View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by UNH Scholars' Repository

University of New Hampshire
University of New Hampshire Scholars' Repository

UNH Personality Lab | Research Institutes, Centers and Programs

1-31-2019

# Technical Supplement for the article "Advances in Measuring Personal Intelligence"

John D. Mayer
*University of New Hampshire, Durham*, jack.mayer@unh.edu

David R. Caruso
*Yale University*

Abigail T. Panter
*University of North Carolina, Chapel Hill*

Follow this and additional works at: https://scholars.unh.edu/personality_lab

Part of the Personality and Social Contexts Commons, and the Social Psychology Commons

Recommended Citation

Mayer, John D.; Caruso, David R.; and Panter, Abigail T., "Technical Supplement for the article "Advances in Measuring Personal Intelligence"" (2019). *UNH Personality Lab*. 31.
https://scholars.unh.edu/personality_lab/31

This Supplement is brought to you for free and open access by the Research Institutes, Centers and Programs at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in UNH Personality Lab by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.

# *Technical Supplement*

for the article

# "Advances in Measuring Personal Intelligence"

8 March 2019

John D. Mayer

University of New Hampshire

David R. Caruso

Yale University

Abigail T. Panter

University of North Carolina

**Please note**: This technical supplement was developed alongside the article it cross-references. The supplement is provided as a more extensive resource for those interested in the details of the project carried out and reported by the authors. The article covers the key points; the technical supplement covers the key points and includes tables and figures from the article, as well as additional tables and figures, further analyses, and other extra detail.

The article to which this supplement refers is published in the *Journal of Intelligence* as:

Mayer, J.D., Caruso, D.R., & Panter, A.T. (2019). Advancing the measurement of personal intelligence with the Test of Personal Intelligence, Version 5 (TOPI 5), *Journal of Intelligence*, 7(1), 1-17, doi: https://doi.org/10.3390/jintelligence7010004

# Table of Contents

# Chapter 1. Key Elements of the Research Program Through December 2017

## Overview of the Test of Personal Intelligence (TOPI) and its Item Clusters

### Description of the TOPI

The Test of Personal Intelligence (TOPI) is an ability-based measure of personal intelligence consisting of a number of multiple choice items. Scoring of correct answers is based on research on personality. We've developed several versions of the TOPI and this Technical Supplement is intended to accompany the article "Advances in Measurement with the Test of Personal Intelligence, Version 5 (TOPI 5). That said, it contains information relevant to the TOPI versions 2 through 5.

All versions of the Test of Personal Intelligence share several characteristics in common. First, all versions are divided into the four areas of problem solving that the original theory of personal intelligence delineates to define and demarcate the intelligence. The areas are: (a) the identification of personality-relevant information, (b) forming models of personality, (c) using personality-relevant information to guide choices, and (d) systematizing plans and goals. Second, each problem-solving area is further divided into subsets of items. Those items are divided into item clusters in the early forms of the test, and more explicitly into subtests beginning with the TOPI 4.

## Breadth in Measurement

The present article poses the question of whether the TOPI 5 is broader in its measurement than the TOPI 4. That raises the question of "How broad is personal intelligence?". This question can be approached in two ways. The first approach asks, "How many different content areas of reasoning about personality might personal intelligence encompass?" The second approach inquires "How many distinct human mental abilities does a person draw on to answer questions about personality?" This is answered with factor analysis and related mathematical models.

### Breadth of Personal Intelligence in Terms of Problem-Solving Areas

The theory of personal intelligence specifies four areas in which people must problem-solve about personality:

> (a) to recognize personally relevant information from introspection and from observing oneself and others, (b) to form that information into accurate models of personality, (c) to guide one's choices by using personality information where relevant, and (d) to systematize one's goals, plans, and life stories for good outcomes.

In past versions of the Test of Personal Intelligence such as the TOPI 1.4, about 9 to 11 item clusters/tasks were examined. In two exploratory studies, we examine the feasibility of adding five more areas that could also be keyed to the research area in personal intelligence. These "new areas of breadth" will be described shortly.

### Breadth in Terms of Human Mental Abilities

The second way to examine breadth is to determine whether personal intelligence draws on more than one area of human mental abilities to solve. Factor analyses of the TOPI indicate that it is composed of two factors, although we believe that these factor analytic results may represent artifacts

of test design and that, in fact, a single unitary factor may provide a better representation (see TOPI 5 article).

***[Text continues after table(s)].***

**buffer text contiguous with table**

Table 1.1

Overview and Reference Guide to TOPI Clusters by Area in the Test of Personal Intelligence versions 2 through 5

| Abbreviated Task Name (with code) | Brief Task Description | |
|---|---|---|
| | **Given:** | **Solve for:** |
| | ***Identifying Personality-Relevant Information (RVx Tasks)*** | |
| Identifying Motives (RVA) | ●several behaviors and/or pursuits | ●the common motive among them |
| Inner States (RVF/RVB) | ●a situation, activity, or role in which a person is engaged. | ● infer a person's inner state from information |
| Evidence about the Self (RVC) | ●a need for information about oneself | ● ways to receive accurate feedback. |
| Inner Experience-to-Behavior (RVD) | ●a person is carrying out a common activity | ● identify an inner experience that likely accompanies that activity. |
| | ***Forming Models of Personality (FMx Tasks)*** | |
| Trait Knowledge (FMA/FMB) | ● that a person possesses two traits, | ● the person's third likely trait. |
| Integrating Information (FMD) | ● several personality-relevant pieces of information | ● a characteristic of the person's knowledge, intellect, or beliefs |
| Discrepancies-Defense (FME) | ● a discrepancy between a person's words and behavior. | ● infer something about a person's defense and coping |
| Act Frequencies (FMF) | ● a person's trait. | ● behaviors associated with it |
| | ***Guiding Choices Using Personality-Relevant Information (GCx Tasks)*** | |
| Trait Inferences (GCA) | ● someone's trait(s) | ● the person's likely reaction in a situation |
| Observers' Trait Ascriptions (GCB) | ● an observer's plans or behaviors around a target individual | ● identify the trait that an observer ascribes to the target person |
| Motivating Memories (GCC) | ● a person's motivational need, identify. | ● the personal memory that will enhance the individual's motivation |
| | ***Systematizing Plans and Goals (SGx Tasks)*** | |
| Goal-Related Subsidiary Actions (SGA) | ● a longer-term goal | ● an intermediate or subsidiary goal, attitude or behavior that could satisfy it |
| Goal Evaluation (SGB/SGC) | ● a person's objective (e.g., to make friends), | ● a goal that likely will create conflicts for the person because it is unrealistic, hard to fulfil, or contradicts the aim. |
| Personality Change (SGE/SGF) | ● a person's intentions and behaviors | ● how ready they are to change |
| | ***Discontinued Tasks*** | |
| Room with a Cue (RVE) | ● a person's physical surrounding | ● infer some relevant traits |
| Trait Judgeability (FMG) | ● several traits | ● which are most visible/judgeable. |
| Misc. Hard TOPI Questions[e] | | |

[a]The exploratory loadings are for the complete task; [b]the TOPI final are for the tasks after they were screened such that only "functional" items remained, i.e., > .35 loadings on their chief factor; < .25 on any secondary. [c]A version of the of the FME task that included several borderline items was re-introduced to augment the task's reliability. [d]The SGE items are indicated in brackets because they were not counted as items in the TOPI 5-ABR proper but instead included as an experimental task-in-development. [e]Not including the augmented 5 from FME. E. From Forms 1.0 and 1.1 (MSC)

**buffer text contiguous with table**

Table 1.2 contains much the same information as in Table 1.1, but also includes examples of items within each item cluster.

**section break next page here:**

**buffer text contiguous with table**

Table 1.2

Overview and Reference Guide to TOPI Tasks by Area in the TOPI 4 through the TOPI 5R, including 2013-2014 Item Research Trials, with Sample Items

| Task Abbr. | ***Task Name*** | Brief Task Description | ***Item Examples*** |
|---|---|---|---|
| | | ***Identifying Personality-Relevant Information (RVx Tasks)*** | |
| RVA | ***Identifying Motives.*** | Find the common motive among several behaviors and/or pursuits | rva1 1. If a person wants to be around people to talk to them and to have a good time, the person is likely going to:<br>❍ be in love (1)<br>❍ express warmth toward someone (2)<br>❍ meet a goal of excellence (3)<br>❍ socialize (4) |
| RVF (with RVB) | ***Inner States*** | Infer a person's inner state from information about their situation, activity, or role in which they are engaged. | rvb1 1. A minute of time will pass most quickly for:<br>❍ A cook who is watching the clock regularly, carefully timing how long to boil asparagus (1)<br>❍ A student in a classroom who is carefully following the teacher's directions to combine chemicals that will fizz when they are mixed (2)<br>❍ A musician who is in a business meeting in which he feels uninvolved and which his agent will need to explain to him later (3)<br>❍ A bus driver who is transporting people on his regular route toward the end of the day (4) |
| RVD | ***Inner Experience-to-Behavior.*** | A person is carrying out a common activity; identify an inner experience that likely accompanies that activity. | rvd1 1. When a person puts his/her best foot forward he or she often:<br>❍ views him or herself as better than before (1)<br>❍ feels worried about being "found out" as a fraud (2)<br>❍ feels ashamed of his or herself (3)<br>❍ comes to resent the effort (4) |
| | | ***Forming Models of Personality (FMx Tasks)*** | |
| FMA (with FMB) | ***Trait Knowledge.*** | Given a person's two traits, identify a third likely trait | fma1 1. A person is depressed and self-conscious. Most likely, she also could be described as:<br>❍ calm and even-tempered (1)<br>❍ anxious and impulsive (2)<br>❍ self-controlled (3)<br>❍ fairly thick-skinned (4) |
| FMD | ***Integrating Information.*** | Deduce something about a person's knowledge, intellect, or | fmd1 1. Given That: A student believes he understands the material for an upcoming math exam. His teacher, who likes him, says the student doesn't understand it well |

| | | | |
|---|---|---|---|
| | | beliefs from several personality-relevant pieces of information about the individual. | enough to do well.The student's friend doesn't know whether the student understands the material or not.<br>The student should conclude:<br>❍ He is unlikable given that is friend is so unhelpful. (1)<br>❍ His teacher might be right and he doesn't know as much as he thought he did. (2)<br>❍ He knows the material – he knows himself best (3)<br>❍ His friend's uncertainty means no one can know until he takes the exam. (4) |
| FME | ***Discrepancies-Defense*** | Infer something about a person's defense and coping from a discrepancy between their words and behavior. | fme1 1. A teacher's performance in class has declined recently. She has also been having a stressful time at home and, recently, her husband asked for a trial separation, but she does not spend much time thinking about this issue. She is critical of her school, her principal and many of her fellow teachers. She is:<br>❍ taking her hard feelings out on her co-workers (1)<br>❍ behaving impulsively and recklessly (2)<br>❍ ignoring her true feelings about the other teachers – she is jealous of their skill (3)<br>❍ denying her real problems which plainly are at home (4) |
| FMF | ***Trait-Behavior Associations*** | Identify what behaviors are associated with a given trait. | fmf1 1. A person who is aloof would most clearly exhibit that quality by:<br>❍ responding to a question or other conversational comment with a monosyllabic response (1)<br>❍ saving money for the future (2)<br>❍ interrupting others during a conversation (3)<br>❍ drawing in her chair at a meeting closer than the others (4) |
| ***Guiding Choices Using Personality-Relevant Information (GCx Tasks)*** | | | |
| GCA | ***Trait Inferences.*** | Given someone's traits, predict how they are likely to react or behave | gca1 1. Ned's boss, Alan, is highly conscientious and orderly. When Alan finds out Ned was late for work, Alan likely:<br>❍ won't care (1)<br>❍ will - at minimum - make a note of it, and may be disturbed by it (2)<br>❍ greet Ned enthusiastically (3)<br>❍ feel distressed and anxious (4) |
| GCB | ***Observers' Trait Ascriptions*** | Identify the trait that an observer ascribes to another person, given the observer's plans or behaviors around the person. | gcb1 1. A college student returned to his room and noticed a scratch on his desk he never had seen before. He immediately suspected his roommate. The student's reaction makes sense if his roommate is...<br>❍ rigid (1)<br>❍ careless (2)<br>❍ deceitful (3)<br>❍ studious (4) |
| GCC | ***Motivating Memories*** | Given a person's motivational need, identify the personal memory that will enhance the individual's motivation. | gcc1 1. When younger, Sam remembered being cut from his baseball team and the humiliation he felt, and how he wondered if he had practiced enough. Sam used this memory to help himself:<br>❍ work harder to achieve a goal (1) |

