby, Other interactions, DIY/Manual activities, Interactions with strangers, Unusual, Traditions and customs, and Errands.

After analyzing the results of this verification, we decided to modify the categorization. From the first 23 categories (22 categories and a group of unclassifiable answers) we removed the confusing categories: Traditions and customs, Unusual, and Interactions with strangers. The category School/Work was divided into three categories: Work, School/University classes, and Studying and extra classes. We also divided the category Interactions with family into two categories: Family time and Time with a partner. Three other categories were redefined: the categories Hobby and DIY/Manual activities were changed to Creative hobbies and Repairing. The category Physiology was extended by adding the usage of stimulants.

The final categorization consists of 22 groups of behaviors. Among them, 21 represent behaviors that are easy to categorize, and the last category merges other, undefined behaviors. Answers that cannot be included in any of the categories were labeled as *Unclassifiable* (23rd category). All categories, their descriptions, and examples are presented in Table 5.

Table 5

The Final Categorization of Behaviors, Including 23 Categories, With Descriptions and Examples

No	Name	Description	Examples
1	Work	Working and activity during working hours, i.e., contact with clients, superiors, or colleagues; socializing after work is not included	I was having a work meeting. I was working as a coach.
2	School/University classes	Participation in school and university activities, taking exams	I've had an exam. I've been participating in a lecture.
3	Studying and extra classes	Extra classes, extra language lessons, attending conferences, learning at home, doing homework, reading books related to school or work	I have been studying for school. I was participating in a training course.
4	Housework	Cooking, cleaning, washing, gardening	I was cooking/baking. I was cleaning the kitchen. I am washing my car.
5	Shopping	All kinds of shopping, ordering and picking up purchases	I was buying a TV set. I bought a book. I was buying tickets for concert online.
6	Errands	Dealing with issues apart from shopping, photocopying and printing outside work, visits to the office, the dressmaker, the pharmacy; appointments are not included	I was at the post office. I was organizing a medical appointment. I was picking my car up from the repair shop.

No	Name	Description	Examples
7	Family time	Various forms of spending time with family including childcare	I was changing a diaper. I've been eating a dinner with my family.
8	Time with a partner	Spending time with a partner (husband/wife, fiancé/fiancée, boyfriend/girlfriend), including sexual behaviors	I was watching a movie with my wife. I was cuddling with my fiancée.
9	Time with friends	Spending time with friends, going out with them, partying, hosting friends at home	I talked to my friend. I was chatting with my friend online. I was talking about my studies with my colleagues.
10	Time with pets	Spending time with pets (also someone else's); taking care of animals	I was walking my dog. I stroked a cat. I was walking with the dogs in the shelter.
11	Other socializing	Socializing with undefined or unknown person, besides appointments and official contacts	I helped a lady on the bus. I talked to a young man. I talked to the nurse.
12	Indoor leisure and resting	Reading (unrelated to work or school), watching TV, playing computer and board games, doing crossword puzzles, using the Internet (except calls over the Internet), resting	I am watching a documentary film. I am listening to the radio. I was reading a book.
13	Outdoor leisure	Mass and cultural events, cinema, fishing, scout rally, eating out, sightseeing, travelling, walking	I was at the cinema/theatre. I was at a concert. I was eating out.
14	Physical activity	Sport workout, fitness classes, dancing, jogging, intense walking	I was at yoga classes. I was cycling. Swimming-pool.
15	Repairing	Minor repairs of, i.e., electronic equipment, assembling of furniture, house minor repairs, painting the walls	I was repairing the car. I was framing a book.
16	Creative hobbies	Playing instruments, singing, drawing, sewing, making jewelry	I was singing. I am at the theatrical workshop. I am making beer. I am knitting.
17	Religious practice	Religious practices, also those connected with holidays and festivity	I was at holy mass. I was praying. I was worshipping God.

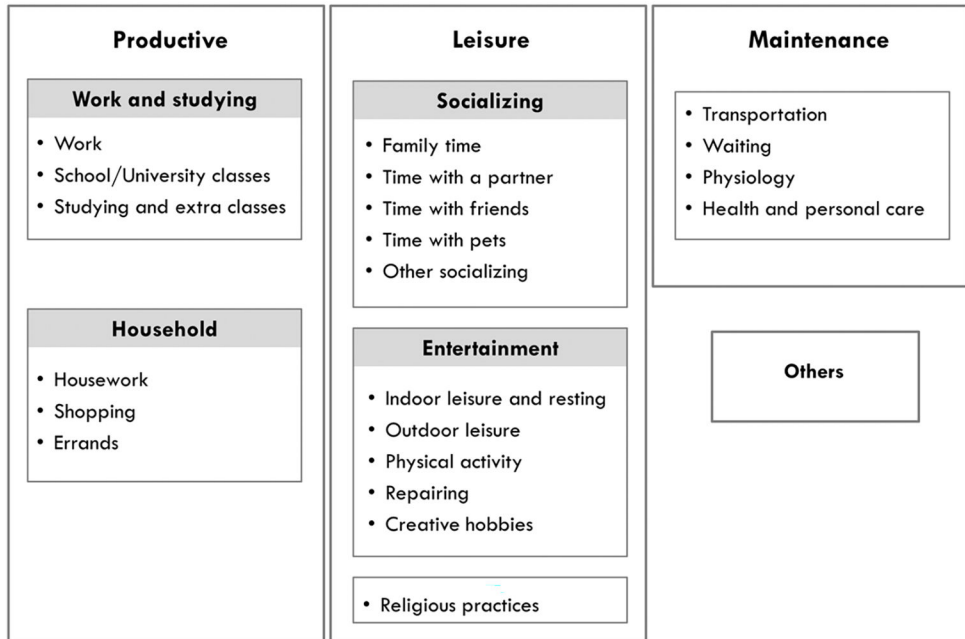
No	Name	Description	Examples
18	Transportation	Being on the way somewhere (means of transport or walking); walking around and biking are not included	I was walking. I am using public transport. I was running to the bus.
19	Waiting	Waiting for something or somebody, no matter where	I am waiting for a meal. I am sitting in the waiting room. I am waiting in a line.
20	Physiology	Eating or drinking, sleeping, using a toilet, using stimulants	I was sleeping. I am drinking coffee. I was in the toilet.
21	Health and personal care	Personal hygiene, dressing, getting ready to go out, make-up, visiting doctors	I was shaving. I was taking a bath. I am taking pills.
22	Others	Behaviors mismatched to any other category	
23	Unclassifiable	Answer that cannot be classified, i.e., complex, ambiguous, confusing	

The categorization system and its hierarchical structure is shown in Figure 1. On the top of the hierarchy there are three meta-categories, which we also name higher-order categories: Productive, Leisure, and Maintenance. The first higher-order category—Productive—contains classes of behaviors related to responsibilities, duties, and everyday tasks, but also extra learning. It can be simply divided into two middle-level groups of behavior, which are Work and studying and Household. If needed, these categories can be further divided: Work and studying into Work, School/University classes, and Studying and extra classes; Household into Housework, Shopping, and Errands. We decided to differentiate learning at school or at university during compulsory classes and learning during free time—at home or during additional classes. It allows one to differentiate between obligatory and chosen activities. In line with our assumption, foreign language lessons should also be included in the latter category (if they are not obligatory lessons at school, or work).

