

Kids in Taiwan: National Longitudinal Study of Child Development and Care (KIT): KIT-M36 at 60 months old

User Guide

I. Overview

Principal Investigator	Chien-Ju Chang
Execution Institution	Center for Educational Research and Evaluation & Department of Human Development and Family Studies, National Taiwan Normal University
Funding Agency	Ministry of Science and Technology
Span of Project Execution	2017-11-01~2020-07-31
Span of Investigation	2018-03-01~2019-04-30

II. Released File Description

1. User guide

userguide.doc	Includes description of released files, guideline for data validation and usage, guideline for citation and contact us etc.
---------------	---

2. Questionnaire

ques1.pdf	Questionnaire for parents or primary caregivers who provide care for the child for the majority of time
ques2.pdf	Questionnaire for relatives, friends, or domestic nannies who provide care for the child during the day
ques3.pdf	Questionnaire for nannies/caregivers/educarers at daycare centers or kindergartens

3. Datasets

data1.csv	Comma delimited dataset (parents or primary caregivers)
data1.sav	SPSS dataset (parents or primary caregivers)
data1.dta	STATA dataset (parents or primary caregivers)
data2.csv	Comma delimited dataset (relatives, friends, or domestic nannies)
data2.sav	SPSS dataset (relatives, friends, or domestic nannies)
data2.dta	STATA dataset (relatives, friends, or domestic nannies)
data3.csv	Comma delimited dataset (for nannies/caregivers/educarers at daycare centers or the kindergartens)

data3.sav	SPSS dataset (for nannies/caregivers/ educarers at daycare centers or the kindergartens)
data3.dta	STATA dataset (for nannies/caregivers/ educarers at daycare centers or the kindergartens)

4. Other

M36-M72 questionnaire domains_reliability&validity.pdf	Domains, reliability and validity of questionnaires
weighting.pdf	Weighting calculation

III. Guideline for data validation and usage

1. CSR has checked wild codes and out-of-range values, to validate and clean data. At each subsequent wave of data collection, the KIT project cross-checked response accuracy across waves. However, some inconsistent data could not be verified as the respondents could not be reached. Please be aware that these data have not been removed and are included in the datasets. Please also refer to 7(5) below for more details. For example, variables that asked about times of certain events such as parentsam, parentspm, sleep01b, sleep01c (in data1) or am, pm (in data3) may not have been converted to the 24-hour time format (e.g., “900” may need to be converted to “2100”) because they could not be verified after the interviews. In addition, other variables that asked about times may appear unreasonable, such as sleep01a, sleep01d (in data1). Responses that do not appear to be reasonable are listed in the table below:

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part A II. 2.1	data1	pfb0202a	2008	1
			2012	1
ques1 Part A II. 2.1	data1	pfb0202b	1922	1
			2012	1
ques1 Part A II. 2.1	data1	pfb0204a	2013	1
ques1 Part C I. 35	data1	famevic15a	6	3
ques1 Part C I. 35	data1	famevic1501a	4	5
ques1 Part E II.2	data1	postn02	48	1
ques1 Part E II.3	data1	postn03	216	1
			260	1
ques1 Part E II.6	data1	postn06a	888	1