| | | | |
|---|---|---|---|
| | | | ❍ recall that self-doubt just isn't helpful (2)<br>❍ perform well in a job interview (3)<br>❍ cope with the challenges of shopping for sports equipment (4) |
| ***Systematizing Plans and Goals (SGx Tasks)*** | | | |
| SGA | ***Goal-Related Actions*** | **.** Identify the intermediate or subsidiary goal, attitude or behavior that could satisfy a longer-term goal. | sga1 1. A person wants "to perform at work with excellence". What goal might most promote this?<br>❍ to take a training course to learn to do the job better (1)<br>❍ to be a good leader to others (2)<br>❍ to use forceful, strong actions so as to become a good leader (3)<br>❍ to try to be a good friend (4) |
| SGB (with SGC) | ***Goal Evaluation.*** | Given a person's objective (e.g., to make friends), identify a goal around that objective that likely will create conflicts for a person because it is unrealistic, hard to fulfil, or contradicts the objective. | sgb1 1. A person wants to make friends. Which goal might cause him problems when he pursues new friendships?<br>❍ be a good friend to his friends (1)<br>❍ to be all things to all people (2)<br>❍ to be myself (3)<br>❍ to spend time meeting new people (4) |
| SGE/SGF | ***Personality Change.*** | Given a person's intentions and behaviors, judge how ready they are to change | sge1 1. On average, people reach the peak of their social dominance – their capacity to make a powerful impression on others – during their:<br>❍ entire lifespan (this characteristic doesn't change over time) (1)<br>❍ teens (2)<br>❍ twenties (3)<br>❍ forties (4) |
| ***Tasks Discontinued as to Their Study*** | | | |
| RVC (Discontinued after the 12Rf) | ***Evidence about the Self*** | Identify good methods for getting feedback about oneself. | rvc1 1. Someone who you don't get along with too well at work mentions to your boss that you completed a project very well. Later that day, he asks you for a favor. One reasonable interpretation is that:<br>❍ he wants you to turn down his request to prove that you are no good (1)<br>❍ he wants to tell the boss you turned down his request so as to "take away" the praise he felt he had to give (2)<br>❍ he said positive things to your boss to help persuade you to grant the favor (3)<br>❍ none of the above (4) |
| RVE (Never exited "item research trials") | ***RVE "Room with a Cue"*** | Given a person's physical surrounding, infer something about his/her personality | rve1 1. If you wanted to convince people that you were open and flexible, you would make an office that looked:<br>❍ well lit (1)<br>❍ clean (2)<br>❍ stylish (3)<br>❍ in good condition (4) |

| | | | |
|---|---|---|---|
| FMG (Never exited "item research trials") | ***Trait Judgeability*** | Give several traits, identify which are most visible/judgeable. | fmg1 11. A group of friends are most likely to agree that someone they know:<br>❍ Is direct and honest (1)<br>❍ Gets defensive easily (2)<br>❍ Is a negative kind of person (3)<br>❍ Values being an intellectual (4) |
| MSC (Never exited "item research trials") | ***Misc. Hard TOPI Questions From Forms 1.0 and 1.1*** | | msc1 Which are most likely to go together:<br>❍ shame and desire for companionship (1)<br>❍ aggression and interest (2)<br>❍ contempt and aggression (3)<br>❍ guilt and a desire to play (4) |
| [a]The exploratory loadings are for the complete task; [b]the TOPI final are for the tasks after they were screened such that only "functional" items remained, i.e., > .35 loadings on their chief factor; < .25 on any secondary. [c]A version of the of the FME task that included several borderline items was re-introduced to augment the task's reliability. [d]The SGE items are indicated in brackets because they were not counted as items in the TOPI 5R proper but instead included as an experimental task-in-development. [e]Not including the augmented 5 from FME. | | | |

**buffer text contiguous with table**

**section break next page here:**

# Version Numbers of the TOPI and TOPI Factor Models

## Renaming the TOPI Versions 1.0 to 1.4 to the TOPI Versions 1 through 4R

The first TOPI scales were labeled the 1.0 through the 1.4, and often with variants, i.e., the 1.4R is revised version of the 1.4. To simplify our account of versions here, we will drop the leading identifier ("1.") relabeling version 1.4 as version 4 and, when forms are quite similar, refer to a slightly-revised (e.g., abridged) version with an "R" as a modifier, such that the version referred to in earlier research as the 1.4R will become the 4R. This will become helpful as we discuss new forms in the present article.

**buffer text contiguous with table**

Table 1.3

Version Control of the Test of Personal Intelligence

| Current Name | Original name | Comments |
|---|---|---|
| TOPI 0 | TOPI 1.0 | Initial tests of TOPI items |
| TOPI 1 | TOPI 1.1 | Expanded item-testing |
| TOPI 2 | TOPI 1.2 | Further test expansion with validity tests |
| TOPI 2R | TOPI 1.2 Reformatted | TOPI 2 Revised. Items were reordered and the order of answer choices counterbalanced to ensure an equal number of As Bs, Cs and Ds were correct |
| TOPI 4 | TOPI 1.4 | Streamlined (i.e., abridged and reformatted version of the TOPI 1.2Rf |
| TOPI 4R | TOPI 1.4R | IRT-revised version of the TOPI 1.4 with a focus on scoring for two mental ability factors |
| TOPI 5 | TOPI 1.5 | The TOPI 4 expanded to include 13 clusters with 205 items (compared to 93 for the TOPI 4). |
| TOPI 5R | TOPI 1.7 | Reduced, 145-item version of the TOPI 1.5 removing items that were (a) less-functional or (b) loaded on more than one factor |
| TOPI 5G | T14-RG47[a] | TOPI 5 General, A one-factor scale using items shared in common across TOPI forms 2 through 5R (1.2 through 1.7) |
| TOPI 5E | T15-G66[a] | TOPI 5E Another one-factor scale but this one using items newly introduced in the TOPI 5 so as to better assess higher levels of personal intelligence |

a. Also appears as the name of the one-factor model (see Table 1.4)

**buffer text contiguous with table**

## Versions of Two-Factor Models of the TOPI

During the course of our research with the TOPI, we also have named the factor models we employ of the test, particularly several two-factor models of the TOPI 4 and 5. The version control for those factor models are indicated in Table 1.4.

**buffer text contiguous with table**

Table 1.4

Version Control of the One- and Two-Factor Models of the Test of Personal Intelligence

| Current Name | Original name | Comments |
|---|---|---|
| Two-factor model of the TOPI 4 | Two-factor model of the TOPI 1.4 | A two-factor model of the full Test of Personal Intelligence, Version 4, trimming the original 93-item test to 67 items |
| Model MDT | Model MDT | A task-based two-factor model of the "Most Distinct Tasks" across samples; this, sadly, also was challenging to interpret (see Chapter |
| TOPI 4 Model; Model 4 | Model T14-58 | The two-factor model of the TOPI 4 trimmed such that only 58 items that are compatible with the TOPI 5 are used. This 58-item model fit the original TOPI 4 data even better than the full 67 Two-Factor model of the TOPI 4. |
| TOPI 5 Model; Model 5 | Model T15-58 | The two-factor model of the TOPI 5 constructed from just the 58 items of the TOPI 5 that, as indicated later in this Technical Supplement, this two-factor model is nearly identical to the two-factor model of the TOPI 5 that employed all 205 test items of version 5. |

**buffer text contiguous with table**

## Overview of the Test Versions with Clusters and their Reliabilities

A summary of the forms, their clusters, and reliabilities can be seen in Table 1.5.

***[Text continues after table(s)].***

**buffer text contiguous with table**

Table 1.5

Spreadsheet of Item Clusters Across Test Versions, with Alpha Reliabilities (and parenthetical no. of items)

| Task Abbr. | ***Abbrev. Task Name*** | TOPI Version | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 2 | 2Rf | 4 | Item Res. Studies | | 5 | | 5R |
| | | | | | 2013 | 2014 | All | Final[a] | Initial |
| *N of Study* | | 384 | 1114 | 10318 | | | 961 | 961 | 548 |
| ***Recognizing Personality-Relevant Information*** | | | | | | | | | |
| RVA | ***Identifying Motives*** | .29 (12) | .37 (10) | .38 (10) | -- | -- | .71 (16) | .77(14) | .69 (14) |
| RVF/ RVB | ***Inner States*** | .14 (8) | .08 (7) | T-DPI | .12 (12) | .50 (22) | .73 (22) | .78 (15) | .70 (15) |
| RVC | | .29 (4) | .17 (4) | D-CS | -- | -- | -- | -- | -- |
| RVD | ***Inner Exp.-to-Beh.*** | .49 (4) | .33 (4) | .37 (4) | -- | -- | .76 (14) | .79 (12) | .82 (12) |
| RVE | ***RVE Room with a Cue*** | -- | -- | -- | .16 (11) | -- | -- | -- | -- |
| Faces | | .18 (18) | -.01 (12) | D-RL | -- | -- | -- | -- | -- |
| Spaces | | .09 (12) | .01 (7) | D-RL | -- | -- | -- | -- | -- |
| Pets | | .48 (12) | .09 (7) | D-RL | -- | -- | -- | -- | -- |
| FMA/FMB[b] | ***Trait Knowledge*** | .33 (5) / .46 (8) | .19 (5) /.42 (8) | .51 (13) | -- | -- | .73 (15) | .74 (13) | .80 (13) |
| FMC | ***Trait Knwldg., Abstract*** | .48 (6) | .39 (6) | D-CS | -- | -- | | -- | -- |
| SGA | ***Iden. Goal-Rel. Planning*** | .47 (7) | .38 (7) | .42 (7) | -- | -- | .80 (15) | .80 (14) | .73 (14) |
| FMD | ***Integrating Inform.*** | .67 (9) | .46 (9) | .58 (.60) | -- | -- | .77 (16) | .78 (12) | .77 (12) |
| FME | ***Discrepancies-Defense*** | -- | -- | -- | .53 (12) | .63 (15) | .75 (15) | .62 (5) | .68 (8) |
| FMF | ***Act Frequencies*** | -- | -- | -- | .47 (11) | .65 (16) | .71 (16) | .50 (3) | -- |
| FMG | ***Trait Judgeability*** | -- | -- | -- | .14 | -- | -- | -- | -- |
| ***Guiding Choices*** | | | | | | | | | |
| GCA | ***Trait Inferences*** | .54 (8) | .43 (8) | .46 (8) | -- | -- | .75 (15) | .77 (12) | .78 (12) |
| GCB | ***Observers' Trait Ascriptions*** | .56 (8) | .38 (8) | .46 (8) | -- | -- | .76 (15) | .78 (14) | .66 (14) |
| GCC | ***Motivating Memories*** | .73 (9) | .54 (9) | .58 (9) | -- | -- | .83 (16) | .85 (13) | .87 (13) |
| GCD | ***Self-Mdls. and Choices*** | .67 (7) | .53 (7) | .50 (7) | -- | -- | -- | -- | -- |
| ***Systematizing Goals*** | | | | | | | | | |
| SGA??? | | | | | | | | | |
| SGB / SGC[b] | ***Goal Evaluation*** | .56 (6)/.45 (6) | .45 (6)/.32 (6) | .60 (12) | -- | -- | .85 (16) | .85 (16) | .84 (16) |
| SGE/ SGF | ***Personality Change*** | -- | -- | -- | .16 (10) | .46 (11) | .43 (11) | .46 (3) | [.55 (6)][d] |
| ***Other*** | | | | | | | | | |
| MSC | ***Hard Items Dropped from TOPI 0 and 1.1*** | -- | -- | -- | -.09 | -- | -- | -- | -- |
| TOPI Total | | | .84 | | -- | -- | .97 (205) | | .96 (140) |

T-DPI: Temporarily dropped pending item research; D-RL: Dropped owing to low reliability; D-CS: Dropped to conserve space, usually because area was oversampled.

