Kids in Taiwan: National Longitudinal Study of Child Development and Care (KIT): KIT-M3 at 12 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	2014-08-01~2020-12-31
Span of Investigation	2017-03-01~2018-07-31

II. Released File Description

1. User guide

userguide.docx	Including released file description, 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 preschools

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 (nannies/caregivers/ educarers at daycare centers or preschools)
data3.sav	SPSS dataset (nannies/caregivers/educarers at daycare centers or preschools)
data3.dta	STATA dataset (nannies/caregivers/educarers at daycare centers or preschools)
IRT.csv	Comma delimited dataset (ability estimates of cognitive, language, social, emotional, physical developments)
IRT.sav	SPSS dataset (ability estimates of cognitive, language, social, emotional, physical developments)
IRT.dta	STATA dataset (ability estimates of cognitive, language, social, emotional, physical developments)

4. Other

M3-M24 questionnaire domains_reliability&validity. pdf	The result of the reliability and validity analysis of the Questionnaire
weighting.pdf	Weighting calculation
list_variables_by_wave.xlsx	Full list of variable information (variable names, variable labels, scale/domains) by wave

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 (See Table 1):

Table 1

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part A I.7.1.2.3	data1	pfa070203b	24	2
ques1 Part C I.38.2	data1	famevic150208a	1	1
ques1 Part E III.1	data1	heigh	11.2	1
			11.6	1
not printed in ques2. pdf	data2	relationship	3	13
			28	16
ques3 Part B I.1	data3	am	442	1
			1800	1

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.

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 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 is used

with county and person as the primary and secondary sampling units respectively. Because the population in the east region of Taiwan is smaller and concentrated in fewer areas, only the personal level is used during the sampling process. Stratification is based on the system created by Center for Survey Research, Academia Sinica 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 (birth cohort) and KIT-M36 (child cohort). KIT-M3 is interviewed at 3, 6, 12, 18, 24, 36, 48, 60, 72 months old, and every year at the elementary school level until Grade 6 (a total of 15 waves). KIT-M36 is interviewed at 36, 48, 60, 72 months old, and every year at the elementary school and junior high school levels until Grade 9 (a total of 13 waves). The data released is the 3rd wave (12 months old) of data collection for KIT-M3. As the goal of the project is to retain as many participants as possible, those who did not complete the interview in prior wave(s) were not excluded in subsequent wave(s).

6. Questionnaires

The instruments that KIT project uses for data collection include parent questionnaire, relative-and-friend questionnaire, and educarer questionnaire. Researchers can combine the data from the parent questionnaire and the friends-and-relatives questionnaire by matching the variable, *release_id*.

(1) If the child is cared for by parent(s) during the day, the parent questionnaire is administered and is 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 on weekends, two questionnaires are administered: the parent questionnaire is answered by the parent; the friends-and-relatives questionnaire is answered by the daytime caregiver.

(3) When 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 on weekends, the parent questionnaire is answered by the parent.

(4) If the child is cared for by a relative or a live-in nanny 24 hours a day for a prolonged period, the parent questionnaire is administered and is answered by the caregiver.

(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 weekends for a prolonged period, the parent questionnaire is answered by the caregiver.

(6) Data collection for the KIT project is mixed-mode; respondents completed surveys via face-to-face interviews or online (with online responses coded as 99999 for *interviewer_id*).

7. Others

(1) In order to observe the accurate status of child development, the proposed interview should be executed 15 days before and after the child reaches 3 months old and 6 months old. For children over 1 year old, the proposed interview date is 30 days before and after the child's birthday. Sometimes it is difficult to schedule an interview in the proposed interview period. In the datasets, *int_months* provides the number of months (i.e., child's age in months) at the time of interview completion. Researchers may decide whether to keep data collected outside the proposed 30-day interview range. In this data wave (i.e., 18M), due to a birthday error, the interview for one particular child's primary caregivers took place six months earlier than intended. Researchers may decide whether to include the information in the analysis.