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part E III.1	data1	heigh	10.5	1
			11.0	1
			17.0	1
ques1 Part E III.1	data1	heighy	109	2
ques1 Part E III.2	data1	weight	2.9	1
			42	1
ques1 Part E III.2	data1	weighty	109	2
ques1 Part E VI	data1	growheight01	height value too high at 6M	2
ques1 Part E VI	data1	growweight01	weight value too low at 6M	1
			weight value too high at 6M	11
ques1 Part E VI	data1	growy01	reported time of measurement too far from 6M	24
ques1 Part E VI	data1	growheight02	height value too high at 12M	2
ques1 Part E VI	data1	growweight02	weight value too high at 12M	9
ques1 Part E VI	data1	growy02	reported time of measurement too far from 12M	26
ques1 Part E VI	data1	growheight03	height value too low at 15M	1
ques1 Part E VI	data1	growweight03	weight value too high at 15M	3
ques1 Part E VI	data1	growy03	reported time of measurement too far from 15M	17
ques1 Part E VI	data1	growheight04	height value too high at 24M	1
ques1 Part E VI	data1	growweight04	weight value too low at 24M	1
			weight value too high at 24M	1
ques1 Part E VI	data1	growy04	reported time of measurement too far from 24M	14
ques2 Part A I.8	data2	pfc08	number of weekly hours too low	1
ques3 I.2	data3	basic02 basic0201	1929	1
ques3 I.5	data3	basic08 basic0801	0	1
ques3 I.5	data3	basic09 basic0901	non-integer	2
ques3 Part II I.1	data3	pm	2330	2

For responses that do not appear to be reasonable across question items, please see Appendix A.

2. The column widths of variable label and value description in the STATA are limited to 80 and 40 bytes, respectively, so the STATA dataset shows only the first 80 bytes as variable label, and the first 40 bytes as value description. In case of cut-off contexts, please see questionnaire (or code book).
3. The questionnaires developed by KIT project cannot be used elsewhere for data collection without formal permission. If you intend to use items from our questionnaires, please contact the KIT team via phone (02)77493142 or email kidsintw@gmail.com.

4. Population and Sampling Principle:

Two groups of children are followed, and their developmental condition and family and care information were collected by the Kids in Taiwan: National Longitudinal Study of Child Development and Care (KIT) project. Both samples are nationally representative. The sampling frame of KIT-M3 consists of children who live in Taiwan (excluding offshore islands) and were registered in the Household Registration System and whose birthdays are between 2016/4/1 and 2017/6/30. The sampling frame of KIT-M36 consists of children who live in Taiwan (excluding offshore islands) and were registered in the Household Registration System and whose birthdays are between 2013/4/1 and 2014/3/31. The stratified two-stage probability-proportional-to-size sampling method was used with county and individual as the primary and secondary sampling units respectively. Because the east region of Taiwan is less populated and the population located mostly in fewer populated areas, only individual level was used during the sampling process. Stratification was based on the system created by Center for Survey Research, Research Center for Humanities and Social Science, Academia Sinica and Project of Taiwan Social Change Survey in 2014. The original sample of KIT-M3 includes 13,692 children, and the original sample of KIT-M36 includes 4,260 children.

5. Sample

The KIT project follows two national representative samples, KIT-M3 and KIT-M36. Data collection plan is to follow up with the sample from 0 to 8 years old. KIT-M3 is scheduled to be interviewed at 3, 6, 12, 18, 24, 36, 48, 60, 72, 84, and 96 months old, and KIT-M36 is scheduled to be interviewed at 36, 48, 60, 72, 84, and 96 months old. The data released is the third wave (60 months old) of data collection for the KIT-M36.

6. Questionnaires

The instruments that KIT project uses for data collection include parent questionnaire, relative-and-friend questionnaire, and educator questionnaire. Researchers can combine the data from the parent questionnaire, the relative-and-friend questionnaire, and the educator questionnaire by matching the variable, `baby_id`.

- (1) If the child is cared for by parents during the day, the parent questionnaire is administered and answered by the parent.
- (2) If the child is cared for by a relative or a domestic nanny during the day and is cared for by parent(s) during the night or weekend, two questionnaires are administered, i.e., (1) parent questionnaire is answered by the parent, and (2) the relative-and-friend questionnaire is answered by the daytime caregiver.
- (3) If the child goes to a nanny's home, a daycare center or a kindergarten, during the day and is cared for by parent(s) during the night or weekend, two questionnaires are administered, i.e., (1) parent questionnaire is answered by the parent, and (2) the educator questionnaire is answered by the nanny or the caregiver at the daycare center or the kindergarten.
- (4) If the child is cared for by a relative or a live-in nanny 24 hours a day for a prolonged period, parent questionnaire is administered and is answered by the relative or nanny.
- (5) If the child goes to a nanny's home, a daycare center or a kindergarten during the day and is cared for by a relative or a live-in nanny during the night and most of the weekends for a prolonged period, two questionnaires are administered, i.e., (1) parent questionnaire is answered by the relative or live-in-nanny, and (2) the educator questionnaire is answered by the daytime nanny or the caregiver at the daycare center or the kindergarten.