The second higher-order category—Leisure—refers to spare time. We distinguish between two main groups of activities typical for spending free time: Socializing and Entertainment. The first group (Socializing) includes categories such as Family time, Time with a partner, Time with friends, Time with pets, and Other socializing, whereas the second group (Entertainment) includes Indoor leisure and resting, Outdoor leisure, Physical activity, Repairing, and Creative hobbies. The Leisure higher-order category also includes Religious practice as activities typically conducted in one's spare time.

Figure 1

A Hierarchical Structure of the Categorization System (23 Categories Assigned to Broader Ones)



The third higher-order category—Maintenance—refers to basic, daily or habitual activities that are common in daily life. This higher-order category includes: Transportation, Waiting, Physiology, and Health and personal care.

Besides the categories mentioned above, we also distinguished between two other groups of answers: Others and Unclassifiable. The first one takes behaviors that do not fit into any category described above into account. The Unclassifiable category includes ambiguous answers and unintelligible statements which cannot be called behaviors (e.g., “game”).

Validation of the Main Categorization

To validate the categorization system developed in Steps 1–4, we used two datasets. First, we verified the categorization in the same way as we verified its previous version. Each of the five judges classified 196 responses selected from data collected in Study 5 to one category. The judges’ rating agreement was 100% – which means assignment to the same category by all—in the case of 109 behaviors (55.6% of responses). Five or four judges assigned 158 behaviors (80.6% of responses) to the same category (agreement rate at least

80%). The mean agreement rate for all 296 behaviors was 86.6%. These results are slightly better than in the first version of the categorization.

Then we used the data collected in Study 6 to assess the reliability and validity of the final version of the main categorization. We assessed the reliability by calculating inter-rater agreement among three independent judges (two of them were new and did not participate in the categorization development in previous steps). To assess the validity of the categorization, we compared different activity categories in terms of emotional states and value-states (importance ascribed to personal values; see Skimina, Ciecuch, Schwartz, Davidov, & Algesheimer, 2018) during activities assigned to these categories.

Reliability

In order to assess the reliability of the developed categorization system, we calculated inter-rater agreement indices among three raters: an expert (a person who participated in the development of the categorization) and two independent raters who did not participate in the development of the categorization. They received a Polish version of Table 5, containing descriptions of the categories and example behaviors. They were instructed to assign each behavior from a list to one category, if possible. If, in their opinion, a response could be assigned to more than one category, they were asked to list all those categories.

The list of behaviors used in this task was selected from the pool of all 13,873 responses provided by participants of Study 6. We first randomly selected 10% of the whole pool of 13,873 responses and then reduced the pool to 520 responses by removing duplicates (identical responses or responses differing only in tense or gender).

In the first step, we examined how many responses could be assigned to more than one category. We found that 79.6% of 520 responses were assigned to only one category by all three raters, 12.7% were assigned to two categories by one rater and to only one category by the rest of the raters, 7.7% were assigned either to three or more categories by one rater or to more than one category by two or three raters. A response was assigned to more than one category in one of the following situations: (a) it described more than one activity, (b) it could not be understood properly without the context, (c) it described being on the way to somewhere, or (d) it described getting ready to go somewhere. Thus, we concluded that assigning each activity to only one category was justified.

In the second step, we calculated inter-rater agreement indices for the whole pool of 520 responses. All three raters assigned a response to the same category in 82.3% of cases (either as the only one or as one of the possible categories to describe the activity). In 16.2% of cases, two of three raters chose the same category and 1.5% of responses were assigned to a different category by each rater. The responses with no inter-rater agreement were: "Preparations for a trip," "I was reading survey instructions," "I'm writing a note," "I was unpacking a package," "I was decorating the Christmas tree,"

“I’m preparing for a business trip,” “I was preparing for work,” “I’m going to pick up a package.”

Then we calculated Cohen’s kappa (McHugh, 2012) for each pair of the three raters. The results were the following: .76, .83, and .84 ($M = .81$). They can be interpreted as satisfactory. The level of agreement was higher when one of the raters in a pair was an expert, which may suggest that the accuracy of coding behaviors using this categorization improves with practice.

In the last step, we calculated reliability indices for each category separately. In order to do this, we used the expert’s codes as benchmarks—we checked how many responses assigned to each category by the expert were assigned to the same category by the other raters. The results are presented in Table 6.

Table 6

Inter-Rater Agreement for Particular Categories

Category (N of responses)	Perfect inter-rater agreement	Half inter-rater agreement	No inter-rater agreement
1. Work (93)	73.1%	23.7%	3.2%
2. School/University classes (28)	71.4%	28.6%	0.0%
3. Studying and extra classes (40)	90.0%	10.0%	0.0%
4. Housework (39)	84.6%	5.1%	10.3%
5. Shopping (23)	91.3%	8.7%	0.0%
6. Errands (6)	16.7%	66.7%	16.7%
7. Family time (26)	96.2%	3.8%	0.0%
8. Time with a partner (5)	80.0%	20.0%	0.0%
9. Time with friends (25)	88.0%	12.0%	0.0%
10. Time with pets (4)	100.0%	0.0%	0.0%
11. Other socializing (4)	50.0%	50.0%	0.0%
12. Indoor leisure and resting (46)	87.0%	13.0%	0.0%
13. Outdoor leisure (19)	63.2%	36.8%	0.0%
14. Physical activity (10)	90.0%	10.0%	0.0%
15. Repairing (2)	100.0%	0.0%	0.0%
16. Creative hobbies (4)	100.0%	0.0%	0.0%
17. Religious practice (2)	100.0%	0.0%	0.0%
18. Transportation (50)	94.0%	6.0%	0.0%
19. Waiting (10)	80.0%	20.0%	0.0%
20. Physiology (29)	96.6%	3.4%	0.0%
21. Health and personal care (39)	87.2%	12.8%	0.0%
22. Others (10)	30.0%	70.0%	0.0%
23. Unclassifiable (6)	50.0%	50.0%	0.0%

Note. Perfect inter-rater agreement = both other raters assigned the response to the same category as the expert; half inter-rater agreement = one of the other raters assigned the response to the same category as the expert; no inter-rater agreement = none of the other raters assigned the response to the same category as the expert.