(2) Researchers should pay attention to the variable *relationship* in the parent questionnaire and the friends-and-relatives questionnaire. The variable identifies the person who answers the questionnaire. The parent questionnaire may be answered by people other than the parent, and the friends-and-relatives 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" for this question, but the interviewer later found that the child did go to a center or school. Therefore, the educarer questionnaire is still administered. Researchers may decide how to use inconsistent information discovered between 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, etc. Change of respondents could lead to inconsistency among these variables: *pfb13* and *fammar07*. Researchers will need to decide how to approach the information.

(5) Before releasing the datasets, the project staff checked for any inconsistent information between different waves. Inconsistent answers were verified with the respondents via phone call and the answers were corrected. If the inconsistent answers were unable to verify, 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) Data labels for non-responses include 7777 ("refused to answer"), 8888 ("don't know"), 9996 ("not applicable/question skipped"), 9999 ("missing").

(7) Social and emotional development: In ques1 (parents or primary caregivers), there are four areas/scales that measure children's competence: cognition, language, social-emotional, and physical development. For cognition, language, and physical development, researchers can obtain a total (mean) score from each scale as an overall estimate of a child's competence. However, for social-emotional development, such an approach does not apply, as (1) social and emotional development are two separate sub-scales; and (2) the entire scale includes sub-domains besides social and emotional competence (such as temperament, secure attachment, withdrawal, aggression, shyness, and shame). To obtain a score for social competence and another score for emotional competence separately, please refer to questionnaire items listed under these domains in the following two files: *M3-M24 questionnaire domains_reliability&validity* (PDF file) and *list_variables_by_wave* (xlsx file).

(8) More notes about data1 (parents or primary caregivers):

  i. Some question items collecting basic information that was not likely to change over a period of six months (i.e., the interview time gap between 6M and 12M) were not asked again for this wave. These variables are *pfa0101, pfa0101a, pfa0102, pfa0102a, pfa0201, pfa0202, pfa03, pfa03a, pfa0301, pfa0301a, pfa0302, pfa0302a, pfa0303, pfa030401, pfa030402, pfa030403, pfa030404, pfa04, pfa04a, pfa0401, pfa0401a, pfa0402, pfa0402a, pfa0403, pfa040401, pfa040402, pfa040403, pfa040404, pfa05, pfa05a, pfa08, pfa09, pfa1501, pfa1501a, pfa1502, pfa1502a, pfb0101, pfb0102, pfb0103, pfb0104, pfb0105, pfb0106, pfb0107, pfb0108, pfb0109, pfb0110, pfb0111, pfb0112, pfb0113, pfb0114, pfb0115, pfb0116, pfb0117, pfb0118, pfb0119, pfb0120, pfb0121, pfb0122, pfb0112a, pfb0113a, pfb0114a, pfb0115a, pfb0116a, pfb0117a, pfb0120a, pfb02, pfb0201, pfb0202, pfb0203, pfb0204, pfb0201a, pfb0201b, pfb0201c, pfb0201d, pfb0201e, pfb0201f, pfb0201g, pfb0201h, pfb0201i, pfb0202a, pfb0202b, pfb0202c, pfb0202d, pfb0202e, pfb0202f, pfb0202g, pfb0202h, pfb0202i, pfb0203a, pfb0203b, pfb0203c, pfb0203d, pfb0203e, pfb0203f, pfb0203g, pfb0203h, pfb0203i, pfb0204a, pfb0204b, pfb0204c, pfb0204d, pfb0204e, pfb0204f, pfb0204g, pfb0204h, pfb0204i, pfb12*, med0902, *med0902a*. The KIT project integrated responses from KIT-M3 at 3M and 6M into this dataset (i.e., 12M) when the surveys were answered by the same respondent. The following variables are made available accordingly: *pfa0101, pfa0101a, pfa0102, pfa0102a, pfa0201, pfa0202, pfa03, pfa0301, pfa0301a, pfa0302, pfa0302a, pfa03a, pfa0303, pfa030401, pfa030402, pfa030403, pfa030404, pfa04, pfa0401, pfa0401a, pfa0402, pfa0402a, pfa04a, pfa0403, pfa040401, pfa040402, pfa040403, pfa040404, pfa05, pfa05a, pfa08, pfa09, pfa1501, pfa1501a, pfa1502, pfa1502a*. With the data integration, however, approximately four percent of the respondents still have missing data for these variables. Researchers may consider integrating survey responses provided by different respondents (across waves), but please be aware that data accuracy could be an issue.