7. Others

- (1) In order to observe the accurate status of child development, the proposed interview should be executed 15 days before or after the child reaches 3 months old or 6 months old. For children over 1 year old, the proposed interview date is 30 days before or after the child's birthday. Sometimes it is difficult to schedule an interview in the proposed interview period. The variable, `int_date`, is the date when the interview was conducted. Researchers

can use `int_date` and child's birthday, `baby_dob`, to identify if the interview was conducted in the proposed period.

- (2) Researchers should pay attention to the variable, `relationship`, in parent questionnaire and the relative-and-friend questionnaire. The variable identifies the person who answers the questionnaire. Parent questionnaire may be answered by people other than parents, and the relative-and-friend questionnaire may be answered by a domestic nanny.
- (3) The parent questionnaire asks if the child goes to a day care center or a kindergarten. Some respondents reported no on this question, but later the interviewer found that the child did go to a center or school. Therefore, an educator questionnaire was still administered. Researchers may decide how to use inconsistent information across different questionnaires.
- (4) Only one person should answer the parent questionnaire. However, for a number of cases, there was a change of respondents during the interview. These errors could be related to any event that occurred during the interview, such as attending to a crying baby of the first respondent, etc. For example, changing respondents would lead to inconsistency between variable `pfb13` and `fammar07`. Researchers may decide how to use the information based on research purposes.
- (5) Before releasing the datasets, the project staff checked for any inconsistent information across different waves. Inconsistent answers were verified with the respondents via phone call and the answers were corrected. If the inconsistent answers were unable to be verified, the original answers were retained in the file. Researchers should check the data when conducting cross-wave analysis, and prepare the data according to their own research purpose.
- (6) The special codes for the response options (i.e., the value labels) are 7777 ("refused to answer"), 8888 ("don't know"), 9996 ("not applicable/question skipped"), 9999 ("missing").
- (7) Some notes about `data1` (parents or primary caregivers):
 - i. Part E is Health Questionnaire. In the released copy of the parent questionnaire (`ques1`), it was specified that some sections or question items—question items that asked about the condition of the mother during pregnancy and of the child after birth (e.g., Sections I and II), and the health

condition of the child at the time of the interview (e.g., questions 3 and 4 in Section III)—were intended to be collected only once at the beginning wave of data collection (i.e., children aged 36 months). However, since some samples were not interviewed in the first wave, such question items were then asked again in this wave (i.e., 60M): prec01, prec02, prec03, prec04, prec0401, prec0402, prec0403, prec05, prec06, postn01, postn02, postn03, postn04, postn0501, postn0502, postn0503, postn0504, postn0505, postn0506, postn0507, postn0508, postn0509, postn0510, postn0511, postn0512, postn0513, postn06, postn06a, postn07, health01, health0101, health0102, health03, health0301, health030101, health030102, health030103, health030104, health030105, health030106, health030107, health030108, health030109, health030110, health030111, health0302, med0902, med0902a, vac01, growd01, growd02, growd03, growd04, groweigh01, groweigh02, groweigh03, groweigh04, groweigh01, groweigh02, groweigh03, groweigh04, growm01, growm02, growm03, growm04, growy01, growy02, growy03, growy04.