[a]the TOPI final 1.5 are for the tasks after they were screened such that only "functional" items remained, i.e., > .35 loadings on their chief factor; < .25 on any secondary.

[b]The tasks were merged between the TOPI 2Rf and 1.4 owing to their reasonable similarity and similar behavior in early factor analyses

[c]A version of the of the FME task that included several borderline items was re-introduced to augment the task's reliability.

[d]The SGE alpha and items are indicated in brackets because they were not counted as items in the TOPI 5R proper but instead included as an experimental task-in-development.

[e]Not including the augmented 5 from FME.

**buffer text contiguous with table**

# Chapter 2. Development and Construction of the TOPI 5

## First Step: Modifications of the TOPI 4

The TOPI 4 was composed of 11 clusters and 93 items, but in the course of moving forward to the TOPI 5, we revised the 1.4, combining two pairs of clusters that appeared closely related to one another to form a 9 cluster test with little loss of breadth of content. : the first pair consisted of clusters that asked about which personality traits typically go together (FMA and FMB, merged into FMA), the second pair both asked test-takers to evaluate whether personal goals were potentially problematic (SGB and SGC, merged into SGB). That dropped the scale to 9 clusters.

## Second Step: Composition of New Tasks and Their Study in Two Pilot Studies

Concurrently over the years 2013 and 2014, we conducted two studies (*Ns* = 446 and 381) to trial seven new item clusters as possible candidates for a revised TOPI, and to revive an earlier cluster we had discarded owing to an insufficient number of items remaining.

### Overview of the Pilot Studies

Beginning in the summer of 2013, we developed a series of additional problem-solving tasks beyond those found in the TOPI 2 and the TOPI 4R (a subset of the TOPI 2 with more functional items).

### TOPI Item Research 2013 (Pilot Study 1)

The seven tasks upon which we began work included two tasks that "looked back" to TOPI versions 0 through 2: (a) The RVF "Inner States" task expanded on a promising TOPI 2 task that had been dropped en route to the 4 because it had lacked a sufficient number of well-performing items. (b) The second MSC "Miscellaneous Hard Questions" included items on early versions of the TOPI (versions 0 and 1) that fewer than 10%-15% of our participants could answer correctly and had therefore been dropped on the original scale, and that we now hoped to take a further look at.

The five new tasks included, (c) RVE "Room with a cue", (d) FME "Discrepancies-defense mechanisms, (e) FMF "Act frequencies", (f) FMG "Trait judgeability", and (g) SGE "Personality change". These seven tasks were piloted in a study called "TOPI Item Research-2013." The earlier table, Table 1.5 (in the last chapter), provides an overview of the personal intelligence tasks that we had been working with up to that time, including the additional problem-solving tasks we tested in the pilot studies.

**Participants.** Five-hundred seventy-eight individuals logged into our survey; we screened out 132 who had substantially incomplete data or completed the survey overly quickly, leaving 446 respondents. Based on these participants, we calculated the scale reliabilities as indicated in the middle columns of Table 1.5.

**Results.** Three of the six tasks were problematic: Although correct answers for each were keyed to published research (e.g., Gosling, Ko, Mannarelli, & Morris, 2002). FMG (Trait Judgeability), contained no items that loaded above $r = .10$ with the overall TOPI and an overall

reliability of α = .14; RVE contained just 3 items correlated with the overall TOPI above $r > .10$, and overall exhibited a reliability of α = .16. MSC, the miscellaneous hard items reinstated from the TOPI 0 through 2 had just three items that performed well, and we regarded those as not worth saving on their own.

The remaining four tasks, RVF, FME, FMF, and SGE, did show promise. Although Inner States (RVF) and Change (SGE) both had equally low reliabilities as the discarded scales, Inner States had  except inner states had 5 of 12 items loading above $r = .10$ on the overall test—some substantially so, and SGE similarly had 5 of 10 meeting criteria. These four were carried forward to TOPI Item research 2014.

**TOPI Item Research 2014 (Pilot Study 2)**

**Participants.** Four-hundred and seventy-two people logged onto the survey, of whom 381 met criteria of (a) taking longer than 20 minutes for the items and (b) having fewer than 3 missing values. We reworked each of the four promising tasks based on the item information from the first pilot, increasing the length of each task by from 1 item (for SGE) to 10 (for FMF; see Table 1.5).  Of the four tasks carried forward, RVF, FME, FMF, and SGE, RVF, FME, and FMF improved in reliability. SGE (personality change) dropped somewhat. Nonetheless, we decided to carry forward all four tasks RVF/RVB, FME, FMF and SGE to the new TOPI 5 without further changes to them. Specific cluster reliabilities were indicated in Chapter 1, Table 1.5.

## Third Step: Addition of New Tasks

From the pilot studies, four trial clusters seemed most promising and were prepared for the TOPI 5: These new tasks concerned people's problem-solving in areas of (a) evaluating inner states (RVF), personality change (SGE), identifying discrepancies in behavior, i.e., defense mechanisms (FME), and linking traits to acts (FMF).

# Chapter 3: The TOPI 5 in Study 1

## Sample

**Initial logins to the study.** Participants were tested under three conditions, spread over three academic years from 2014-2015 to 2016-2017. The first sample took the test in its original order online at a time and place of their choosing (N= 739 logins); the second sample (N=31) took the same test online while being proctored in a classroom. The third sample (N=540) took the test online, unproctored, in a "second-half-first" form, for which the first and second halves of the test were switched but otherwise the same. These samples were combined into an overall initial file with N = 1310 logins, which, after screening for reasonable levels of completion rates and of attention, resulted in a combined sample of $N$ = 961 participants.

**Screening in greater detail.** Screening for completeness of data and for attention. Of the 1310 who had logged onto the survey, 69 were non-respondents (completing just a few items) and 94 were partial respondents (completing less than 50% of the TOPI survey), and these were removed, leaving 1147 in the remaining sample.

An additional series of screens flagged participants for signs of extreme inattention: We flagged 11 participants who completed the survey in less than 2 sec per item, 182 who failed more than 50% of the attention items (a surprisingly high number of participants), 4 participants who exhibited longstring responding over more than 10% of the test (i.e., endorsed the same alternative more than 21 times in succession), and 21 who responded with a single answer invariantly more than 2/3rds of the time across the test. In all, we flagged 183 participants: 153 respondents were flagged once, 13 twice, and 20 thrice: In our final screening step, all those flagged were removed, yielding a final sample of $N$ = 961, upon which all further analyses were conducted.

## Procedure

Participants were recruited via SONA software for an online study in which they would "complete items related to personal intelligence." Further information provided defined what personal intelligence is, explained that the researchers were trying to evaluate test items related to the concept. We explained in advance that the survey would take between 1 to 2 hours, and encouraged participants to take a short break during filling it out. Participants received 2 hours of experimental credit toward their course requirement for completing the scale. If they signed up, they were then sent to a Qualtrics survey that provided the consent form, followed by the survey materials.

## Additional Details Regarding Selected Factor Analyses and Results

### Item retention in the factor model process

In our prior work with the TOPI 4 and 4R, most of the items we retained loaded on their primary factors between approximately $r$ = .25 and .90 (unstandardized). The larger number of items in version 5 argued for setting a modestly more stringent criterion, which we did: we retained only items that loaded $r$ = .35 or higher on their primary factor. We also screened out any items that exhibited loadings above $r$ = |.25| on other (non-primary) factors, because items

with high secondary loadings in the exploratory analyses on the TOPI 4/4R often violated the assumptions of the models we subsequently attempted to fit. (The second criterion was unchanged from our earlier development of the TOPI 4/1.4R).

**Exploratory Factor Analysis**

We began with Exploratory Factor analyses of the 205-item version in Mplus. We treated the data as categorical, and employed a Weighted Least Squares, Mean and Variance Adjusted (WLSMV) extraction, with a facparsim rotation—the latter member of the Crawford-Ferguson family of rotations particularly appropriate to large numbers of items (Finch, 2011; Sass & Schmitt, 2010). Table 3.1 lists the first five factor solutions and their fits. The 1-, 2-, 3-, and 5-factor solutions converged to criteria but the 4-factor solution did not, despite our allowing Mplus to employ up to 1.5 million iterations for each factor solution. (Our moderately fast laboratory laptop took 10 days, 16 hours and 17 minutes to complete the analyses for the 1 through the 5 factor solutions).

Given our criteria, *all* the tested models from the 1- to 5-factor model fit well, as can be seen from the fit statistics in Table 3.1. The 1-factor model led off with CFI and TLI both equal to .94, and an RMSEA of .015, with marked improvement for the 2-factor model (CFI, TLI = .984; RMSEA = .008), and continued but more gradual improvement through 5-factors (at which point we stopped our analysis).

**A Further Evaluation of the Models in Terms of Item-Indicators**

Aside from test structure, the factors also can be evaluated by how many items loaded on each factor—and according to whether certain tasks purely represented one factor or another. We examined these characteristics for the 1, 2, and 3 factor solutions. The 2-factor solution represented most of the items well, and, in that solution, items within tasks all loaded on either factor 1 or factor 2 but not on both. By comparison, the 3-factor solution's third factor was mostly defined by the FME task (coping and defense), and by 15 further items that arose across the remaining tasks; these additional items were difficult to interpret. Because items on the 5-factor solution exhibited mostly low loadings, that solution appears unlikely to support five reliable factor-based scales.

**Expanded Version of Exploratory Factor Analysis Table in the article**

Table 3.1 presents a version of Study 1 portion of Table 3 (Exploratory Factor Analyses at the Item level for Studies 1 and 2) that appeared in the paper with additional detail that we did not have room to fully discuss. These additional details include confirmatory factor analyses for 1-, 2-, and 3-factor models, including a variation in which several tasks were removed.