  ii. At the time of 12M interviews, a question (variable *pfb14*) was added to see if primary caregivers were interviewed during the prior year. For those who were, this question served as the skip logic for questions that collect basic info that are not likely to change. These variables are *pfc01, pfc0201, pfc0202, pfc03, pfc03a, pfc0301, pfc0301a, pfc0302, pfc0302a, pfc0303, pfc030401, pfc030402, pfc030403, pfc030404, pfc04, pfc05, pfc06, pfc09*. The KIT project integrated responses from KIT-M3 at 3M and 6M into this

dataset (i.e., 12M) when the surveys were answered by the same respondent. The following variables are made available accordingly: *pfc01, pfc0201, pfc0202, pfc03, pfc04, pfc05, pfc06, pfc09*. With the data integration, approximately six percent of the respondents still have missing data for these variables. Researchers may consider integrating survey responses provided by different respondents (across waves), but please be aware that data accuracy could be an issue.

iii. Variables not included in the prior waves of data collection but added to this wave (i.e., 12M) are *pfb050a, med0901, med0901a, city1, city2*:

① *med0901*, *med0901a*: These questions asked whether the child had otitis media (i.e., middle ear infection) within the six months prior to the interview, and the number of times the child caught such infection. Please refer to Question 10 on p. 46 of the parent questionnaire (under II. "Condition of the child after birth" in Part E. "Health Questionnaire").

② *pfb050a* (Table 2), *city1*, and *city2* (Table 3): These questions were added to this wave of data collection (KIT-M3 at 12 months old). Please see below for these questions, or refer to the dataset.

Table 2

pfb050a:
Relationship between current daytime caregiver and child: □① father □② mother □③ paternal grandfather/grandmother □④ maternal grandfather/grandmother □⑤ friend □⑥ nanny (providing care at nanny's place) □⑦ nanny (providing care at child's place) □⑧ educarer at kindergarten/day care center □⑨ other

Table 3

city1, city2:

Where in Taiwan has the child lived most of his/her life?

□① Taipei City   □② New Taipei City   □③ Keelung City

□④ Taoyuan City   □⑤ Hsinchu City   □⑥ Hsinchu County

□⑦ Miaoli County   □⑧ Taichung City   □⑨ Changhua County

□⑩ Nantou County   □⑪ Yunlin County   □⑫ Chiayi City

□⑬ Chiayi County   □⑭ Tainan City   □⑮ Kaohsiung City

□⑯ Pingtung County   □⑰ Yilan County   □⑱ Hualien County

□⑲ Taitung County

□⑳ Outlying islands of Taiwan (e.g., Penghu, Kinmen, Matsu, Lanyu, etc.)

□㉑ Other countries/Not in Taiwan (Please specify: _____)

(9) More notes about data2 (relatives, friends or domestic nannies):