- ii. Variables not included in the prior waves of data collection but added to this wave (i.e., 60M) are sd1, sd2, sd3, sd4, sd5, care1, online1, online2, online3, online4, online5, online6—among these variables, online1-online6 were collected only at this wave. For these question items, please see the three sets of questions in the following three tables (sd1-sd5 in the first set; care1 in the second set; online1-online6 in the third set):

sd1, sd2, sd3, sd4, sd5:

How often do you feel or do the following?

No.	Item	Often	Sometimes	Seldom	Never
1	There is fun in my life.	1	2	3	4
2	I am happy.	1	2	3	4
3	I feel like taking advantage of others.	1	2	3	4
4	I am envious of others' good fortune.	1	2	3	4
5	I gossip about others.	1	2	3	4

care1:

Is the child care facility (where the child is receiving care) public or private?

☐①public

☐②private

☐③non-profit

☐④don't know

online1, online2, online3, online4, online5, online6 (available only for this wave):

1. Do you have the following equipment at home? 1. Internet connection

☐①no

☐②yes

2. Do you have the following equipment at home? 2. desktop computer(s)

☐①no

☐②yes

3. Do you have the following equipment at home? 3. laptop(s)

☐①no

☐②yes

4. Do you have the following equipment at home? 4. smartphone(s)

☐①no

☐②yes

5. Do you have the following equipment at home? 5. tablet(s)

☐①no

☐②yes

6. Do you have the following equipment at home? 6. none of the above.

☐①no

☐②yes

(8) Some notes about data2 (relatives, friends or domestic nannies):

Variables not included in the prior waves of data collection but added to this wave (i.e., 60M) are sd1, sd2, sd3, sd4, sd5. See the first question set in 7(7)ii for the five question items.

(9) Some notes about data3 (nannies, caregivers or educarers):

i. Questionnaire respondents include the following three categories: preschool educarers (also referred to as educarers), daycare/nursey care center educarers (also referred to as caregivers), and domestic educarers (also referred to as nannies). Please see variable tsorting for the categories (1 = preschool educarer; 2 = nursey care center educarer; 3 = domestic educarer).

- ii. For some variables that provide basic information by educator category (e.g., gender, year of birth, years of work experience), the information is available (and sometimes repeated) through multiple variables. For example, gender information is provided through the following four variables: tgender (gender for preschool educators and nursery care center educators combined), tgender01 (gender for preschool educators alone), tgender02 (gender for nursery care center educators alone), tgender03 (gender for domestic educators alone). Such variables also include basic0501, basic0502, basic0503, basic0504, basic0505, basic0506, basic0507, basic0508, basic0101, basic0201, basic030101, basic030201, basic0801, basic0901, basic0102, basic0202, basic030102, basic030202, basic0802, basic0902, basic0103, basic0203.
- iii. For researchers who are interested in looking at data for child care programs and organizations altogether, data3 provides basic information for preschool and nursery care center combined. Please refer to basic02, basic0301, basic0302, basic08, basic09. Such aggregated data are provided for your convenience.
- iv. In the released copy of the educator questionnaire, two questions asked the respondents whether their organizations participated in particular Ministry of Education programs. At the time of the 60M interviews, these two questions no longer applied and were omitted. For the two questions, please see Question 6 in Part I, “In the past three years did the institution you work participate in any program sponsored by the Ministry of Education?” (See p.1-2, end of Basic Information for preschool educators) and Question 6 in Part I, “In the past three years did the institution you work participate in any program sponsored by the Ministry of Education?” (See p. 3-4, end of Basic Information for nursery care center educators).
- v. Variables not included in the prior waves of data collection but added to this wave (i.e., 60M) are sd1, sd2, sd3, sd4, sd5. See the first question set in 7(7)ii for the five question items.

IV. Guideline for citation

Citation, based upon the requirement of various journals, theses or dissertations, can be revised.