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 3.1

Factor Models of the TOPI 5 ($N$ = 961) at the Item Level

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | ***Item-Level Exploratory Factor Analyses*** | | | | | | | | |
| | *1- to 3-Factor Solutions Facparsim-Rotated, Oblique* | | | | | | | | |
| | Items used[a] | Item split of factors[b] | Dep. variables/ free parameters | Fit Indices | | | | | |
| | | | | Chi-2 | df | RMSEA | CFI | TLI | $r_{factors}$ |
| One factor model | 0 | – | 205/205 | 25135.28 | 20705 | .015 | .940 | .940 | na |
| Two factor model | 0 | – | 205/409 | 21673.45 | 20501 | .008 | .984 | .984 | $r$ = .40 |
| Three factor model | 0 | – | 205/612 | 21143.92 | 20298 | .007 | .989 | .988 | $r$ = .23 to .47 |
| Four factor model | 0 | – | NA | NA | NA | NA | NA | NA | NA |
| Five factor model | 0 | – | 205/1015 | 20477.52 | 19895 | .006 | .992 | .992 | $r$ = .15 to .51 |
| | ***Initial Item-Level Confirmatory One-, Two-, and Three-Factor Models*** | | | | | | | | |
| | *Items loading < .35 on their primary factor and/or > .25 on any other factor removed.* | | | | | | | | |
| One factor model | 162 | 162 | 162/324 | 17556.09 | 12879 | .019 | .938 | .937 | *na* |
| Two factor model | 146 | 74/72 | 146/293 | 11520.11 | 10438 | .010 | .983 | .982 | $r$ = .688 |
| Three factor model | 90 | 33/31/26 | 90/183 | 4607.46 | 3912 | .014 | .977 | .977 | *rs* = .60, .73, .74 |
| ***Item-Level Confirmatory for Two-Factor Model with (Three) Stub Tasks[c] Removed*** | | | | | | | | | |
| Two factor model | 135 | 69/66 | 135/271 | 9894.20 | 8909 | .011 | .983 | .982 | $r$ = .672 |
| ***Confirmatory Models for the Two-Factor Model with FME plus its Augmented Items and Stub Tasks[c] Removed*** | | | | | | | | | |
| Two factor model | 145 | 69/66 | 145/291 | 11615.62 | 10294 | .010 | .979 | .979 | $r$ = .698 |

[a]Out of an original total of 205 items to start

[b]Item splits for EFAs were indeterminate as some items loaded on more than one factor.

[c]We defined stub tasks as those with 5 or fewer remaining items

**buffer text contiguous with table**

## Further Exploring the 2 Factors of the TOPI 5

### Do the Factors Yield Reliable Scales?

We had hypothesized further that we could construct factor-based scales of the TOPI 5 that would exhibit good reliability. The reliability for the overall test was α = .97, and that for the two factor-based subscales was α = .94 for Factor 1 and α = .95 for Factor 2. Moreover, most of the individual tasks formed reliable composites as well, indicating the possibility of creating content-based scales. For the five tasks on each factor with more than 5 items, the reliabilities fell in the range of α = .74 to .85. For the remaining three tasks: FME, SGE and FMF, the reliabilities were lower.

### Correlation between the 2 Factors of the TOPI 4 and TOPI 5

We wondered if the two factors of the TOPI 5 that emerged would exhibit some backwards compatibility TOPI 4R in the form of high correlations across tests on comparable factor-based scores. We therefore next examined the correlation between the TOPI 5 factors with those of TOPI 4R, which we earlier had labeled Consistency-Congruency and Dynamic-Analytic factors (TOPI 4R Factors 1 and 2, respectively). Table 3.2 indicates the correlations between the

initial scales of the TOPI 5 and the final scales of the TOPI 4R. These correlations are based on 62 items (this was before we had eliminated three items for non-functionality, and discovered another item had been a "stowaway" that shouldn't have been carried forward…leading to the final 58) from the TOPI 5 that were carried over from the complete set of 67 TOPI 4R items, representing 92.5% of the earlier tests' items.

Importantly, the TOPI 5 overall correlated $r$ = .943 with the TOPI 4R. Also, the first factors of the two instruments and the second factors of the two instruments appeared to correlate somewhat more highly with one another, at $r$ = .84 and .90 than they did across factors, at $r$ = .77 and .77. that were included in the TOPI 5. Factors 1 and 2 of the earlier TOPI 4R had been labeled "Consistency-Congruency" and "Analytic-Dynamic" factors respectively.

**buffer text contiguous with table**

Table 3.2

TOPI 5 and TOPI 4R Overall and Factor-Based Scale Correlations

| Scales | TOPI 5 Scales | | | TOPI 4R (Abbr.) Scales | | |
|---|---|---|---|---|---|---|
| | TOPI 5 Tot | Factor 1 | Factor 2 | TOPI 4R Tot | Consistency | Dynamic |
| TOPI 5 Total | 1.0 | | | | | |
| Factor 1 | .875 | 1.0 | | | | |
| Factor 2 | .896 | .632 | 1.0 | | | |
| TOPI 4 Total (Abbr) | .943 | .853 | .891 | 1.0 | | |
| Consistency-Congr. | .864 | .844 | .774 | .940 | 1.0 | |
| Dynamic-Analytic | .916 | .772 | .904 | .961 | .785 | 1.0 |

**buffer text contiguous with table**

## Developing a two-factor TOPI 5 model

Working from the two-factor exploratory factor analysis for the TOPI 5, we re-divided the 58 items depending upon whether their highest loading was on Factor 1 or 2 in the present sample. The placements were relatively straightforward because 51 of the 58 items possessed loadings > .20 higher on one factor than the other. By coincidence, the TOPI 5 model had 27 items on Factor 1 and 31 on Factor 2, the same numerical split as the TOPI 4 model, but with some items changing factors.

### Testing the model on the present data

As can be seen in Table 3 (TOPI 5 Model, bottom right), the TOPI 5 model fit the TOPI 5 sample quite well with an RMSEA of .016 and CFI and TLIs of .977 and .976, respectively. The correlation between the two factors dropped to $r$ = .61, indicating their greater independence than before. This was a viable overall fit for a two-factor model of mental ability in the area.

### A fly in the ointment

However, we found the factors nearly impossible to interpret. We note retrospectively that the two-factor model of the TOPI 4 also had been challenging to interpret, but that we had gone forward and characterized the two factors in a way that appeared satisfactory at the time. The

two factors here, however, seemed to defy description, despite our collective experience at interpreting factors (see the Technical Supplement, Chapters 3 and 5).

## Issues of Factor Interpretability

We found the factors nearly impossible to interpret. We note retrospectively that the two-factor model of the TOPI 4 also had been challenging to interpret, but that we had gone forward and characterized the two factors in a way that appeared satisfactory at the time. The two factors here, however, seemed to defy description, despite our collective experience at interpreting factors. Although Factor 1 at first appeared to concern the understanding of goal states (a group of high-loading items came from the Goal-Related Actions and Identifying Motives tasks), other high-loading items concerned drawing conclusions about a person from diverse information (from the Integrating Information task). Collectively, we thought, this might bear some resemblance to our earlier, Dynamic-Analytic factor. The interpretation of Factor 2 was complicated by the issue that, similar to Factor 1, it also led off with goal-related questions from the Goal Evaluation task, and then continued with diverse questions about inner states, labeling traits, and recalling motivating memories from Inner-Experience-to-Behavior, Trait Knowledge, and Motivating Memories tasks—a diversity of content that, to us, defied characterization beyond concerning personal intelligence in general.

## Post-Hoc Search for Artifactual Causes due to Test Form Qualities in (Slightly) Greater Detail.

Given the difficulty we found earlier in interpreting the two-factor solution in the TOPI 5 factors, we next turned to the question of whether the two-factor solution could be due to some method artifact. To do so, we selected the top-10 items loading on factors 1 and 2, and compared them on a number of characteristics. These items were highly defining of their factors: Factor 1 items loaded on factors 1 and 2 $r_{mean}$ = .75 and .03, respectively; Factor 2 items reversed the pattern, $r_{mean}$ = .12 and .77. All our items were multiple choice, ruling out changes in the response scales as a possible explanation here (cf., Legree et al., 2014), but we examined several other possible artifacts.

### Item Difficulty

The overall difficulty levels for the two sets of 10 items were very similar at $M_{correct}$ = .87 and .86, $t$ = .52, n.s.. We conducted a parallel check of the 10 highest loaders of the TOPI 4, Archive A-Odd sample (from which the 2-Factor TOPI 4R model was developed), and which was administered in a fixed, uniform fashion (i.e., not counterbalanced). As with the TOPI 5, we identified the 10 highest-loading items on Factors 1 and 2 from the initial 2-factor EFA of the scale. There were again no striking differences at $M_{correct}$ = .85 and .87, $t$ = .51, n.s..

**Readability**

We entered the two sets of 10 items from the TOPI 5 into a readability calculator that scored the text simultaneously according to the Flesh-Kincaid Grade Level, Fog Scale measure, and four others (Scott, 2018). The average among six (default) estimates that provided grade level readability estimates was, for factor 1, 12.5, or midway through senior year high school, with a range of estimates from eighth grade to college graduate. The readability for the 10 items of Factor 2 was 11.2, or the beginning of junior year, with a range from the ninth grade to the end of 1st-year college. Given the wide variability and reasonable proximity of the mean difficulty, these seemed unlikely to be the basis of the different factors.

**Item Position**

We also examined item position. Recall that the TOPI 5 administration order was counterbalanced according to the first and second halves of the test: The first half ran: RVA, FMD, GCB, SGA, RVB, and FME; the second half, GCA, SGB, RVD, FMA, GCC, SGE and FMF. We first observed that the top 10 items of Factor 1 all came from the first half of the test, whereas the top 10 Factor 2 items all came from the second half—a clear difference. Second, we assessed the distance of each item from the beginning of its test-half, in units of survey screens (each survey screen displayed about 3 items). The Factor 1 items all fell within a 29-screen span at the beginning of their half (Screens 3 through 32) of 53 screens; the Factor 2 items all came within a 23-screen span (Screens 12 through 35) of 57 screens, roughly toward the middle of their half.

We examined the TOPI 4, Archive A-Odd sample for position effects there (TOPI 4 was presented in a fixed order). Nine of the 10 items of Factor 1 were on screens 1 through 11 of of the 31 screens of the test —roughly the first third (the 10th item was on screen 16). All 10 of the Factor 2 items extended between screen 11 through screen 28 of the 31 screens of the test—roughly the last two thirds. The finding is again suggestive that position could be a factor in the extraction process of the 2-factor model.

# Chapter 4. The TOPI 5R in Study 2

## Participants and Screening

Participants were students taking psychology courses at the same large New England pubic university from which the participants for Study 1 were drawn. They were tested on-line during the fall, 2017 semester and screened using procedures identical to Study 1.

We used the same data screening process as in Study 1: There were N= 686 initial logins. We first removed 20 non-respondents and 20 non-respondents, leaving $N = 633$. Still using procedures of Study1, we three participants for undue speededness, three for longstring responding on 10% or more of the test, and the remainder for failing to pass the attention-check items, yielding a final sample of $N = 548$, upon which all further analyses were conducted.

### Measures

***Demographic questions.*** Four demographic items asked about participants' gender, age, educational level and ethnicity.

## Developing the TOPI 5R from the TOPI 5

### The Test of Personal Intelligence 5R (a.k.a., TOPI 1.7 or 7)

The TOPI 1.7 represents an abridged version of the TOPI 1.5, containing 145 items of the original 205-itemsTOPI 1.5. We abridged the test by (a) removing 55 items that failed to load > .35 on either factor of the 2-factor model or that correlated > .25 on its secondary factor--the same criteria we used earlier to construct the TOPI 1.4R (Mayer et al., 2017), and (b) removing SGE/SGF and FMF after the item-level screen left so few items on them that they lacked reliability.

### Removing Tasks When Abridging the TOPI 5 to the TOPI 5R (a.k.a. TOPI 7)

To guide test development moving forward, we also further explored the three tasks that had five or fewer items remaining after our selection criteria: FME, FMF and SGE. FME, FMF and SGE had loaded 11, 11, and 4 items respectively above $r = .35$ on the 1-factor representation of personal intelligence—indicating the general level of the task's functioning. The personality change task, SGE, was clearly unworkable as it stood. (We did, however, note that the three items that *had* worked in the 2-factor solution and potentially could be supplemented by rewrites of three "next-best" items that were similar in content and had exhibited some promise in their loadings (although failing to meet our criteria).