As can be seen, some categories were poorly represented in the sample of 520 responses, for instance Repairing or Religious practice (only two responses for each), whereas others were over-represented, for instance Work (93 responses) and Transportation (50 responses), which reflects the distribution of different categories in the general population of responses (cf. Table 7). For the majority of categories, the inter-rater indices of agreement were high (> 80% agreement among all raters). Lower indices were found for Work and School/University classes (> 70% agreement among all raters). In the case of Work, some of the disagreement seems to be a consequence of the lack of context. When it comes to School/University classes, one of the independent raters assigned some of the responses to the category Studying and extra classes, which are similar in content. There was also lower inter-rater agreement for Outdoor leisure (63.2% agreement among all raters). Some responses assigned to this category could also be assigned to Time with friends (e.g., participating in a party). The reason for disagreement in Other socializing was assigning two responses as unclassifiable by one of the independent raters (the responses were: “I was at the meeting” and “Meeting”). The Errands category appeared to be the most troublesome—four of six responses assigned to this category by the expert were assigned to the same category by one of the other raters; one was assigned to the same category by both other raters, and one was assigned to different categories by all raters. The likely reason for this is that in the sample of responses there was only one response typical for the Errands category—it was “I was at the post office” and it was identically assigned by all raters. The other responses were difficult to interpret precisely without context (e.g., “Printing”) or described being on the way to somewhere (e.g., “I’m going to pick up a package”). We attribute this poor index of inter-rater agreement to an unfortunate coincidence rather than to the ambiguity of this category. The level of agreement in the cases of Others and Unclassifiable categories appeared to be quite large. Some responses coded as Others by the expert were coded as Unclassifiable by one of the other raters.

We conclude that the categorization system is a reliable tool. The ambiguity regarding some categories could be reduced while coding responses in the dataset, when contextual information is provided.

Validity

To validate the categorization, the whole pool of 13,873 responses was coded by the expert and each response was assigned to only one category. While coding, the expert was provided with some information about participants and particular prompts that were available in the dataset. For instance, time of measurement and previous responses provided by the same person served as a context for the activity which was helpful in assigning responses to particular categories.

In Table 7, we present the frequencies of all categories in the pool of 13,873 responses. As can be seen, participants spent most of their time resting, working, or fulfilling their

physiological needs. They rarely reported spending time with a partner, spending time with pets, waiting, doing errands, repairing, or doing something creative.

Table 7

Frequencies of Categories in the Pool of 13,873 Responses to Open-Ended Question About Recent Activity

Category	Frequency	%
Work	2306	16.6
School/University classes	669	4.8
Studying and extra classes	1103	8.0
<i>Work and studying (total)</i>	<i>4078</i>	<i>29.4</i>
Housework	1052	7.6
Shopping	619	4.5
Errands	75	0.5
<i>Household (total)</i>	<i>1746</i>	<i>12.6</i>
Family time	341	2.5
Time with a partner	102	0.7
Time with friends	275	2.0
Time with pets	63	0.5
Other socializing	291	2.1
<i>Socializing (total)</i>	<i>1072</i>	<i>7.7</i>
Indoor leisure and resting	2375	17.1
Outdoor leisure	330	2.4
Physical activity	202	1.5
Repairing	20	0.1
Creative hobbies	54	0.4
<i>Entertainment (total)</i>	<i>2981</i>	<i>21.5</i>
Transportation	1267	9.1
Waiting	96	0.7
Physiology	1904	13.7
Health and personal care	523	3.8
<i>Maintenance (total)</i>	<i>3790</i>	<i>27.3</i>
Religious practice	49	0.4
Others	73	0.5
Unclassifiable	84	0.6
Total	13873	100.0

The categorization validity was assessed in two steps. First, by comparing higher-order categories (Work and studying, Household, Socializing, Entertainment, and Maintenance) in terms of the importance of values pursued during these activities. Second, by examining whether there is a reason to distinguish categories based not only on activity, but also on the interaction partner mentioned in the response. We examined this by comparing

emotional states and value importance reported when interacting with a partner or a family member based on participants' response to the question: "Who participated with you in this activity?" Participants could either mention an interaction partner in the open-ended question about the activity, or not.

In the first step of validation, we compared the hierarchies of values, from Schwartz et al.'s refined theory (2012), pursued during activities from different higher-order behavioral classes. Schwartz (1992) defines values as trans-situational goals, varying in importance, that guide perception and behavior. This definition refers to values as dispositions (value-traits). In the current study, values were treated as states, which means we examined the situational importance of goals in single behavioral acts. The differentiation between value-traits and value-states was proposed by Skimina et al. (2018).

The hierarchies are based on the mean scores of ESM items measuring the momentary importance of particular values (see: Table 3). The comparison is presented in Table 8.

Table 8

Hierarchies of Values Important in Activities Assigned to Different Higher-Order Categories

Work and studying		Household		Socializing		Entertainment		Maintenance	
Value	M (SD)	Value	M (SD)	Value	M (SD)	Value	M (SD)	Value	M (SD)
POR	2.87 (1.08)	POR	2.81 (1.04)	HE	2.82 (1.14)	HE	2.85 (1.08)	POR	2.75 (1.08)
COI	2.74 (1.06)	BEC	2.35 (1.15)	BEC	2.79 (1.13)	POR	2.74 (1.09)	SEP	2.28 (1.17)
SDT	2.57 (1.10)	HE	2.18 (1.09)	POR	2.42 (1.08)	ST	2.29 (1.11)	HE	2.07 (1.09)
AC	2.37 (1.10)	SEP	2.17 (1.10)	COI	2.38 (1.09)	SDT	2.17 (1.09)	ST	1.80 (1.01)
ST	2.11 (1.04)	COI	2.11 (1.04)	SDT	2.38 (1.10)	SEP	1.83 (1.06)	BEC	1.79 (1.04)
SEP	2.10 (1.13)	ST	1.88 (1.02)	ST	2.25 (1.10)	BEC	1.64 (0.96)	COI	1.77 (0.98)
UNC	2.00 (1.10)	SDT	1.79 (0.99)	SEP	2.17 (1.19)	AC	1.59 (0.95)	SDT	1.71 (0.96)
HE	1.79 (0.96)	AC	1.67 (0.94)	AC	1.59 (0.93)	COI	1.56 (0.90)	AC	1.54 (0.89)
BEC	1.74 (1.00)	UNC	1.41 (0.81)	UNC	1.56 (0.94)	UNC	1.36 (0.75)	UNC	1.44 (0.81)

Note. POR = Power-Resources; COI = Conformity-Interpersonal; SDT = Self-Direction-Thought; AC = Achievement; ST = Stimulation; SEP = Security-Personal; UNC = Universalism-Concern; HE = Hedonism; BEC = Benevolence-Caring.