Chien-Ju Chang. (2022). Kids in Taiwan: National Longitudinal Study of Child Development and Care (KIT): KIT-M36 at 60 months old (D00223) [Data file]. Available from Survey Research Data Archive, Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica.
doi:10.6141/TW-SRDA-D00223-2

IRB/REC

The KIT project was approved by the research ethics committee of National Taiwan University (No. 201408ES007) and by the research ethics committee of National Taiwan Normal University (No. 201707HS003).

V. Contact us

Survey Research Data Archive (SRDA), Center for Survey Research, Research Center for Humanities Social Sciences, Academia Sinica, Taipei, Taiwan

Email: srda@gate.sinica.edu.tw

The above contents are edited by the Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica, Taipei, Taiwan. Please respect research ethics. Data use is limited to applicants only. Please do not copy or forward to others.

Appendix A

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part A I. 7	data1	pfa07	1	4
		pfa070101	anything but 9996	
ques1 Part A I. 7	data1	pfa07	1	5
		pfa070201	anything but 9996	
ques1 Part A I. 7	data1	pfa07	1	2
		pfa070101a	anything but 9996	
ques1 Part A I. 7	data1	pfa07	1	2
		pfa070101b	anything but 9996	
ques1 Part A I. 7	data1	pfa0701	0	2
		pfa070101	anything but 9996	
		pfa070101a		
		pfa070101b		
ques1 Part A I. 7	data1	pfa070101a	greater than 3 years (i.e., maximum number of years allowed) collectively/maximum number of years larger than 2	3
		pfa070101b		
ques1 Part A I. 7	data1	pfa07	1	1
		pfa070102a	anything but 9996	
ques1 Part A I. 7	data1	pfa07	1	1
		pfa070102b	anything but 9996	
ques1 Part A I. 7	data1	pfa070102a	greater than 3 years (i.e., maximum number of years allowed) collectively	1
		pfa070102b		
ques1 Part A I. 7	data1	pfa07	1	5
		pfa070201	anything but 9996	
		pfa070201a		3
		pfa070201b		
ques1 Part A I. 7	data1	pfa0702	0	2
		pfa070201	anything but 9996	
		pfa070201a		
		pfa070201b		
ques1 Part A I. 7	data1	pfa070201a	greater than 3 years (i.e.,	12

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
		pfa070201b	maximum number of years allowed) collectively/maximum number of years larger than 2	
ques1 Part A I. 7	data1	pfa070203a	greater than 3 years (i.e., maximum number of years allowed) collectively/maximum number of years larger than 2	2
		pfa070203b		
ques1 Part A I. 7	data1	pfa070204a	greater than 3 years (i.e., maximum number of years allowed) collectively/maximum number of years larger than 2	3
		pfa070204b		
ques1 Part A II. 1-2	data1	pfb0115	1	1
		pfb0115a	2	
		pfb0204	0	
		pfb0204a	9996	
		pfb0204b	9999	
ques1 Part A II. 3-5	data1	pfb03	1	7
		pfb050a	8	
ques1 Part A II. 4	data1	pfb04	1	1
		pfb04a	2	
ques1 Part A II. 5	data1	pfb0501a	0	1
		pfb0501b	1	
ques1 Part A II. 7	data1	pfb07	2	1
		pfb0703	anything but 9996	
ques1 Part C I. 31-32	data1	famevic1101	1	21
		famevic1202	1	
ques1 Part C I. 31-32	data1	famevic1102	1	10
		famevic1202	1	
ques1 Part C I. 31-32	data1	famevic1105	1	3
		famevic1202	1	
ques1 Part C I. 31-32	data1	famevic1101 - famevic1105	0 for all	9
		famevic1202	anything but 1	
ques1 Part C I. 31-32	data1	famevic1106	0	57
		famevic1201	anything but 1	