The FME and FMF tasks were, by comparison, sound measures of personal intelligence as indicated in the 1-factor solution. The issue with FMF was that most its items loaded on both factors of the 2-factor solution, and therefore failed to contribute to the scale's measurement quality at the 2-factor level. (Specifically, 14 of FMF's 16 items loaded higher on Factor 2 than Factor 1, but the items' *comparative* loadings on both factors were within $r = .15$ of one another in half of those cases.

Finally, the discrepancies task, FME, loaded 9 items above .35 on Factor 1 on the 2-factor solution, but 4 of those 9 items were initially excluded because 3 items loaded above our

cut-off point of .25 on Factor 2 (3 items between $r$ = .25 and .30; and 1 item $r$ = .37 versus. $r$ = .53 on Factor 1). Here, we elected to relax our initial screen to allow those four additional items in a revised model including that "augmented FME;" that revised model continued to fit the data well overall, as indicated in the bottom row of Table 1.5.

It was on that basis that we removed these tasks moving forward to the TOPI 5R.

## Additional Measurement Scales

An additional four scales were used for various criterion tests. We judged, however, that establishing a robust and generalizable measure took priority over any further validation of our TOPI scales (which already had a good deal of evidence for their capacity to correlate and likely predict criteria). For that reason, we did not analyze these criterion scales in the present research, although we anticipate returning to the data at a later point in time in a subsequent article.

**Self-Estimated Personal Intelligence—16-items (SEPI-16).**

Participants completed the *Self-Estimated Personal Intelligence-16* (SEPI-16, Mayer et al., 2017), a 16-item scale that includes statements such as, "I read people's intentions well" and "I understand who I am" and answer indicating whether the item describes them on a 5-point response scale, from "1 strongly disagree" to "5 strongly agree".

***The Rosenberg Self-Esteem Scale (RSE).*** This 10-item scale includes questions such as "On the whole, I am satisfied with myself" and "I feel I do not have much to be proud of." (Rosenberg, 1965). Participants answered on a 5-point scale from "1 strongly disagree" to 5 "strongly agree". that departs from the original 4-point scale but is now the convention (e.g., Donnellan, Ackerman, & Brecheen, 2016).

***WordsumPlus (WsP).*** This 14-item vocabulary test was refined using item response theory and serves as a proxy for more general intelligence (Cor, Haertel, Krosnick, & Malhotra, 2012).

***HEXACO-60.*** This 60-item measure of the Big Six personality traits Honesty, Emotional Stability, Extraversion, Agreeableness, Consciousness and Openness; the scale contains items such as "I would be quite bored by a visit to an art gallery" (low Openness), which respondents answer on a 5-point scale from 1 "strongly disagree" to 5 "strongly agree" (Ashton & Lee, 2009).

### Item-Level Exploratory Factor Analysis

We began with Exploratory Factor analyses of the 145-item version in Mplus, again classifying the data as categorical, and applying a Weighted Least Squares, Mean and Variance Adjusted (WLSMV) extraction, with a facparsim rotation. Table 4.1 lists the first five factor solutions and their fits. All five models—from the 1- to the 5-factor model— fit well, as can be seen from the fit statistics in Table 4.1. The 1-factor model led off with CFI and TLI both equal to .94, and an RMSEA of .015, with marked improvement for the 2-factor model (CFI, TLI = .984; RMSEA = .008), and continued but more gradual improvement through 5-factors (at which point we stopped our analysis).

**buffer text contiguous with table**

Table 4.1

Exploratory and Confirmatory Factor Models of the TOPI 5R ($N$ = 548) at the Item Level

| | | | | ***Item-Level Exploratory Factor Analyses*** | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | *1- to 3-Factor Solutions Facparsim-Rotated, Oblique* | | | | | |
| | Items used | Item split of factors[a] | Dep. variables/ free parameters | Fit Indices | | | | | |
| | | | | Chi-2 | df | RMSEA | CFI | TLI | $r_{factors}$ |
| One factor model | 145 | – | 145/145 | 11759.49 | 10295 | .016 | .951 | .950 | na |
| Two factor model | 145 | – | 145/289 | 10579.86 | 10151 | .009 | .986 | .985 | $r$ = .403 |
| Three factor model | 145 | – | 145/432 | 10336.32 | 10008 | .008 | .989 | .988 | $r$ = .26, .37, .54 |
| Four factor model | 145 | – | 145/574 | 10147.44 | 9866 | .007 | .991 | .990 | $r$ = .24 to .52 |
| Five factor model | 145 | – | 145/715 | 9975.38 | 9725 | .007 | .992 | .991 | $r$ =.22 to .47 |
| ***Cross Validations of a General 1-Factor Model[b] and the TOPI 5 2-Factor Model (Model T15)[c]*** | | | | | | | | | |
| One factor model | 145 | 145 | 145/290 | 11759.49 | 10295 | .016 | .951 | .950 | *na* |
| Two factor model | 145 | 74/72 | 145/291 | 11157.46 | 10294 | .012 | .971 | .970 | $r$ = .79 |
| ***Cross-Check Tests of Personal Intelligence Models T14-58 and T15-58*** | | | | | | | | | |
| Model T14-58 | 58 | 27/31 | 58/117 | 2345.01 | 1594 | .029 | .904 | .901 | $r$ = .94 |
| Model T15-58 | 58 | 27/31 | 58/117 | 1839.63 | 1594 | .017 | .969 | .968 | $r$ = .62 |

=./738== [a]Item splits for EFAs were indeterminate as some items loaded on more than one factor.

[b]This model assigned all 145 items of the TOPI 5R to the same general factor.

[c]Model T15 employed the 2-factor Model T15 based on the earlier-conducted exploratory factor analysis of the TOPI 5.

**buffer text contiguous with table**

## Item-Level Confirmatory Factor Analysis

The confirmatory factor analysis at the item level generated similarly encouraging results. Both the general 1-factor model and 2-factor model from Model T15 fit the data well. The two factor model, for example, fit with model (CFI, TLI = .964; RMSEA = .016), nearly as good as on the earlier TOPI 5 data set. One matter that gave us pause, however, was that the correlation between factors rose to $r$ = .81. The T15-58 model exhibited a somewhat better correlation between the two factors at $r$ = .62, indicating the persistent independence from one another of those two abbreviate factors. Model T14-58 fit this new sample less well, with a fit below even the 1-factor model, and exhibiting a correlation between factors of $r$ = .94.

# Chapter 5: Study 1 and 2 Analyses of the TOPI 5 and 5R at the Task Level

## Exploratory Factor Analyses of the Tasks Paralleled Those of the Items

We further wondered whether analyses of the TOPI 5 at the level of the 13 tasks would be consistent with analyses we had conducted of the 205 TOPI items. We regarded the 13 task scores as yielding continuous data and, of course, the 13 tasks were far fewer in number than the 205 items. For these reasons, we used the Mplus factor analysis defaults of continuous data, and Geomin rotation (this was different from our treatment of the test at the item level). Results from the fits of 1- through 5-factor fits in our exploratory analyses are indicated in the top rows of Table 5.1.

## TOPI 5 Analyses at the Task Level

### Dimensionality of the Individual Tasks of the TOPI 5

Before the TOPI 5, small sets of items were grouped in clusters; the clusters were brief and not expected to rise to acceptable levels of reliability as stand-alone indices. The TOPI 5 was constructed around tasks that were of sufficient length to attain levels of individual reliability.

The TOPI 5 was designed with the possibility in mind to be tractable for analysis at the level of its tasks—which we considered viable if the tasks individually reached levels of reliability around $r = .70$ or higher. We did accomplish this (see Table 1.5 in Chapter 1).

### Exploratory Factor Analyses at the Individual Task Level

We carried out exploratory factor analyses for each of the 13 tasks. A one-factor EFA fit each of the tasks individually at our criterion fit levels, although several items within each task often failed to load on the single factor (see Table 5.1). This suggested minimally that that each task was a reasonably pure index of the factor it assessed, and perhaps, that the overall TOPI might be unifactorial. Supporting our hypothesis as to task reliability, thirteen of the fourteen tasks exhibited coefficient alphas between $\alpha = .71$ to .85, with only SGE (personality change) below at $\alpha = .43$ (see Table 1.5 in Chapter 1).

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 5.1

Confirmatory Factor Analyses for One-Factor Models of Each Individual Task of the TOPI 5

| *Task* | *Basics of the Analyses* | | | *Fit Statistics* | | | | |
|---|---|---|---|---|---|---|---|---|
| | Factors | Dep. Vars | Param | Chi-2 | df | RSMEA | CFI | TLI |
| | | | | | | | | |
| FMA | 1 | 18 | 18 | 227.13 | 135 | .027 | .969 | .965 |
| FMD | 1 | 16 | 16 | 182.27 | 104 | .028 | .977 | .974 |
| FME | 1 | 15 | 15 | 130.84 | 90 | .022 | .989 | .987 |
| FMF/FMB | 1 | 16 | 16 | 146.15 | 104 | .021 | .987 | .985 |
| GCA | 1 | 15 | 15 | 122.40 | 90 | .019 | .990 | .988 |
| GCB | 1 | 15 | 15 | 125.70 | 90 | .020 | .988 | .986 |
| GCC | 1 | 16 | 16 | 353.65 | 104 | .050 | .964 | .958 |
| RVA | 1 | 16 | 16 | 152.88 | 104 | .022 | .984 | .981 |
| RVB* | 1 | 22 | 22 | 285.27 | 209 | .019 | .965 | .962 |
| RVD | 1 | 14 | 14 | 280.77 | 77 | .052 | .941 | .930 |
| SGA | 1 | 15 | 15 | 121.89 | 90 | .019 | .993 | .992 |
| SGB | 1 | 16 | 16 | 133.77 | 104 | .017 | .996 | .996 |
| SGE-analy failed | 1 | 11 | | | | | | |

*Looks like an error on RVB with this run, from the data (one variable continuous, or variables not labeled as categorical

**buffer text contiguous with table**

## Test of a One-Factor Model at the Level of the 13 Tasks

A 1-factor exploratory factor analysis at the level of the 13 tasks TOPI 5 tasks failed to fit the data well (Table 5.2); however, the 2-factor confirmatory factor analysis using the same 13 tasks, did fit, with an RMSEA = .049; TLI = .986; CFI = .979, and a $r$ = .69 between Factor 1 and 2. The 3- and 4-factor solutions fit even better, but no task loaded above $r > .35$ on Factor 3 on either solution, or above $r > .45$ on Factor 4, so the latter two solutions were non-starters.

