TwinLife

Documentation TwinLife Data: Cognitive Abilities

v.2.0.0

F2F1, F2F2, & F2F4

by Christoph H. Klatzka & Lena Paulus

Christoph.klatzka@uni-saarland.de

UNIVERSITÄT BIELEFELD
Fakultät für Soziologie

Universität Bremen

UNIVERSITÄT DES SAARLANDES

Christoph H. Klatzka & Lena Paulus
Documentation TwinLife Data: Cognitive Abilities v2.0.0
F2F1, F2F2, & F2F4

TwinLife Technical Reports are refereed scholarly papers. Submissions are reviewed by the general editors before a final decision on publication is made.

The Technical Report Series is a forum for presenting technical works (e.g., data documentation, field reports) in progress. Comments on the manuscript should be addressed directly to the author(s).

The papers can be downloaded from the project website:
https://www.twin-life.de/twinlife-series

TwinLife "Genetic and social causes of life chances"   Phone: +49 (0)521 106-4309
University of Bielefeld   Email: martin.diewald@uni-bielefeld.de
Faculty of Sociology   Web: https://www.twin-life.de
PO Box 100131
D-33501 Bielefeld
Germany

UNIVERSITÄT BIELEFELD
Fakultät für Soziologie

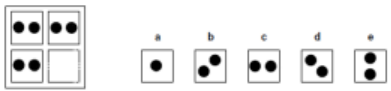Universität Bremen

UNIVERSITÄT DES SAARLANDES

Description of the cognitive test battery

To test cognitive abilities, the CFT-1-R (Culture Fair Test) with three subtests (figural reasoning, figural classification, and matrices) and the CFT-20-R with four subtests (figural reasoning, figural classification, matrices, and reasoning) were used. These tests are designed to measure non-verbal fluid intelligence as a proxy for general cognitive ability (Weiß, 2006; Weiß & Osterland, 2012).

In the first and second face-to-face household survey (F2F1 & F2F2), the CFT-20-R was guided by the interviewer and was conducted on a laptop in the participant's home, while the CFT-1-R was conducted via interviewer guided paper-pencil test. In F2F1, the cognitive abilities test was conducted for all children aged 9 or younger (CFT-1-R, N= 2,399) and respondents aged 10 or older (CFT-20-R, N= 13,775). In F2F2, new respondents were also given the opportunity to take the cognitive abilities test using the same test batteries (CFT-1-R: N = 19; CFT-20-R: N = 677). In rare cases, participants were falsely treated as a new respondent to the sample in F2F2 and therefore took the test a second time (N = 25).

Due to the COVID pandemic during the F2F4 interviews, the household interviews were converted into a mixed mode of telephone interviews and online-based survey methods. Therefore, the test administration had to be modified and the CFT assessment was conducted online via the Hogrefe test system, but for twins and siblings only. Since all participants were now over 9 years old, only the CFT-20-R was used (N = 3,115). In contrast to the first two measurements, the instructions were given in the test system and the participants completed the test on their own.

CFT-20-R (age range: 8;5 to 60)   **CFT-1-R** (age range: 5;3 to 9;11)

Target age: 10+	Target age: 5-9
F2F1 + F2F2: Computer administrated	F2F1+ F2F2: Paper-pencil administrated
F2F4: Hogrefe test system (online)	F2F4: Not administrated
Total test time: approx. 15-25 min	Total test time: approx. 30 min
Test time per subtest:	Test time per subtest:
Subtest 1-3: 4 min + 1 min	Subtest 1-3: 3 min + 1 min
Subtest 4: 3 min + 1min	

In both test versions, there is a short and a long test time. The long test time is one minute longer than the short test time. If the short test time was over, the answers of the participants were stored in one set of variables (igf*40 – ig*54). All answers given after the short test time were recorded in another set of variables (igf*60- igf*74). As participants always had the opportunity to go back to the items, answers could also be changed, potentially leading to deviations in answers for the long and short version. However, using the Hogrefe test system in F2F4, no distinction was made between long and short versions, so that there is just one set of item answers.

Calculation of the test scores

Correct answers were coded with 1 and wrong answers were coded with 0. Non-Response was coded -88, while items that were answered only in the short version were coded -89 in the long version's variables. Multiple responses on an item were coded -87 for the CFT-1-R. -94 indicates technical errors in the CFT-20-R. Participants that were not in the target population for the CFT tests were assigned other missing codes (-95, -90, -92).

The generated variables with the information as to whether an item was solved correctly are named as following for F2F1 and F2F2:

- CFT-20-R:
  - Short version: igf0140-igf0154, igf0240-igf0254, igf0340-igf0354, igf0440-igf0450
  - Long version: igf0160-igf0174, igf0260-igf0274, igf0360-igf0374, igf0460-igf0470
- CFT-1-R:
  - Short version: igf0540-igf0554, igf0640-igf0654, igf0740-igf0754
  - Long version: igf0560-igf0574, igf0660-igf0674, igf0760-igf0774

For each subtest, the number of correctly solved items were summed to a sum score. Three different scores are available for the short version (correct answers given in the short test time), the long version (change in correct answers during the additional minute), and the total version (correct answers during the total test time).