There are some similarities among all hierarchies in the different higher-order categories: Power-Resources is situated on the top and Universalism-Concern on the bottom. It is important to note that the item that was supposed to measure Power-Resources was "to get some advantage for yourself." This was instead a measure of Self-Enhancement (as a broader higher-order value) rather than Power-Resources (as a more narrowly defined value that is a part of Self-Enhancement) in particular, according to Schwartz et al.'s theory (2012; see Skimina, Ciecuch, Schwartz, Davidov, & Algesheimer, 2019). For this reason, it is not surprising that participants wanted to get some advantage for themselves in a variety of situations. On the contrary, the item meant to measure Universalism-Con-

cern was formulated in a way that made it difficult to be found as important in everyday activities by participants.

Despite similarities, there are many substantial differences among hierarchies. For instance, Achievement is situated at a low hierarchy level for most categories, but on a much higher level for Work and studying behavioral categories. By way of contrast, Hedonism is situated low in Work and studying, but high in the other categories. Benevolence-Caring was most important during activities from Household and Socializing and less important in Work and studying. Conformity-Interpersonal was the most important in Working and studying and the least important in Entertainment. Security-Personal was higher in the hierarchy of Maintenance than in other hierarchies. The possible explanation for this is that one of the frequently reported lower-order categories included in Maintenance is Transportation. All these differences among hierarchies show that people's experiences vary depending on activities they are engaged in. Work and studying, Household, Socializing, Entertainment, and Maintenance are related differently to the momentary importance of personal values.

In the second step of validation, we examined whether there is something exceptional in responses that describe activities including information on a social partner and, therefore, that they should be assigned to a distinct category. For this purpose, we compared the average emotional states and the average momentary importance of values reported during activities assigned to Time with partner category (when spending time with partner was mentioned in the response to the open-ended question about the activity) and during activities assigned to Indoor leisure category when the partner was present (the information about company was provided in another question included in the ESM form) but not mentioned in the response to the open-ended question about the activity. In a similar way, Family time was compared to Indoor leisure, when a family member was present, but not mentioned in the response to the open-ended question. This comparison is presented in [Table 9](#).

As can be seen in [Table 9](#), situations in which partner/family member is present but not mentioned in the description of the most important activity provided by the respondent and situations in which interaction with partner/family member is described as the most important activity are experienced differently. When a participant includes their interaction with a partner in their response to the question about activity, they also report a higher level of enthusiasm, cheerfulness, and arousal, as well as higher importance of Conformity-Interpersonal, Benevolence-Caring, Stimulation, and Hedonism values. Similarly, including interaction with family in the description of the most important activity is associated with a higher level of cheerfulness, higher importance of Security-Personal, Conformity-Interpersonal, Benevolence-Caring, and lower importance of getting some advantage for themselves (Power-Resources).

Table 9

Comparison of Experience During Activities Done in Companionship of Partner/Family Member When Interaction Partner Was Mentioned in the Description of the Most Important Behavior and When He or She Was Not Mentioned in Spite of His or Her Presence Reported in Other Question

Variable	Interaction with partner				Interaction with family			
	Partner mentioned <i>M</i> (<i>SD</i>) <i>n</i> = 90	Partner not mentioned <i>M</i> (<i>SD</i>) <i>n</i> = 516	<i>Z</i> (Mann-Whitney's test)	Cohen's <i>d</i>	Family mentioned <i>M</i> (<i>SD</i>) <i>n</i> = 276	Family not mentioned <i>M</i> (<i>SD</i>) <i>n</i> = 290	<i>Z</i> (Mann-Whitney's test)	Cohen's <i>d</i>
Emotional states								
Enthusiasm vs. apathy	5.93 (1.49)	5.35 (1.35)	-4.78***	0.41	5.50 (1.34)	5.25 (1.51)	-1.80	0.18
Cheerfulness vs. worry	5.88 (1.64)	5.52 (1.29)	-3.88***	0.24	5.65 (1.42)	5.41 (1.43)	-2.40*	0.17
Calmness vs. anger	5.74 (1.47)	5.68 (1.29)	-1.03	0.04	5.74 (1.28)	5.67 (1.35)	-0.58	0.05
Restraint vs. arousal	4.10 (2.18)	5.25 (1.61)	-4.56***	-0.60	5.29 (1.62)	5.40 (1.55)	-0.65	-0.07
Value-states								
Achievement	1.64 (1.01)	1.53 (0.93)	-1.13	0.11	1.41 (0.77)	1.38 (0.72)	-0.22	0.04
Power-Resources	2.88 (0.98)	2.63 (1.10)	-1.87	0.24	2.24 (1.05)	2.51 (1.01)	-3.24**	-0.26
Security-Personal	2.13 (1.12)	1.91 (1.09)	-1.93	0.19	2.53 (1.23)	1.83 (1.05)	-6.95***	0.61
Conformity-Interpersonal	2.37 (1.10)	1.72 (0.95)	-5.57***	0.63	2.62 (1.06)	1.63 (0.83)	-10.89***	1.05
Benevolence-Care	2.93 (1.13)	1.92 (1.05)	-7.48***	0.93	3.28 (0.87)	1.84 (1.00)	-14.43***	1.54
Universalism-Concern	1.32 (0.75)	1.36 (0.77)	-0.42	-0.05	1.36 (0.75)	1.31 (0.69)	-0.82	0.69
Self-Direction-Thought	2.27 (1.14)	2.08 (1.04)	-1.37	0.17	2.21 (1.05)	2.14 (1.01)	-0.71	0.07
Stimulation	2.69 (1.07)	2.20 (1.10)	-3.84***	0.45	2.12 (1.03)	2.19 (1.03)	-0.92	-0.07
Hedonism	3.19 (1.05)	2.86 (1.05)	-3.10**	0.31	2.74 (1.15)	2.90 (1.03)	-1.45	-0.15

* $p < .05$. ** $p < .01$. *** $p < .001$.