At the time of the 12M interviews, a question (variable *pfb14*) was added to see if primary caregivers were interviewed during the prior year. For those who were, this question served as the skip logic for questions that collect basic info that are not likely to change. These variables are *pfc01, pfc0201, pfc0202, pfc03, pfc03a, pfc0301, pfc0301a, pfc0302, pfc0302a, pfc0303, pfc030401, pfc030402, pfc030403, pfc030404, pfc04, pfc05, pfc06, pfc0701, pfc0702, pfc09, pfc09a, pfc10, pfc1101, pfc1102*. The KIT project integrated responses from KIT-M3 at 3M and 6M into this dataset (i.e., 12M) when the surveys were answered by the same respondent. The following variables are made available accordingly: *pfc01, pfc0201, pfc0202, pfc03, pfc0301, pfc0301a, pfc0302, pfc0302a, pfc03a, pfc0303, pfc030401, pfc030402, pfc030403, pfc030404, pfc04, pfc05, pfc06, pfc09*. With the data integration, approximately four to six percent of the respondents still have missing data for these variables. Researchers may consider integrating survey responses provided by different respondents (across waves), but please be aware that data accuracy could be an issue.

(10) More notes about data3 (nannies, caregivers or educarers), added to the third wave (i.e., 12M) of KIT-M3 data collection:

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 info by educarer 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 info for preschool educarers and nursery care center educarers combined), *tgender01* (gender info for preschool educarers alone), *tgender02* (gender info for nursery care center educarers alone), *tgender03* (gender info for domestic educarers 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 info 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 educarer questionnaire, two questions asked the respondents whether their organizations participated in particular Ministry of Education programs. At the time of the 12M 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?" (p.1-2, end of Basic Information for preschool educarers) 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?" (p. 3-4, end of Basic Information for nursery care center educarers).

IV. Guideline for citation

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

Chien-Ju Chang. (2024). Kids in Taiwan: National Longitudinal Study of Child Development and Care (KIT): KIT-M3 at 12 months old (D00214) [Data file]. Available from Survey Research Data Archive, Center for Survey Research, Research Center for Humanities and Social Sciences, Academia Sinica. https://doi.org/10.6141/TW-SRDA-D00214-3

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

Appendix A

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part A I.7	data1	pfa07	1	2
		pfa0702	1	
ques1 Part A I.7	data1	pfa07	2	2
		pfa0701, pfa0702	0	
ques1 Part A I.7	data1	pfa0701	0	3
		pfa070101	1, 2	
ques1 Part A I.7	data1	pfa0702	0	2
		pfa070201	1, 3	
ques1 Part A I.7	data1	pfa070201	2	1
		pfa070201a, pfa070201b	5, 6	
ques1 Part A I.7	data1	pfa070201	2	1
		pfa070201a, pfa070201b	4, 6	
ques1 Part A I.7	data1	pfa070201	2	1
		pfa070201a, pfa070201b	2, 9999	
ques1 Part A I.7	data1	pfa070201	2	1
		pfa070203a, pfa070203b	3, 3	
		pfa070204a, pfa070204b	3, 3	
ques1 Part A I.7	data1	pfa070201	2	1
		pfa070203a, pfa070203b	2, 9999	
		pfa070204a, pfa070204b	2, 24	
ques1 Part B II.3-5	data1	pfb03	1	6
		pfb03a, pfb03b, pfb04, pfb04a, pfb050a, pfb05a, pfb05b, pfb0501a, pfb0501b, pfb0502	anything but 9996	
ques1 Part A II.6	data1	pfb06	1	1
		pfb0601 - pfb0608	0 for all	
ques1 Part A II.7	data1	pfb07	2	1
		pfb0701	1	
ques1 Part A II.7	data1	pfb07	2	4
		pfb0702	1	