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 5.2

Fit Levels of Factor Models of the TOPI 5 ($N$ = 961) at the Task Level

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | ***Exploratory Factor Analyses of the TOPI 5 at the Task Level*** | | | | | | |
| | | *1- to 3-Factor Solutions Facparsim-Rotated, Oblique* | | | | | | |
| | Tasks used[a] | Item split of factors | Dep. variables/ free parameters | Fit Indices | | | | |
| ***Exploratory Models for Unaltered Tasks (All Items)*** | | | | | | | | |
| | | | | Chi-2 | df | RMSEA | CFI | TLI | $r_{factors}$ |
| One factor model | 13 | | 13/39 | 1619.52 | 65 | .158 | .820 | .784 | -- |
| Two factor model | 13 | | 13/51 | 173.81 | 53 | .049 | .986 | .979 | *r = .690* |
| Three factor model | 13 | | 13/62 | 86.63 | 42 | .033 | .995 | .990 | *rs = .65, .43* |
| Four factor model | 13 | | 13/72 | 48.85 | 32 | .023 | .998 | .995 | *rs = .14 to .72* |
| ***Confirmatory Models for Unaltered Tasks (All Items)*** | | | | | | | | |
| One factor model | 13 | | 13/39 | 1619.52 | 65 | .158 | .820 | .784 | -- |
| Two factor model | 13 | | 13/39 | 2293.91 | 65 | .189 | .742 | .690 | *r = -.97* |
| ***Exploratory Models "Task-Level-Distilled" Tasks[a, b, c] Based on Task Factor Analyses*** | | | | | | | | |
| One factor model | 14 | | 14/42 | 1681.46 | 77 | .147 | .820 | .787 | -- |
| Two factor model | 14 | | 14/55 | 230.14 | 64 | .052 | .981 | .973 | *r = .659* |
| ***Confirmatory Models "Task-Level-Distilled" Tasks[a, b, c] Based on Task Factor Analyses*** | | | | | | | | |
| One factor model | 14 | | 14/42 | 1681.46 | 77 | .147 | .820 | .787 | -- |
| Two factor model | 14 | | 14/43 | 419.29 | 76 | .069 | .961 | .954 | *r = .721* |
| ***Confirmatory Models for "Item-Level-Distilled" Tasks[d] Based on the Original TOPI 5 Item Analyses*** | | | | | | | | |
| One factor model | 12 | | 12/36 | 1792.00 | 54 | .18 | .785 | .738 | -- |
| Two factor model | 12 | | 12/37 | 259.83 | 53 | .06 | .974 | .968 | *r = .68* |

[a]Based on the factor analyses of the original 13 tasks.
[b]Both the 1- and 2-factor task-purified analyses raised warnings of non-definite positive first-order derivative product matrices, and consequently, raised doubts as to the reliability of these fit statistics.
[c]Based on individual task analyses, removing items that loaded < |.35| on their primary factor and/or > |.25| on any other factor.
[d]Task purification based on original item analyses, with items loading < .35 on their primary factor and/or > .25 on any other factor removed.

**buffer text contiguous with table**

### Confirmatory Factor Analysis at the Task Level

A 1-factor confirmatory factor analysis at the level of the 13 tasks TOPI 5 tasks failed to fit the data well; however, and—in distinction to the exploratory analysis—the 2-factor CFA also performed poorly (see Table 5.2, middle rows). confirmatory factor analysis using the same 13 tasks, failed to fit—with an RMSEA = .189; TLI = .742; CFI = .690. Imposing the additional constraints of simple structure (i.e., each task loading one factor) stretched the model too far (Factors 1 and 2, using all the items, correlated $r$ = -.97 (details in Table 5.2)—as if they all measured the same general personal intelligence factor.

## TOPI 5R Analyses at the Task-Level

**Results from the task-level analyses**

Whereas the item-level analyses of the TOPI 5R were more-or-less as anticipated—with reservations as to the correlation between factors—the task-level 2-factor analysis was a bit more concerning. Replicating the 2-factor solution using the unaltered task assignments to factors as with the TOPI 5, the fit appeared to suffer a bit, with the

**Exceptions to the rules**

Examining the tasks item-by-item, the items on tasks RVA, SGA, GCB, FMD, and RVB loaded fairly exclusively on Factor 1 as before (with the exception of RVB's last 3 items). Many of the tasks loading on Factor 2 again exhibited item-by-item loadings consistent with that: FMA, SGB (excluding the first 2 items), GCC, and RVD loaded fairly exclusively on Factor 2 as before. That said, the items from tasks GCA (of Factor 2) and FME (of Factor 1) were not so clear in the present data set: Although GCA loaded *more* on Factor 2 (as before), many of its items loaded near-equally on Factor 1, and the items of FME were so evenly split across the two factors that it was challenging to pick a predominant factor on which it loaded.

**buffer text contiguous with table**

Table 5.3

Factor Models of the TOPI 5R ($N$ = 548) at the Task Level

| ***Cross-Check of the 2-Factor Confirmatory Model T15 at the Task Level for the TOPI 5R*** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Two factor, all tsks | 11 | 6/5 | 11/34 | 278.23 | 43 | .100 | .939 | .922 | $r$ = .79 |
| Two factor, wo FME | 10 | 5/5 | 10/31 | 117.02 | 34 | .067 | .976 | .968 | $r$ = .74 |

[a]Item splits for EFAs were indeterminate as some items loaded on more than one factor.
[b]This model assigned all 145 items of the TOPI 5R to the same general factor.
[c]Model 5 employed the 2-factor Model 5 based on the earlier-conducted exploratory factor analysis of the TOPI 5.

**buffer text contiguous with table**

## Comparative Fit Levels for TOPI 5 and 5R Analyses

The comparative fit levels of the various solutions are shown in two different fashions (sequentially and side-by-side for comparative purposes) in Tables 5.4.

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 5.4

Fit Levels of the Exploratory and Confirmatory Factor Models of the TOPI 5 (*N* = 961) and TOPI 5R (*N* = 548) at the Task Level Arranged Side-by-Side

| Fit Statistics for Exploratory Factor Models | | | | | | | | | | Fit Statistics for Confirmatory Factor Models | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Dep. Vars | Param | Chi-2 | df | RSMEA | CFI | TLI | | | Dep. Vars | Param | Chi-2 | df | RSMEA | CFI | TLI |
| Study 1 | | | | | | | | | | Study 1 | | | | | | | |
| 1 | | 13 | 39 | 1619.52 | 65 | .158 | .820 | .784 | | 1 | 13 | 39 | 1619.52 | 65 | .158 | .820 | .784 |
| 2 | | 13 | 51 | 173.81 | 53 | .041 | .986 | .979 | | 2 | 13 | 39 | 2293.91 | 65 | .189 | .742 | .690 |
| $r_{12}$ | | | | | | | | .67 | | $r_{12}$ | | | | | | | -.97 |
| Study 2 | | | | | | | | | | Study 2 | | | | | | | |
| 1 | | 11 | 33 | 595.35 | 44 | .151 | .858 | .822 | | 1 | 11 | 32 | 2073.91 | 45 | .287 | .477 | .361 |
| 2 | | 11 | 43 | 99.56 | 34 | .06 | .983 | .973 | | 2 | 11 | 34 | 278.23 | 43 | .100 | .939 | .922 |
| $r_{12}$ | | | | | | | | .67 | | $r_{12}$ | | | | | | | .79 |

***Tech note:*** Low fit for the TOPI 5-Abr 1-factor model in Study 2 is because RVA doesn't fit the single factor well/mod index of 260.83.

**buffer text contiguous with table**

## (Non-) Interpretation of the Meaning of the Two Factors at the Task Level

A general idea of what the two-factor model looks like at the task level can be gathered from Tables 5.5 and 5.6 which includes both the two-factor task loadings for the exploratory and confirmatory models. The "Exploratory Model" includes the complete, initial tasks. The Confirmatory Model used the tasks from which the most function items had been retained, based on the item-level analyses.

The two-factor structure was by-and-large resistant to interpretation: For example, as Tables 5.5 and 5.6 indicate, RVA and GCB loaded most highly on the first factor (*r*s = .84 and .81). Reference to Table 1 explains these tasks measure inferring a motive from a set of behaviors and or pursuits (RVA) and identifying the traits that an observer ascribes to a target person, given how the observer behaves (GCB). By comparison, SGB and RVD loaded most highly on the second factor (*r*s = .96 and .89); consulting Table 1, these tasks measure the ability to identify a goal or goals that may conflict with a person's aims (SGB), and to identify an inner experience that may accompany carrying out an activity. Examining further high-loading tasks failed to clarify the picture for us. Moreover, when subjected to the constraints of a simple structure confirmatory factor analysis, the two factors exhibited a correlation of *r* = -.97, suggesting that there might be one factor to the TOPI despite the better 2-factor fits. (But please note that these analyses were on "untampered-with" tasks, i.e. no items deleted, no assumptions relaxed, etc.). We continued to track the task-level analyses in Study 2, but the results failed to become clearer, and because the task-level analyses have the drawback of being less comparable to the item-level analyses we have used in the past, we focus on the item level and do not further report the tasks here).

**buffer text contiguous with table**

Table 5.5

Task Loadings on the Exploratory and Confirmatory Factor Two-Factor Models of the TOPI 5 and 5R

*Note: FME, which the table indicates changed factors between Studies 1 and 2, actually did so; this is not a reporting error.*

| | | | Exploratory Factor Analyses | | | | | | | | Confirmatory Factor Analysis--Simple Structure | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 1-Factor Solution | | 2-Factor Solution | | | | | | 1-Factor Solution | | 2-Factor Solution | | | | | |
| | items | coef. α | Study 1 | Study 2 | Study 1 | | Study 2 | | | | Study 1 | Study 2 | Study 1 | | Study 2 | | | |
| | | | **I** | **I** | **I** | **II** | **II[a]** | **I[a]** | | | **I** | **I** | **I** | **II** | **I** | **II** | | |
| FMA | 15 | .73 | .74 | .73 | .11 | .70 | .02 | .74 | | | .74 | .97 | | .98 | | .76 | | |
| FMD | 16 | .77 | .70 | .60 | .79 | .02 | .73 | .00 | | | .70 | .95 | -.97 | | .70 | | | |
| FME | 15 | .75 | .73 | .76 | .50 | .17 | .21 | .59 | | | .73 | .98 | -.95 | | .70 | | | |
| FMF | 16 | .71 | .75 | -- | .25 | .56 | -- | -- | | | .75 | -- | | .97 | -- | -- | | |
| GCA | 15 | .75 | .74 | .82 | -.01 | .76 | .16 | .70 | | | .74 | .98 | | .98 | | .81 | | |
| GCB | 15 | .76 | .77 | .62 | .81 | .07 | .75 | .00 | | | .77 | .95 | -.98 | | .72 | | | |
| GCC | 16 | .83 | .77 | .77 | .05 | .76 | -.09 | .88 | | | .77 | .98 | | .98 | | .81 | | |
| RVA | 16 | .71 | .62 | .55 | .84 | -.01 | .82 | -.13 | | | .62 | .97 | -.94 | | .68 | | | |
| RVF/ RVB | 22 | .73 | .78 | .77 | .76 | -.01 | .62 | .26 | | | .78 | .98 | -.98 | | .83 | | | |
| RVD | 14 | .76 | .79 | .83 | -.00 | .89 | -.03 | .90 | | | .79 | .99 | | .99 | | .87 | | |
| SGA | 15 | .80 | .76 | .71 | .79 | .01 | .75 | .09 | | | .76 | .97 | -.98 | | .80 | | | |
| SGB | 16 | .85 | .77 | .84 | -.13 | .96 | .07 | .81 | | | .77 | .99 | | .99 | | .86 | | |
| SGE/ SGF | 11 | .43 | .59 | -- | .14 | .45 | -- | -- | | | .59 | -- | -- | -- | | -- | | |

a. Factors I and II exchanged positions across studies (which is not unusual and reflects their similarity in accounting for variance).