The generated variables with sum scores are named as following for F2F1 and F2F2:

- CFT-20-R:
  - Short version sum score: igf0180, igf0280, igf0380, igf0480
  - Long version sum score: igf0181, igf0281, igf0381, igf0481
  - Total sum score: igf0182, igf0282, igf0382, igf0482
- CFT-1-R:
  - Short version sum score: igf0580, igf0680, igf0780
  - Long version sum score: igf0581, igf0681, igf0781
  - Total sum score: igf0582, igf0682, igf0782

Since the Hogrefe test system did not distinguish between a short and a long version of the test, only variables combining both short and long test time into one variable and total sum scores are provided. To facilitate the use of the data on item level, a harmonized version of the variables as the complete version (igf1*80 - igf1*94, e.g. igf1180) is also provided for the data collections in F2F1 and F2F2. However, the names of the total sum scores (igf0*82) remain unchanged and correspond to the names of F2F1 and F2F2. For more information, please refer to the scales manual of *TwinLife* (Klatzka et al., 2023).

The harmonized variables are named as following for F2F1, F2F2, and F2F4 for each subtest:

- CFT-20-R:
  - Complete version: igf1180-igf1194, igf1280-igf1294, igf1380-igf1394, igf1480-igf1490
  - Total sum score: igf0182, igf0282, igf0382, igf0482

- CFT-1-R:
  - Complete version: igf1580-igf1594, igf1680-igf1694, igf1780-igf1794,
  - Total sum score: igf0582, igf0682, igf0782

Starting with data release v8.0.0, we also provide a variable with the actual answers of the participants in a subtest (raw data) so that users can check the response pattern of the participants themselves. We only provide this for the harmonized variant (total test time). Nines (9) in the variable indicate a missing in the variable (the item was not answered). The variables are named as following:

- CFT-20-R:
  - igf1199, igf1299, igf1399, igf1499

- CFT-1-R:
  - igf1599, igf1699, igf1799

To check for response patterns, it is advisable to convert the variable into a character string and replace the 9 with a "." so that response patterns are easier to recognize visually.

Example pattern for igf1299:
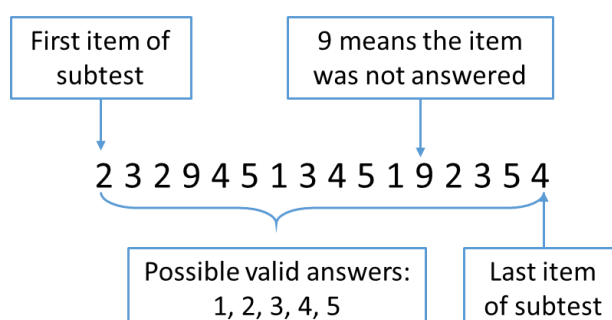
Checking routines

A new approach was developed, in order to avoid a lengthy rating process and inconsistencies in the identification of invalid cases. This new approach was used to ensure the consistency of the checking routine for all data collections with cognitive data. As a result, the checking routine provided in the first technical report (Gottschling et al., 2017) is no longer applied starting with v8.0.0.
This new procedure was applied uniformly to all data collections with cognitive data (i.e. F2F1, F2F2, and F2F4). While many cases previously categorized as invalid remain invalid under the new procedure, the status of invalidity has changed in several cases. Frequencies of valid-missing changes for the data collections are depicted in Table 1. Changes were only made in the generated sum scores; the item level was unaltered.

Table 1.

Frequencies of valid-missing status changes in the subtests per data collection.

	CFT-20-R				CFT-1-R		
	igf0182	igf0282	igf0382	igf0482	igf0582	igf0682	igf0782
F2F1	13,775				2,399		
Changes from missing to valid	44 (0.3%)	37 (0.2%)	17 (0.1%)	54 (0.4%)	145 (6.0%)	148 (6.1%)	160 (6.65)
Changes from valid to missing	13 (0.1%)	22 (0.1%)	11 (0.1%)	120 (0.9%)	159 (6.6%)	67 (2.8%)	63 (2.6%)
No changes in missing status	287 (2.1%)	141 (1.0%)	85 (0.6%)	172 (1.2%)	150 (6.2%)	115 (4.8%)	123 (5.1%)
No changes in valid scores	13,431 (97.5%)	13,575 (98.5%)	13,662 (99.1%)	13,429 (97.5%)	1,945 (81.1%)	2,069 (86.2%)	2053 (85.6%)
F2F2	667				19		
Changes from missing to valid	7 (1.0%)	2 (0.2%)	2 (0.2%)	22 (3.3%)	0	0	0
Changes from valid to missing	0	0	0	4 (0.6%)	0	1 (5.3%)	0
No changes in missings	75 (11.2%)	55 (8.2%)	50 (7.5%)	54 (8.1%)	0	0	0
No changes in valid scores	582 (87.3%)	610 (91.5%)	615 (92.2%)	587 (88.0%)	19 (100%)	18 (94.7%)	19 (100%)

Note. Percentages may not add up to 100 % due to rounding.