Development of the Lower Levels of the Categorization System

The data collected in Study 6 were also used to develop a set of narrow categories that can be perceived as a lower-level categorization of the system described above. The narrow categories were developed by Skimina et al. (2019). Their development process is summarized in Table 4, Step 6. Skimina and colleagues were looking for associations between value-states and real-time behaviors. From the set of 13,873 descriptions of behaviors provided by participants, they selected those that were related to the highest rates

of value-states. First, they person-centered all value-state scores and then sampled ESM records in which participants assigned the largest importance to particular value-states. They used a score of two standard deviations above the mean importance assigned to a particular value-state as a threshold. This way they selected between 232 and 565 descriptions of behaviors that were most strongly related to each of nine value-states. They gave 3,914 descriptions of behavioral acts in total. This pool of responses was reduced to 646 by merging descriptions of the same behavioral act that were written using varying grammar. The reduced pool seems to be representative of everyday behaviors, including routine activities, as well as those more specific and rarely reported. Skimina et al. asked two judges (the same who participated in the assessment of reliability described as Step 5) to group the 646 descriptions of behaviors into approximately 100 categories. They instructed the judges to group activities that were similar in purpose and performance together (if provided). Working together, following the instructions, the judges developed 98 categories.

This set of narrow categories was used by Skimina and Cieciuch (2020) to code all 13,873 responses collected in Study 6. They found that two of the 98 categories were indistinguishable, so they merged these two categories into one (everyday shopping and shopping into shopping). Because some of the 97 categories were very infrequent, they further merged very similar categories into broader ones. For instance, they merged hosting friends, visiting friends, and meeting friends in a public place into spending time with friends. This reduced the 97 categories into 63. The Cohen's kappa for the 97 categories was .81 and for the 63 categories it was .88 (Skimina & Cieciuch, 2020).

The sets of 63 and 97 categories can be perceived as two lower levels that can be added to the hierarchical system of behavioral categories, presented in Figure 1. All narrow categories are listed in Table A in the Appendix and descriptions of 63 categories are presented in Table B in the Appendix. The lower levels of the categorization system were validated in studies conducted by Skimina et al. (2019) and by Skimina and Cieciuch (2020) who found that narrow categories differ in terms of their associations with value-states, metraits of personality, and the preferences of higher-order values. For instance, visiting a family member was more strongly related to the benevolence-caring value-state than hosting a family member or meeting a family member in a public place. Playing with a child was strongly related to security value-state, whereas other child-related activities (e.g., training cognitive skills with a child) were not. The security-personal item was: "How important was it to you to avoid danger," which suggests that, in the case of playing with a child, it was not interpreted as security-personal but rather as caring for a child's safety (Skimina et al., 2019). Skimina and Cieciuch (2020) confirmed that narrow categories of behaviors can reveal associations with personality variables that are not observable at a level of broader behavioral categories. They found, for example, that talking with a partner and spending time with a partner are related differently to metraits of personality.

Discussion

Our aim was to develop a categorization system for everyday behaviors. We worked on pools of responses to an open-ended question about recent behavior provided by people aged 12 to 66 years old in six experience sampling studies. We used a bottom-up approach, trying to group a large pool of behaviors into a smaller number of categories based on the similarities between those behaviors. We first developed a basic categorization, consisting of 23 categories, and then created a hierarchical system by adding both higher and lower levels. At the higher levels the 23 categories are grouped into five or three broader ones. The lower levels consist of 63 and 97 categories that were developed in separate analyses. As a result, we introduce a hierarchical categorization system that can be used to code responses to open-ended questions about a behavior/activity in experience sampling and diary studies, conducted on people of different ages—from adolescents to older adults. Depending on the study purpose one can use the categorization with more narrowly defined categories or less broad categories. The reliability and validity of the broad and narrow categories were confirmed in the current study as well as in other work (Skimina et al., 2019; Skimina & Ciecuch, 2020).

The analyses showed a high level of inter-rater agreement for the sets of 23, 63, and 97 categories (Skimina & Ciecuch, 2020), indicating that the categorization system is a reliable tool. Analyses conducted on the pool of 13,873 responses to the ESM form showed that activities assigned to different categories from the pool of 23 vary in terms of emotional states and the momentary importance ascribed to personal values during these activities. Skimina et al. (2019) also showed that the lowest-level categories differ in the importance ascribed to personal values in real time. Skimina and Ciecuch (2020) found associations between the frequencies of a large part of 63 categories and personality dispositions: metatraits and preferences of higher-order values.

It might seem surprising that the system of 22 categories does not include a separate category of using the Internet, especially taking into account the fact that adolescents and young adults—who use the Internet very frequently on a daily basis (Twenge, 2017) — constituted a large part of the samples. In fact, the words *Internet* or *smartphone* were rarely used in the descriptions of activities. Indeed, young people use the Internet for many purposes, for example, to watch videos, to play computer games, to shop, or to read news. We assume that participants in our studies focused on the *purpose* of using the Internet, and did not find it important to mention the medium in the description of activity. From the analysis of the content of the developed categorization, we infer that the usage of the Internet has become so common that it cannot exist as one category of behavior. It includes too diverse subcategories which should be separated (e.g., communicating with friends, watching videos, or reading news). At the level of narrow categories, using the Internet was reduced to surfing the Internet (browsing websites), and this category was reported infrequently (Skimina & Ciecuch, 2020).

The categorization we propose is based not only on activity, but also on situational cues such as social partners and location. The validation analysis showed that mentioning an interaction partner in the response to the open-ended question about activity is related to different experiences (emotional and value states) compared to not mentioning an interaction partner who is present (which is known from a response to another question). Combining information about an activity with information about a social partner enables us to distinguish between activities done in company from activities done alone. However, our results indicate that the class of activities distinguished in this way is qualitatively different from the class distinguished based on the person's description of their behavior. This suggests that mentioning an interaction partner in the description of a behavior indicates that this is important for the participant in their activity. Therefore, this is also an argument for including categories such as Family time or Time with a partner in the coding system of activities.

The categorization system we propose may be used to assign each response describing a behavior to a single category. In our analyses, conducted on the pool of 520 responses, 79.6% were assigned to only one category by all three independent judges. However, some responses may be assigned to more than one category. For instance, drinking coffee in a café with a friend may be assigned to three categories: Time with friends (spending time with a friend), Outdoor leisure (being in a café), and Physiology (drinking coffee). A researcher may decide whether to assign responses like this to only one or several categories at the same time.