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part A II.7	data1	pfb07	2	1
		pfb0705	1	
ques1 Part A II.7	data1	pfb07	1	1
		pfb0701 - pfb0708	0 for all	
ques1 Part A II.8	data1	pfb080301	1	~~8~~ 9
		pfb080302	1	
ques1 Part A II.8	data1	pfb080301	1	1
		pfb080304	1	
ques1 Part A II.8	data1	pfb080301	1	1
		pfb080305	1	
ques1 Part A II.8	data1	pfb080301	1	2
		pfb080307	1	
ques1 Part A II.8	data1	pfb080301	0	4
		pfb080302 - pfb080307	0 for all	
ques1 Part A II.8	data1	pfb080302	0	13
		pfb08030201	1, 3	
ques1 Part A II.8	data1	pfb080401	1	2
		pfb080402	1	
ques1 Part A II.8	data1	pfb080401	1	1
		pfb080405	1	
ques1 Part A II.8	data1	pfb080401	1	2
		pfb080406	1	
ques1 Part A II.8	data1	pfb080401	0	5
		pfb080402 - pfb080406	0 for all	
ques1 Part A II.8	data1	pfb080402	0	2
		pfb08040201	1	
ques1 Part A II.9	data1	pfb09	1	1
		pfb09a, pfb09a	0, 4	

ques1 Part A III.3
    Part A III.3.1
    Part A III.3.2
    Part A III.3.3
	data1	pfc03	1	1

		pfc0301	1	
		pfc0302	2	
		pfc0303	1	

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part A III.3 Part A III.3.1 Part A III.3.2 Part A III.3.3	data1	pfc03	1	1
		pfc0301	1	
		pfc0302	1	
		pfc0303	3	
ques1 Part C I.34	data1	famevic1107	1	6
		famevic1101	1	
ques1 Part C I.34	data1	famevic1107	1	1
		famevic1102	1	
ques1 Part C I.34	data1	famevic1107	1	1
		famevic1105	1	
ques1 Part C I.34	data1	famevic1107	1	9
		famevic1106	1	
ques1 Part C I.34	data1	famevic1107	1	1
		famevic1108	1	
ques1 Part C I.35.2 Part C I.34	data1	famevic1202	1	161
		famevic1101	1	
ques1 Part C I.35.2 Part C I.34	data1	famevic1202	1	23
		famevic1102	1	
ques1 Part C I.35.2 Part C I.34	data1	famevic1202	1	1
		famevic1103	1	
ques1 Part C I.35.2 Part C I.34	data1	famevic1202	1	3
		famevic1104	1	
ques1 Part C I.35.2 Part C I.34	data1	famevic1202	1	12
		famevic1105	1	
ques1 Part C I.35.2 Part C I.34	data1	famevic1201	1	33
		famevic1106	1	
ques1 Part C I.38	data1	famevic15	2	6
		famevic1501	2	
ques1 Part C I.38	data1	famevic15	2	1
		famevic1501	1	
ques1 Part C I.38	data1	famevic15	2	1
		famevic150201 - famevic150210	0 for all	

Questionnaire item	Dataset	Variable	Value (Not reasonable)	Count
ques1 Part C I.40	data1	famevic17	2	1
		famevic1701 - famevic1706	0 for all	
ques1 Part E III. 8.2	data1	med0201	1	1
		med0204	1	
ques1 Part E III.11.1 Part E III.11.1.1 Part E III.11.1.1	data1	med07	1	1
		med07a	1	
		med07b	6	
ques1 Part E IV.1	data1	food01	2	1
		food01a	1	
ques1 Part E V.1 Part E V.1.1	data1	sleep01	2	1
		sleep01a	14	
ques2 Part A I.4 Part A I.7	data2	pfc04	1	1
		pfc0701	8	
ques2 Part A I.4 Part A I.11	data2	pfc04	1	23
		pfc1101	7	
ques2 Part A I.4 Part A I.11	data2	pfc04	1	2
		pfc1102	7	
ques2 Part A I.34	data2	famevic1107	1	3
		famevic1101	1	
ques2 Part A I.34	data2	famevic1107	1	1
		famevic1106	1	
ques2 Part A I.35 Part A I.34	data2	famevic1201	1	19
		famevic1106	1	
ques2 Part A I.35 Part A I.34	data2	famevic1202	1	30
		famevic1101	1	
ques2 Part A I.35 Part A I.34	data2	famevic1202	1	5
		famevic1102	1	
ques2 Part A I.35 Part A I.34	data2	famevic1202	1	4
		famevic1105	1	