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	29
		famevic150201a		
		famevic150201b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	6
		famevic150202a		
		famevic150202b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	1
		famevic150203a		
		famevic150203b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	3
		famevic150205a		
		famevic150205b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	4
		famevic150206a		
		famevic150206b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	34
		famevic150207a		
		famevic150207b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	17
		famevic150208a		
		famevic150208b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	40
		famevic150209a		
		famevic150209b		
ques1 Part C I. 35	data1	famevic15a	greater than child's age (i.e., 60 months) collectively	8
		famevic150210a		
		famevic150210b		
ques1 Part C I. 35	data1	famevic15	2	4
		famevic150201 - famevic150210	0 for all	
ques1 Part E I. 2	data1	prec02	difference between reported and derived maternal age at childbirth (i.e., from mother's and child's year of birth) greater than 3	23
		pfa0102a		
		baby_dob		
ques1 Part E I. 3-4	data1	prec03	1	4
		prec04	anything but 9996	

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part E I. 4	data1	prec04	1	15
		prec0401	anything but 9996	
ques1 Part E I. 4	data1	prec04	1	14
		prec0402	anything but 9996	
ques1 Part E I. 4	data1	prec04	1	15
		prec0403	anything but 9996	
ques1 Part E I. 4	data1	prec04	2	2
		prec0401 - prec0403	0 for all	
ques1 Part E II. 5	data1	postn0501	1	2
		postn0502	1	
ques1 Part E II. 5	data1	postn0501	1	14
		postn0503	1	
ques1 Part E II. 5	data1	postn0501	1	2
		postn0504	1	
ques1 Part E II. 5	data1	postn0501	1	2
		postn0505	1	
ques1 Part E II. 5	data1	postn0501	1	2
		postn0506	1	
ques1 Part E II. 5	data1	postn0501	1	5
		postn0510	1	
ques1 Part E II. 5	data1	postn0501	1	12
		postn0511	1	
ques1 Part E II. 5	data1	postn0501	1	1
		postn0513	1	
ques1 Part E II. 2, 3, 5	data1	postn0502	1	4
		postn02	1	
		postn03	birth weight value high enough	
ques1 Part E II. 2, 3, 5	data1	postn0502	0	1
		postn02	born pre-term	
		postn03	birth weight value too low	
ques1 Part E II. 3, 5	data1	postn0504	0	11
		postn03	birth weight lower than 2000	
ques1 Part E II. 3, 5	data1	postn0504	1	2
		postn03	birth weight higher than 2500	

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part E II. 3, 5	data1	postn0505	1	1
		postn03	3500	
ques1 Part E II. 6-7	data1	postn06a	30	1
		postn07	1	
ques1 Part E II. 6-7	data1	postn07	5	1
		postn06	1	
ques1 Part E III. 8	data1	med01	1	6
		med0101	anything but 9996	
ques1 Part E III. 8	data1	med01	1	5
		med0201	anything but 9996	
ques1 Part E III. 8	data1	med01	1	6
		med0301	anything but 9996	
ques1 Part E III. 8	data1	med01	1	6
		med0401	anything but 9996	
ques1 Part E III. 8	data1	med01	1	5
		med0501	anything but 9996	
ques1 Part E III. 8	data1	med01	1	2
		med0601	anything but 9996	
ques1 Part E IV. 1	data1	food01	1	1
		food01a	anything but 9996	
ques2 Part B I. 34-35	data2	famevic1101	1	4
		famevic1202	1	
ques2 Part B I. 34-35	data2	famevic1102	1	3
		famevic1202	1	
ques2 Part B I. 34-35	data2	famevic1101 - famevic1105	0 for all	2
		famevic1202	anything but 1	
ques2 Part B I. 34-35	data2	famevic1106	0	10
		famevic1201	anything but 1	
ques2 Part B I. 34	data2	famevic1107	1	1
		famevic1101- famevic1106	anything but 0	
ques2 Part B I. 34-35	data2	famevic1107	1	2
		famevic1201	anything but 1	
ques2 Part B I. 34-35	data2	famevic1107	1	1

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
		famevic1202	anything but 1	
ques3 I.4	Data3	basic0402	1	1
		basic0407	anything but 0	