**buffer text contiguous with table**

***[Tables continue on next page]***

**buffer text contiguous with table**

Table 5.6

Individual Task Loadings in a Form Easier to Check for Interpretability

| | | Items | | Factor Loadings | | | | |
|---|---|---|---|---|---|---|---|---|
| Task | | Init.[a] | Screen.[b] | Exploratory | | | Confirmatory | |
| *Tasks Loading Predominantly on Factor 1* | | | | | | | | |
| | | | | F1 | F2 | | F1 | F2 |
| RVA | Find the common motive among three or so behaviors or pursuits | 16 | 14 | **.83** | -.14 | | **.74** | -- |
| SGA | Subsidiary Goal-Related Actions | 18 | 14 | **.84** | .01 | | **.86** | -- |
| GCB | Given a person's decision to act with another person in a particular way (that could entail some anticipation of their reaction), identify a trait they believe the other person has. | 15 | 14 | **.85** | .06 | | **.86** | -- |
| FMD | Integrate several personality-relevant pieces of information about a person to conclude something about a person's knowledge, intellect, or beliefs | 16 | 12 | **.83** | -.05 | | **.81** | -- |
| RVB | Given a situation, role, or activity a person is engaged in, infer their inner state | 22 | 15 | **.79** | .07 | | **.83** | -- |
| FME | Infer something about a person's defense and coping from a discrepancy between their words and behavior. | 15 | 5 | **.54** | .25 | | **.62** | -- |
| *Tasks Loading Predominantly on Factor 2* | | | | | | | | |
| FMA | Given a person's two traits, identify a third likely trait | 18 | 13 | .12 | **.68** | | -- | **.79** |
| GCA | Given someone's traits, predict how they are likely to react or behave | 15 | 12 | .00 | **.80** | | -- | **.80** |
| SGB | Give a person's objective (e.g., to make friends), identify a goal around that objective that likely will create conflicts for a person because it is unrealistic, hard to fulfil, or contradicts the objective. | 16 | 16 | -.12 | **.97** | | -- | **.87** |
| GCC | Given a person's intentional calling forth of a specific personal, motivational memory, identify the reason or the goal the person called it forth. | 16 | 13 | .06 | **.77** | | -- | **.86** |
| RVD | A person is carrying out a common activity; identify an inner experience that likely accompanies that activity. | 14 | 12 | -.00 | **.87** | | -- | **.87** |
| FMF | Identify what behaviors are associated with a given trait. | 16 | 3 | .26 | **.55** | | -- | -- |
| SGE | Given a person's intentions and behaviors, judge how ready they are to change | 11 | 3 | .48 | **.58** | | -- | **.66** |
| | | | | | | | | |

[a]The exploratory loadings are for the complete task.

[b]The confirmatory loadings are for the tasks after they were screened such that only "functional" items remained.

**buffer text contiguous with table**

# A Partial Working Through of a "Most Distinct Tasks" Approach to the 2-Factor Solution

## The Most Distinct Tasks Approach

This raises an interesting question…in spite of flip-floppiness of the tasks, are there a few tasks that are stable markers of Factor 1 and Factor 2 (and, by implication, other tasks that more decisively measure both?), To find out, we engaged in an informal "historical" review of prior factor analytic solutions—of the 1.4, the 1.5 and the 1.7. Although the items of the TOPI 4 often split across factors, we went back with a more critical eye to see whether we could sort them into "mostly Factor 1" "mostly Factor 2" and "Both Factor" groups. The tasks for the TOPI 5 and 1.5-Abr (or 1.7) already were designated that way. The results from the TOPI 4, 5 and 5R are shown side-by-side in the next table.

Green highlights are those that "stayed on the same side of the fence" for the 4 and 5; yellow are those that switched sides. Black are those for which tasks were only in one sample or another (moving from 4 to 5). Note that two or three of the green selections included tasks that exhibited some split items: RVD and GCC in particular. Note also that in the transition from the 1.5 to the 1.7, FME, FMF, and GCA exhibited more splitting in the 1.7. For GCA, the level of splitting was more similar for the 1.4 and 1.7 than for the 1.5.

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 5.7 Revised and Expanded

Examination of Item Splits for Individual Tasks in the Mplus Exploratory Factor Analyses for the 2-Factor Models of the TOPI 4R, the 1.5, and the 1.5-Abr (a.k.a. 1.7)[a,b]

| Task Abbr. | ***Abbreviated Task Name*** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | TOPI 4R | TOPI 5 | TOPI 5R | TOPI 4R | TOPI 5 | TOPI 5R | TOPI 4R | TOPI 5 | TOPI 5R |
| | | F1 | F1 | F1 | F2 | F2 | F2 | Both | Both | Both |
| *N of Study* | | 5144 | 961 | 548 | 5144 | 961 | 548 | 5144 | 961 | 548 |
| ***Recognizing Personality-Relevant Information*** | | | | | | | | | | |
| RVA | ***Identifying Motives*** | 6/10 | 14/16 | 14/14 | | | | | | |
| RVF/ RVB (revised) | ***Inner States*** | | 15/22 | 11/15 | | | | | | |
| RVD | ***Inner Experience-to-Behavior*** | | | | *1/4* | 12/14 | 9/12 | *1/4* | | |
| ***Forming Models of Personality*** | | | | | | | | | | |
| FMA/ FMB (A, B merged) | ***Trait Knowledge*** | 1/5A; 3/8B | | | | 13/18 | 13/13 | *1/5A; 1/8B* | | |
| FMC | ***Trait Knowledge, Abstract*** | 4/6 | | | | | | | | |
| FMD | ***Integrating Information*** | | 12/16 | 11/12 | 7/8 | | | | | |
| FME | ***Discrepancies-Defense*** | | *4/15* | | | | | | 4/15 | 6/10 |
| FMF | ***Act Frequencies*** | | | | | *4/16* | | | 6/16 | |
| FMG | ***Trait Judgeability*** | -- | -- | -- | -- | -- | -- | -- | -- | -- |
| ***Guiding Choices*** | | | | | | | | | | |
| GCA | ***Trait Inferences*** | *2/8* | | | | 12/15 | | 3/8 | | 6/12 |
| GCB | ***Observers' Trait Ascriptions*** | 2/8 | 14/15 | 13/14 | *1/8* | | | *1/8* | | |
| GCC | ***Motivating Memories*** | | | | *1/9* | 13/16 | 13/13 | *6/9* | | |
| GCD | ***Self-Models and Choices*** | -- | -- | -- | 5/7 | -- | -- | -- | -- | -- |
| ***Systematizing Goals*** | | | | | | | | | | |
| SGA | ***Goal-Related Actions*** | 5/7 | 14/15 | 12/14 | | | | | | |
| SGB / SGC (merged) | ***Goal Evaluation*** | | | | 3/6B; 3/6C | 16/16 | 15/ 16 | | | |
| SGE/ SGF (revised) | ***Personality Change*** | -- | | | -- | 3/11 | ??? | -- | | |

*Notes:* a. Fractions for factors 1 and 2 indicate how many items exceeded .35 on their assigned factor and were below .25 on the alternative factor. Fractions for the "both" column indicate how many items were above .25 on *both* factors.

b. The visual perception tasks Faces, Spaces and Pets, as well as RVC, Evidence about the Self, RVE, Room with a Cue, and GCD, Self-Models and Choices, were not included in versions 1.4R through 1.7 of the TOPI and do not appear in the table.

c. Item splits in italics represented alternative interpretation to the dominant (i.e., higher count) assignment, and are provided for informational purposes.

**buffer text contiguous with table**

## Stable Indicator Tasks Identified

Based on that review, we identified the tasks shown in Table 5.8 as reflecting mostly Group 1 or Group 2.

**buffer text contiguous with table**

Table 5.8

Most Distinct Tasks?

| Group | **Tasks** |
|---|---|
| Group 1 | **RVA SGA GCB FMD RVB [Maybe: FME]** |
| Group 2 | **FMA, SGB, GCC, RVD, FMF, SGE/F [Maybe: GCA]** |

**buffer text contiguous with table**

## Possible Interpretation of the "Most Distinct Tasks"

The possible interpretation of the two factors remained challenging, even examining just the most distinct tasks. Table 5.9 presents representative items within the maker tasks.

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 5.9

Distinctly Factor A Tasks-- FACTOR 1: RVB, RVA, GCB SGA and Distinctly FACTOR 2 TASKS: RVD, GCC?, AND SGB/SGC

| Task Code | ***Task Name*** | Brief Task Description | ***Abbreviated*** and |
|---|---|---|---|
| **Distinctly Factor 1 Tasks: RVB, RVA, GCB SGA** | | | |
| RVA | ***Identifying Motives.*** | Find the common motive among several behaviors and/or pursuits | rva1 1. If a person wants to be around people to talk to them and to have a good time, the person is likely going to:<br>❍ be in love (1)<br>❍ express warmth toward someone (2)<br>❍ meet a goal of excellence (3)<br>❍ socialize (4) |
| RVF (with RVB) | ***Inner States*** | Infer a person's inner state from information about their situation, activity, or role in which they are engaged. | rvb1 1. A minute of time will pass most quickly for:<br>❍ A cook who is watching the clock regularly, carefully timing how long to boil asparagus (1)<br>❍ A student in a classroom who is carefully following the teacher's directions to combine chemicals that will fizz when they are mixed (2)<br>❍ A musician who is in a business meeting in which he feels uninvolved and which his agent will need to explain to him later (3)<br>❍ A bus driver who is transporting people on his regular route toward the end of the day (4) |
| GCB | ***Observers' Trait Ascriptions.*** | Identify the trait that an observer ascribes to another person, given the observer's plans or behaviors around the person. | gcb1 1. A college student returned to his room and noticed a scratch on his desk he never had seen before. He immediately suspected his roommate. The student's reaction makes sense if his roommate is...<br>❍ rigid (1)<br>❍ careless (2)<br>❍ deceitful (3)<br>❍ studious (4) |
| SGA | ***Goal-Related Actions*** | **.** Identify the intermediate or subsidiary goal, attitude or behavior that could satisfy a longer-term goal. | sga1 1. A person wants "to perform at work with excellence". What goal might most promote this?<br>❍ to take a training course to learn to do the job better (1)<br>❍ to be a good leader to others (2)<br>❍ to use forceful, strong actions so as to become a good leader (3)<br>❍ to try to be a good friend (4) |
| **Distinctly Factor 2 Tasks: RVD, GCC?, AND SGB/SGC** | | | |
| RVD | ***Inner Experience-to-Behavior*** | A person is carrying out a common activity; identify an inner experience that likely accompanies that activity. | rvd1 1. When a person puts his/her best foot forward he or she often:<br>❍ views him or herself as better than before (1)<br>❍ feels worried about being "found out" as a fraud (2)<br>❍ feels ashamed of his or herself (3)<br>❍ comes to resent the effort (4) |
| GCC | ***Motivating Memories*** | Given a person's motivational need, identify the personal memory that will enhance the individual's motivation. | gcc1 1. When younger, Sam remembered being cut from his baseball team and the humiliation he felt, and how he wondered if he had practiced enough. Sam used this memory to help himself:<br>❍ work harder to achieve a goal (1)<br>❍ recall that self-doubt just isn't helpful (2)<br>❍ perform well in a job interview (3)<br>❍ cope with the challenges of shopping for sports equipment (4) |
| SGB (with SGC) | ***Goal Evaluation.*** | Given a person's objective (e.g., to make friends), identify a goal around that objective that likely will create conflicts for a person because it is unrealistic, hard to fulfil, or contradicts the objective. | sgb1 1. A person wants to make friends. Which goal might cause him problems when he pursues new friendships?<br>❍ be a good friend to his friends (1)<br>❍ to be all things to all people (2)<br>❍ to be myself (3)<br>❍ to spend time meeting new people (4) |

[a]The exploratory loadings are for the complete task; [b]the TOPI final are for the tasks after they were screened such that only "functional" items remained, i.e., > .35 loadings on their chief factor; < .25 on any secondary. [c]A version of the of the FME task that included several

borderline items was re-introduced to augment the task's reliability. [d]The SGE items are indicated in brackets because they were not counted as items in the TOPI 5R proper but instead included as an experimental task-in-development. [e]Not including the augmented 5 from FME.

**buffer text contiguous with table**

One possible interpretation of those two factors were:

- Factor 1: Analytic-Dynamic PI, consisting of an understanding of motives, beliefs about people, and
- Factor II Inferring and Analyzing Inner Dynamics text at end of table, plus section break:

That said, the interpretation did not strike us as altogether compelling.