This proposed categorization is designed to be used to code open-ended questions about activities asked in an ESM study, but the list of categories could also be used in closed-ended questions. For example, when a participant is asked to select the option that best describes his or her current activity, the list of 22 categories could be provided. Another possibility is to create a drop-down menu with a category narrowing functionality: from 22 categories to more specific subcategories. A drop-down menu could also be used to determine the purpose of an activity—for instance, whether drinking coffee should be described as resting or socializing. The advantage of providing a list of categories is that it does not demand spending additional time on coding responses in the dataset. However, asking an open-ended question also has its merits. Responses to open-ended questions give researchers more possibilities than closed-ended ones. For instance, they differ in style, and this response style difference may be analyzed as an individual difference: Some people tend to respond briefly, using only one word, whereas others tend to provide long and detailed descriptions. Open-ended responses may be analyzed in various ways, depending on different criteria. Researchers may use responses many times using different coding schemes to suit various purposes.

As the categorization was developed using a bottom-up approach, and almost every description of behavior can be assigned to one or more distinguishable categories (only 1.1% of 13,873 responses were assigned to categories Others or Unclassifiable), we believe

that its use may be broader than just for coding responses in ESM studies. The proposed categorization may also be treated as a system of behavioral or situational classes (including information about activity, social partner, and location). Further, it could be seen as a quite universal classification of the types or dimensions of everyday behaviors in which people engage. Obviously, further research is needed in other countries to verify the cross-cultural replicability of this categorization.

To summarize, we introduce a categorization of behaviors that can be used for coding responses to an open-ended question about behavior in experience sampling and diary studies in Poland and probably in other European countries. This categorization has a hierarchical form, and therefore it offers various sets of categories (from a small number of broad categories to a large number of narrow categories) that can be used depending on the research goals. This categorization is customized to samples at different ages—from adolescents to older adults. It refers to behavior in a broader context—and is based not only on information about the activity, but also on situational cues (social partners and locations), if provided by a respondent. This categorization system was developed in a bottom-up approach using a pool of responses to a question about recent behavior collected at different times of the day over the course of a few months. The validity of the categorization has already been empirically supported in some published studies on the relations between behavior measured by ESM and personality (values and metatraits; Skimina et al., 2019, Skimina & Ciecuch, 2020). We hope that the categorization we proposed can be used in many studies aimed at describing and explaining behavior – a major goal in psychology (Doliński, 2018). As the whole procedure of development of the behavioral categorization was described and documented in detail in this paper, it can be easily improved on in the future if necessary.

Funding: The work of ES, DK, ETS, MKG, KP, RR was supported by Grant 2018/28/T/HS6/00224 from the National Science Centre, Poland. The work of ED, JC was supported by the University Research Priority Program Social Networks of the University of Zurich.

Competing Interests: The authors have declared that no competing interests exist.

Acknowledgments: The authors have no support to report.

Data Availability: For this study, a dataset is freely available (Skimina et al., 2020).

Supplementary Materials

The data used in this research are openly accessible at the Open Science Framework (for access see Index of [Supplementary Materials](#) below).

Index of Supplementary Materials

Skimina, E., Karaś, D., Topolewska-Siedzik, E., Kłym-Guba, M., Ponikiewska, K., Rogoza, R., . . . Ciecuch, J. (2020). *A categorization of behaviors reported in experience sampling studies* [Research data]. OSF. <https://doi.org/10.17605/OSF.IO/9Q3DF>

References

- Bohnert, A. M., Richards, M., Kohl, K., & Randall, E. (2009). Relationship between discretionary time activities, emotional experiences, delinquency and depressive symptoms among urban African American adolescents. *Journal of Youth and Adolescence*, *38*, 587-601. <https://doi.org/10.1007/s10964-008-9336-1>
- Conner, T. S., Feldman Barrett, L., Tugade, M. M., & Tennen, H. (2007). Idiographic personality: The theory and practice of experience sampling. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 79–96). New York, NY, USA: The Guilford Press.
- Csikszentmihalyi, M., & Larson, R. (1984). *Being adolescent: Conflict and growth in the teenage years*. New York, NY, USA: Basic Books.
- Doliński, D. (2018). Is psychology still a science of behaviour? *Social Psychological Bulletin*, *13*(2), Article e25025. <https://doi.org/10.5964/spb.v13i2.25025>
- Fleeson, W. (2007). Situation-based contingencies underlying trait-content manifestation in behavior. *Journal of Personality*, *75*, 825-862. <https://doi.org/10.1111/j.1467-6494.2007.00458.x>
- Furr, R. M. (2009). Personality psychology as a truly behavioural science. *European Journal of Personality*, *23*, 369-401. <https://doi.org/10.1002/per.724>
- Graham, J. M. (2008). Self-expansion and flow in couples' momentary experiences: An experience sampling study. *Journal of Personality and Social Psychology*, *95*, 679-694. <https://doi.org/10.1037/0022-3514.95.3.679>
- Grund, A., Grunschel, C., Bruhn, D., & Fries, S. (2015). Torn between want and should: An experience-sampling study on motivational conflict, well-being, self-control, and mindfulness. *Motivation and Emotion*, *39*, 506-520. <https://doi.org/10.1007/s11031-015-9476-z>
- Hegarty, R. S. M., Treharne, G. J., Stebbings, S., & Conner, T. S. (2016). Fatigue and mood among people with arthritis: Carry-over across the day. *Health Psychology*, *35*, 492-499. <https://doi.org/10.1037/hea0000321>
- Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA, USA: SAGE.
- Kubey, R., & Csikszentmihalyi, M. (1990). *Television and the quality of life: How viewing shapes everyday experience*. Hillsdale, MI, USA: Erlbaum.
- Larson, R. W., Richards, M. H., Sims, B., & Dworkin, J. (2001). How urban African American young adolescents spend their time: Time budgets for locations, activities, and companionship. *American Journal of Community Psychology*, *29*, 565-597. <https://doi.org/10.1023/A:1010422017731>

- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*, 276-282. <https://doi.org/10.11613/BM.2012.031>
- Mueller, S., Ram, N., Conroy, D. E., Pincus, A. L., Gerstorf, D., & Wagner, J. (2019). Happy like a fish in water? The role of personality-situation fit for momentary happiness in social interactions across the adult lifespan. *European Journal of Personality*, *33*, 298-316. <https://doi.org/10.1002/per.2198>
- Robinson, J. P., & Godbey, G. (1999). *Time for life: The surprising ways Americans use their time* (2nd ed.). University Park, PA, USA: Pennsylvania State University.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1-65). New York, NY, USA: Academic Press.
- Schwartz, S. H., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., . . . Konty, M. (2012). Refining the theory of basic individual values. *Journal of Personality and Social Psychology*, *103*, 663-688. <https://doi.org/10.1037/a0029393>
- Shernoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: Quality of experience from the perspective of participants. *Journal of Youth and Adolescence*, *36*, 891-903. <https://doi.org/10.1007/s10964-007-9183-5>
- Skimina, E., & Cieciuch, J. (2020). Explaining everyday behaviors and situational context by personality metatraits and higher-order values. *European Journal of Personality*, *34*(1), 29-59. <https://doi.org/10.1002/per.2230>
- Skimina, E., Cieciuch, J., Schwartz, S. H., Davidov, E., & Algesheimer, R. (2018). Testing the circular structure and importance hierarchy of value states in real-time behaviors. *Journal of Research in Personality*, *74*, 42-49. <https://doi.org/10.1016/j.jrp.2018.02.001>
- Skimina, E., Cieciuch, J., Schwartz, S. H., Davidov, E., & Algesheimer, R. (2019). Behavioral signatures of values in everyday behavior in retrospective and real-time self-reports. *Frontiers in Psychology*, *10*, Article 281. <https://doi.org/10.3389/fpsyg.2019.00281>
- Snippe, E., Simons, C. J., Hartmann, J. A., Menne-Lothmann, C., Kramer, I., Booij, S. H., . . . Wichers, M. (2016). Change in daily life behaviors and depression: Within-person and between-person associations. *Health Psychology*, *35*, 433-441. <https://doi.org/10.1037/hea0000312>
- Sperry, S. H., Lynam, D. R., Walsh, M. A., Horton, L. E., & Kwapil, T. R. (2016). Examining the multidimensional structure of impulsivity in daily life. *Personality and Individual Differences*, *94*, 153-158. <https://doi.org/10.1016/j.paid.2016.01.018>
- Sun, J., Harris, K., & Vazire, S. (2019). Is well-being associated with the quantity and quality of social interactions? *Journal of Personality and Social Psychology*. Advance Online Publication. <https://doi.org/10.1037/pspp0000272>
- Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment*, *21*, 457-462. <https://doi.org/10.1037/a0017653>

- Twenge, J. (2017). *iGen. Why today's super-connected kids are growing up less rebellious, more tolerant, less happy—and completely unprepared for adulthood—and what that means for the rest of us*. New York, NY, USA: Atria Books.
- van Roekel, E., Goossens, L., Verhagen, M., Wouters, S., Engels, R. C. M. E., & Scholte, R. H. J. (2014). Loneliness, affect, and adolescents' appraisals of company: An experience sampling method study. *Journal of Research on Adolescence*, *24*, 350-363. <https://doi.org/10.1111/jora.12061>
- Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher*, *44*, 7-16. <https://doi.org/10.3102/0013189X14566879>

Appendix

Table A
The Sets of Narrow Categories Assigned to Broader Ones

97 categories	63 categories	23 categories	5 categories or other
Creative hobby (e.g., singing, dancing, playing instruments)	Creative hobby	Creative hobbies	Entertainment
Reading	Reading	Indoor leisure and resting	Entertainment
Watching or reading news	News	Indoor leisure and resting	Entertainment
Getting new information on one's hobby	Hobby	Indoor leisure and resting	Entertainment
Hobby (e.g., fishing)	Hobby	Outdoor leisure	Entertainment
Listening to music	Listening to music	Indoor leisure and resting	Entertainment
Intellectual entertainment (e.g., crosswords)	Intellectual entertainment	Indoor leisure and resting	Entertainment
Board games	Board games	Indoor leisure and resting	Entertainment
Computer games	Computer games	Indoor leisure and resting	Entertainment
Resting	Resting	Indoor leisure and resting	Entertainment
Watching TV/movies	Watching TV/movies	Indoor leisure and resting	Entertainment
Surfing the Internet	Surfing the Internet	Indoor leisure and resting	Entertainment
Cultural events (e.g., concert, theatre play)	Cultural events	Indoor leisure and resting	Entertainment
Ability games (e.g., billiards, bowling)	Ability games	Outdoor leisure	Entertainment
Sightseeing	Sightseeing	Outdoor leisure	Entertainment
Hiking	Walking	Outdoor leisure	Entertainment
Physical activity	Sports	Physical activity	Entertainment
Do-it-yourself	Do-it-yourself	Repairing	Household
Formal errands (e.g., at the post office)	Formal errands	Errands	Household
Informal errands (e.g., making an appointment at a beauty salon)	Informal errands	Housework	Household
Caring for a family (e.g., cooking)	Caring for a family	Housework	Household
Cleaning	Cleaning	Housework	Household
Cooking	Cooking	Housework	Household
Shopping online	Shopping online	Shopping	Household
Shopping at a store	Shopping	Shopping	Household
Buying gifts online	Buying gifts	Shopping	Household
Buying gifts at a store	Buying gifts	Shopping	Household
Medical errands	Medical care	Health and personal care	Maintenance
Everyday getting ready to go out	Getting ready	Health and personal care	Maintenance
Beauty treatment at a beauty salon	Beauty treatment	Health and personal care	Maintenance
Beauty treatment at home	Beauty treatment	Health and personal care	Maintenance
Getting ready for a party	Getting ready for a party/date	Health and personal care	Maintenance
Getting ready for a date	Getting ready for a party/date	Health and personal care	Maintenance
Eating	Eating	Physiology	Maintenance