The following checking routines were applied for both test forms:

CFT-20-R:

1. A subtest score was set to "-88: test not completed" when less than three items have been answered in a subtest

2. We used the following rules to decide if participants showed an invalid response pattern:

    a. Response pattern: Only up to two response options out of five were used (e.g. igf1199 was 222222223333333 or 555555555555555). For these cases, we conducted the following check:

        i. Since the first items of subtest 1 to 3 were very easy to solve (mean item difficulty for the first three items in subtest 1: .98; subtest 2: .93; subtest 3: .98, difficulties nearing one are indicating very easy items), a previously detected response pattern in Step 2a was set valid if the participant answered at least two of the first three items correctly, indicating that the response pattern could have occurred because all following items were too difficult to answer for this participant.

        ii. Since subtest four was overall more difficult than the first three subtests (mean item difficulty for the first three items: .70), we loosened the criteria. A case with a previously detected response pattern was set valid if at least two of the first four items were answered correctly.

3. In summary, all cases exhibiting an answering pattern and answering less than two of the first three or four items correctly were set as missing (-85: subtest invalid).

4. If a participant had three invalid or missing subtest scores, the remaining one was also set missing (-86: case invalid).

CFT-1-R:

1. A subtest score was set to "-88: test not completed" when less than three items have been answered in a subtest.

2. We used the following indicator to decide if participants showed an invalid response pattern:

    a. Only up to two response options out of five were used.

    b. The items of the CFT-1-R were significantly more difficult for the targeted age group as the CFT-20-R, making the criteria of correctly solved first items of the subset inapplicable.

    c. All cases showing a response pattern were set as "-85: subtest invalid".

3. If an individual had two invalid or missing subtest scores, the remaining one was also set as missing (-86: case invalid).

Corrections

In the following the numbers of subtest scores that have been set missing by the checking procedure are depicted.

F2F1

CFT-20-R:

N = 13,775	Subtest 1	Subtest 2	Subtest 3	Subtest 4
-88: test not completed	275	117	67	183
-86: case invalid	16	3	6	7
-85: subtest invalid	8	43	23	102

CFT-1-R:

N =2,399	Subtest 1	Subtest 2	Subtest 3
-88: test not completed	99	80	92
-86: case invalid	39	28	28
-85: subtest invalid	171	74	66

F2F2

CFT-20-R:

N = 670	Subtest 1	Subtest 2	Subtest 3	Subtest 4
-88: test not completed	76	55	50	54
-86: case invalid	0	0	0	2
-85: subtest invalid	0	0	0	7

CFT-1-R:

N = 29	Subtest 1	Subtest 2	Subtest 3
-88: test not completed	0	1	0
-86: case invalid	0	0	0
-85: subtest invalid	0	0	0

F2F4

CFT-20-R:

N = 3,115	Subtest 1	Subtest 2	Subtest 3	Subtest 4
-88: test not processed	9	11	8	50
-86: case invalid	4	0	0	0
-85: subtest invalid	2	8	4	30

Recommendation for the use of the cognitive abilities test (adapted from Gottschling, 2017):

- It is recommended to use the sum scores of the subtests as provided in the SUF (Scientific Use Files). The sum scores (igf0*82) are corrected for invalid cases, indicated by scores set to -88, -86 or -85.
- Users can choose to use the uncorrected scores by calculating a new sum score based on the original answers.
- Children younger than 5 should be excluded from the analyses in F2F1 as the CFT 1-R is not normed for this age group.
- Adequate algorithms should be used to substitute missing data, excluding cases that did not take part in the cognitive test.
- Analyses should be conducted within cohorts when dealing with twins from different age cohorts.
- Age effects should be controlled for in analyses across cohorts, including cognitive test data from the twins' parents.
- For comparisons with the normative sample of the CFT-20-R or CFT-1-R, refer to Weiß (2006) or Weiß and Osterland (2012).

References

Gottschling, J. (2017). *Documentation TwinLife Data: Cognitive Abilities* (TwinLife Technical Report Series, 02). Bielefeld: Project TwinLife "Genetic and social causes of life chances" (Universität Bielefeld / Universität des Saarlandes).

Klatzka, C. H., Baum, M. A., Paulus, L., Nikstat, A., Dang, E. T. T., Iser, J., & Hahn, E. (2023). *TwinLife Scales Manual.* (TwinLife Technical Report Series, 08, v3.0.1). Project TwinLife – "Genetic and Social Causes of Life Chances" (Universität Bielefeld; Universität Bremen; Universität des Saarlandes).

Weiß, R. (2006). *CFT 20-R. Grundintelligenztest Skala 2. Manual.* Hogrefe.

Weiß, R. H., & Osterland, J. (2012). *Grundintelligenztest Skala 1 - Revision: CFT 1-R.* Hogrefe.