97 categories	63 categories	23 categories	5 categories or other
Sleeping	Sleeping	Physiology	Maintenance
Hygiene	Hygiene	Physiology	Maintenance
Drinking alcohol	Drinking alcohol	Physiology	Maintenance
Traveling	Traveling	Transportation	Maintenance
Waiting	Waiting	Waiting	Maintenance
Looking for a job	Looking for a job	Other	Other
Other	Other	Other	Other
Planning	Planning	Other	Other
Surveys, lotteries	Surveys, lotteries	Other	Other
Thinking and internal life	Thinking and internal life	Other	Other
Helping someone one knows	Helping	Other	Other/Socializing
Helping a stranger	Helping	Other	Other/Socializing
Someone else's important events (e.g., wedding)	Important events	Other	Other/Socializing
One's important event	Important events	Other	Other/Socializing
Passive engagement in social activity	Social activity	Other	Other/Socializing
Active engagement in social activity (e.g., volunteering)	Social activity	Other	Other/Socializing
Caring for a child (e.g., cooking)	Babysitting	Family time	Socializing
Training cognitive skills with a child	Babysitting	Family time	Socializing
Training physical skills with a child	Babysitting	Family time	Socializing
Playing with a child	Babysitting	Family time	Socializing
Participating in important events in a child's life	Babysitting	Family time	Socializing
Talking with a family member face to face	Talking to family	Family time	Socializing
Talking with a family member on the phone or online	Talking to family	Family time	Socializing
Hosting a family member	Time with family	Family time	Socializing
Meeting a family member in a public place	Time with family	Family time	Socializing
Visiting a family member	Time with family	Family time	Socializing
Talking to a stranger face to face	Talking to strangers	Other socializing	Socializing
Talking with acquaintances face to face	Talking to acquaintances	Time with friends	Socializing
Talking with acquaintances on the phone or online	Talking to acquaintances	Time with friends	Socializing
Talking with friends face to face	Talking to friends	Time with friends	Socializing
Talking with friends on the phone or online	Talking to friends	Time with friends	Socializing
Hosting acquaintances	Time with acquaintances	Time with friends	Socializing
Meeting acquaintances in a public place	Time with acquaintances	Time with friends	Socializing
Visiting acquaintances	Time with acquaintances	Time with friends	Socializing
Hosting friends	Time with friends	Time with friends	Socializing
Meeting friends in a public place	Time with friends	Time with friends	Socializing
Visiting friends	Time with friends	Time with friends	Socializing
Dating	Date	Time with partner	Socializing

97 categories	63 categories	23 categories	5 categories or other
Sexual activity	Sexual activity		Socializing
Talking with a partner face to face	Talking to partner	Time with partner	Socializing
Talking with a partner on the phone or online	Talking to partner	Time with partner	Socializing
Hosting a partner	Time with partner	Time with partner	Socializing
Meeting a partner in a public place	Time with partner	Time with partner	Socializing
Visiting a partner	Time with partner	Time with partner	Socializing
Walking pets	Caring for a pet	Time with pets	Socializing
Caring for pets (e.g., feeding)	Caring for a pet	Time with pets	Socializing
Playing with pets	Caring for a pet	Time with pets	Socializing
Training pets	Caring for a pet	Time with pets	Socializing
Taking exams at school	Exams	Studying and extra classes	Work and studying
Taking exams at a university	Exams	Studying and extra classes	Work and studying
Taking exams at extra classes	Exams	Studying and extra classes	Work and studying
Participating in extra classes	Extra classes	Studying and extra classes	Work and studying
Participating in school lessons	In lecture	School/University classes	Work and studying
Participating in lectures at a university	In lecture	School/University classes	Work and studying
Studying for school individually	Studying	Studying and extra classes	Work and studying
Studying for school in a group of students	Studying	Studying and extra classes	Work and studying
Studying for university classes individually	Studying	Studying and extra classes	Work and studying
Studying for university classes in a group of students	Studying	Studying and extra classes	Work and studying
Studying for extra classes individually	Studying	Studying and extra classes	Work and studying
Studying for extra classes in a group of students	Studying	Studying and extra classes	Work and studying
Working	Working	Work	Work and studying
Religious practices	Religious practices	Religious practices	
Unclassifiable	Unclassifiable	Unclassifiable	

Table B*Descriptions of 63 Categories of Activities*

No	Category	Description
1	in lecture	participating in lectures/classes at university/school
2	studying	studying, preparing for lectures/classes/exams
3	exams	having exams
4	extra classes	participating in extra classes (e.g., foreign language course)
5	working	working, also from home
6	looking for a job	looking for a job, having a job interview
7	babysitting	caring for a child, e.g., feeding, dressing, playing with a child
8	caring for a family	e.g., preparing meals, shopping for family members
9	caring for a pet	e.g., feeding, playing with a pet
10	eating	eating
11	sleeping	sleeping, napping
12	hygiene	having shower, bathing, using toilet
13	cleaning	cleaning house
14	cooking	cooking, preparing meals
15	[formal] errands	errands at the post office, bank, etc.
16	informal errands	e.g., making an appointment at a beauty salon
17	medical care	seeing a doctor, caring for one's health
18	getting ready	dressing, putting on make-up, etc.
19	do-it-yourself	repairing, redecorating
20	shopping online	doing shopping online
21	shopping	doing shopping
22	buying gifts	buying a present for someone
23	cultural events	seeing a movie/play, attending a concert, etc.
24	beauty treatment	having a massage or a beauty treatment at a beauty salon
25	drinking alcohol	drinking alcohol
26	sports	playing sports, jogging, working out, having fitness classes
27	reading	reading books, magazines
28	news	reading/watching news
29	listening to music	listening to music
30	ability games	playing ability games, e.g., table tennis
31	intellectual entertainment	e.g., sudoku, puzzles
32	board games	playing board games, e.g., scrabble
33	computer games	playing computer games
34	hobby	e.g., fishing
35	creative hobby	e.g., playing instrument, singing, painting
36	resting	resting, lying on the bed/sofa
37	watching TV/movies	watching TV/movies/series/videos
38	surfing the Internet	visiting websites
39	thinking and internal life	reflecting, descriptions of internal experiences
40	sightseeing	sightseeing
41	walking	walking
42	talking to family	talking to family members, also on the phone or online
43	talking to partner	talking to a partner, also on the phone or online
44	talking to friends	talking to friends, also on the phone or online
45	talking to acquaintances	talking to acquaintances, also on the phone or online
46	talking to strangers	talking to a stranger, also on the phone or online
47	time with family	spending time with family members
48	time with partner	spending time with a partner
49	time with friends	spending time with friends
50	time with acquaintances	spending time with acquaintances
51	date	being on a date
52	important events	participating in a wedding, christening, etc.

No	Category	Description
53	social activity	participating in a meeting of community organization, volunteering
54	helping	helping someone
55	waiting	waiting for someone or something
56	surveys, lotteries	participating in surveys or lotteries
57	planning	making plans
58	sexual activity	sexual activities
59	getting ready for a party/date	dressng, putting on make-up for a party or date
60	traveling	traveling somewhere, coming back home
61	religious practices	praying, participating in a holy mass
62	other	other activities
63	Unclassifiable	nonwords or unintelligible responses

Note. The table was reprinted from Skimina and Ciecuch (2020).



Social Psychological Bulletin is an official journal of the Polish Social Psychological Society (PSPS).



leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology Information (ZPID), Germany.