## First Test of the "Most Distinct Tasks" Model

We nonetheless, however, would built a 2-factor model that specified the tasks RVB, RVA, GCB and SGA for factor 1, and RVD, GCC, and SGB/SGC for factor 2. Of this selection, GCC is a bit questionable in that it had many blends in the TOPI 4, but perhaps that was because the exploratory rotation was a bit different and less definitive in that earlier work (where the correlations between the two factors were higher). To ensure we had a minimum of 3 indicator tasks for each factor, however, it would need to be included. We then tested it on the TOPI 5-Abr Sample (Study 2). The fit of even this "purified" model for the Study 2/TOPI 5-Abr [a.k.a. 1.7] sample did not meet criteria for the RMSEA, although it did fit well in other regards.

**buffer text contiguous with table**

Table 5.10

**Cross-Check of the 2-Factor Most Distinct Tasks (MDT) Model for the TOPI 5-Abr**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Two factor, MDT | 7 | 4/3 | 7/22 | 50.59 | 13 | .073 | .983 | .972 | *r* = .72 |

**buffer text contiguous with table**

## Rationale for Breaking off Further Analyses

In the end, we broke off this line of investigation for several reasons. Most importantly, the two factors remained challenging to interpret. Beyond that, however, it didn't fit on first try, and the post-hoc nature of the analyses made us concerned we could be capitalizing on chance. Finally, we were increasingly wondering whether a 1-factor model might make more sense.

# Chapter 6. Study 3 "Analyses of Clarification" and the Application of Models 4 and 5 Across the Six Data Sets

In Study 3, we hoped to conduct preliminary analyses that could clarify what caused the difference in fit across samples between the two-factor Model 4 developed from the TOPI 4R and two-factor Model 5 developed from the TOPI 5.

## Hypotheses

We tested two hypotheses to account for the difference. The first hypothesis (a) stated that the difference in Models 4 and 5 was due to differences between college students enrolled in military academies and ROTC, on the one hand, and college students who were civilians, on the other. The suggestion that college students would differ in their structure of intellect depending upon their military-civilian status, however, seemed implausible.

Our second hypothesis (b) was that the difference between Models 4 and 5 might be due to a difference between the 4$^{th}$ and 5$^{th}$ generations of the TOPI in which the 58 common items were embedded—versions 4/4R, on the one hand, and 5/5R, on the other. The hypothesis that the *very same* 58 test items would exhibit different factor structures due to the not-so-different tests in which they appeared also seemed implausible.

## Research Approach

To examine these two hypotheses, we fit the two-factor TOPI 4 and TOPI 5 models to six sets of data: the Study 1 and 2 samples collected here, and four additional sets of data, Archives A, B, C, and D. Fortuitously, these six data sets varied as to their military and civilian composition, as well as to whether the data came from the TOPI 4 or 5 generations of the test, and these variations allowed us to address the questions at hand.

## Method of Analyses

In Study 3, we assembled six data sets we had available to us to conduct tests of the hypotheses, including the two samples collected here and four test archives.

### Test Archives

The four data archives contained predominantly college-age students working toward their baccalaureates, with some from military academies and ROTC, and others with civilian status. Archive A was composed (95%) of college students enrolled in military academies and ROTC; Archive B, of military academy and ROTC students as well ($N$ = 8459). Archives A and B were described in Mayer et al. (2017), where they were referred as the Original and Replication samples, respectively. We constructed Archive C ($N$ = 4922) for this study from data sets from the same military population of test respondents, whose responses were sent to us between July

2015 and June 2018 from the *Office of Economic and Manpower Analysis* of the United States Army (OEMA) for scoring. These data were scored shortly after they were sent, but were otherwise left unanalyzed by our research group until this project.

Archive D was from a predominantly civilian college student group ($N$ = 1072) who participated in graduate-student research projects and whose de-identified data was available to us. The sample was drawn from a large Northeastern campus. A few non-civilian students also may have been among the group as the campus of approximately 12,000 students hosted "over 100" ROTC members at the time (UNH Admissions, 2016).

The average age across the groups varied from 18.5 to 21.1. The three military archives A, B, and C had more men than women (about 3.5 to 1), but women predominated in the civilian samples (about 2.5 to 1). The estimated ethnic composition of the military and civilian samples were 78% and 91% White/Caucasian, respectively, with the remaining groups from Black/African, Asian, Hispanic/Latino groups next, and representations from other groups as well. Details are in Table 6.1.

**buffer text contiguous with table**

Table 6.1

Sample Characteristics of the Four Archives and Two Samples that Make up the Advances Article

| Archive and Source | *N* | Year Collected[b] | Gender Split | Avg. Age | Ethnic Breakdown[a] |
|---|---|---|---|---|---|
| A, mostly military | 5,174 | 2012 to June 15th, 2015 | 4027 M, 1144 F | 21.1 | Not supplied; estimated[d] at 78% White/Caucasian, 17% Black/African American, 4% Asian, 1.2% Multi-racial, |
| B, military | 8,459 | November 2015 to September 2016 | 6539 M, 1920 F | 20.1 | Not supplied; please see above |
| C, military | 4,922 | October 2016 to Early 2018 | 3746 M, 1176 F | 20.0 | Not supplied; please see above |
| D, civilian | 1,072 | 2016-2017 (Jayne's Learning About) N = 644, original; Bryan, 469 prescreened, 2018 | 280 M, 782 F, 10 Unspc. | 18.5 | Not supplied; please see below |
| TOPI 5 Sample | 961 | Fall 2014 to Spring 2017 for Standard Order; Spring 2016 to 2017 for second-half-first order | 288 M, 670 F, 3 Unspc. | 19.5 | (1) 3, (2) 34, (3) 11, (4) 27,(5) 5 (6) 2, (7) 867, (8) 12[c] |
| TOPI 5R Sample | 548 | Fall, 2017 Semester | 175 M, 371 F, 2 Unspc. | 19.0 | (1) 3, (2) 9, (3) 6, (4) 16, (7) 506, (8) 8[c] |

[a]Note that ethnic data was not available from direct survey information collected or provided by the army for several of the samples, in which case we refer to the ethnic breakdown of the populations from which they were drawn.

b.Dates for the military data were reported first in from Mayer, Panter & Caruso, 2017.

[c]1, Native American; 2, Asian; 3; Black/African American; 4, Hispanic/Latino; 5, Middle Eastern; 6, Hawaiian or Pacific Islander; 7, White/Caucasian; 8, Other.

[d]Estimated from U.S. Department of Defense (2013, p. 21, Figure 2.20)

**buffer text contiguous with table**

*Procedure.* The procedure was to test hypotheses *a* (military-civilian differences) and *b* (TOPI 4 and 5 differences) by fitting Models 4 and 5 across the six data sets. We also employed tests of factor invariance that yielded similar results to those reported here (see Technical Supplement, Chapter 6), but were less informative due to the widely different sample sizes of the data sets (Yoon & Lai, 2018).

## Analyses of Clarification: Results

### Did college-aged military and civilian test-takers exhibit different structures of mental ability? (Hypothesis a)

If there were a difference between military and civilian test takers in the structure of their mental abilities in personal intelligence, it ought to have shown up in a difference of fit of the two-factor TOPI 4 and 5 models in the military versus civilian test-takers. Table 3 indicates the fit of the two models to military (A, B, and C) and civilian (D) archives (first rows), all of whom took the TOPI 4 or 4R. The fit for the two-factor TOPI 4 model appear to the left, and those for the TOPI 5 model are to the right. The table clearly indicates that Model 4 fit well across all four sets of data in which participants took the TOPI 4 and that Model 5 fit poorly across them all: Model 5 yielded estimated correlations between the factors of $r$ = .96 to 97 indicating their near-equivalence. There was no evidence of differences in the structure of intellect between military college students (A through C) and civilian students (Archive D) related to personal intelligence.

### Was the model fit dependent upon whether the 58 identical items were embedded in the TOPI 4 or 5 test versions? (Hypothesis b)

All test takers in Archives A through D took TOPI 4$^{th}$ generation tests (Table 3, top rows); all test takers in Samples 1 and 2 took the TOPI 5$^{th}$ generation tests. There were clear differences in the two-factor Model 4 and Model 5 fits depending upon whether generation 4 or 5 of the test was examined—even though the 58 items tested were identical across forms.

The two-factor TOPI 4 model exhibited a good or excellent fit to the 58 items across all samples who took the TOPI 4$^{th}$ generation of tests—and the TOPI 5 model uniformly failed. Examining the TOPI 5$^{th}$ generation of tests, however, the situation reversed: The two-factor TOPI 5 model fit well whereas the TOPI 4 model failed—often estimating the correlation between the two factors at or above $r$ = .90, suggestive of a one-factor solution.

**buffer text contiguous with table**

**Table 6.2**

Confirmatory Fits of the Two-Factor TOPI Models 4 and 5 to the Archival and Study Data Sets

| | | TOPI 4 Model[a] | | | | | | TOPI 5 Model[a] | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Archive/ Sample | N | Dep. Vars./ Free Parm. | Chi-2 | df | RMSEA | CFI | TLI | $r_{factors}$ | Dep. Vars./ Free Parm. | Chi-2 | df | RMSEA | CFI | TLI | $r_{factors}$ |
| ***Military and Civilian Groups Taking the TOPI 4 and 4R*** | | | | | | | | | | | | | | | |
| **Military Groups** | | | | | | | | | | | | | | | |
| Arch. A-Evn[b] | 5,174 | 58/117 | 2982.35 | 1594 | .013 | .957 | .956 | .82 | 58/117 | 5673.19 | 1554 | .016 | .941 | .939 | .97 |
| Archive B | 8459 | 58/117 | 4467.86 | 1598 | .015 | .961 | .959 | .86 | 58/117 | 5463.74 | 1594 | .017 | .947 | .945 | .97 |
| Archive C | 4922 | 58/117 | 3390.61 | 1594 | .015 | .961 | .960 | .86 | 58/117 | 4118.64 | 1594 | .018 | .946 | .944 | .97 |
| **Civilian Groups** | | | | | | | | | | | | | | | |
| Archive D[c] | 1072 | 58/117 | 2115.79 | 1594 | .017 | .951 | .949 | .84 | 58/117 | 2240.21 | 1594 | .019 | .939 | .937 | .96 |
| ***Civilian Groups Taking the TOPI 5 and 5Abr*** | | | | | | | | | | | | | | | |
| Study 1 | 961 | 58/117 | 3636.06 | 1594 | .037 | .882 | .877 | .92 | 58/117 | 1992.54 | 1594 | .016 | .977 | .976 | .61 |
| Study 2 | 548 | 58/117 | 2345.01 | 1594 | .029 | .904 | .901 | .94 | 58/117 | 1839.63 | 1594 | .017 | .969 | .968 | .62 |

[a] Models 4 and 5 refer to the 2-factor models of the TOPI 4 and 5, respectively (there were no Models 1, 2 or 3 tested in these studies). Also of note: Both Model 4 and Model 5 had 27 and 31 items on their 1st and 2nd factors. The fact that they share the same number of items on each factor is coincidental; the specific items on each factor shifted between models. [b] Archive A contained about 5% civilian data. [c] The rate of ROTC among undergraduates at the university is approximately 0.83%.

**buffer text contiguous with table**

## Post-Hoc Search for Artifactual Causes due to Test Form Qualities

To explore whether incidental qualities of the test forms might have generated the two-factor solutions, we selected 10 marker items loading highly on Factor 1 and minimally on Factor 2 and did the same to identify 10 Factor 2 marker items. Comparisons of the two 10-item sets indicated that they were similar in their overall readability (Scott, 2018), with obtained grade-level estimates of $M_{grade}$ = 12.32 v. 11.27, respectively, $t$ = .91, n.s, and equivalent in difficulty level, with mean correct answers of $M_{correct}$ = .87 and .86, $t$ = .52, n.s..

The two item sets did differ, however, in the items that surrounded them. The ten Factor 1 marker items of the TOPI 5 all fell in the "first part" of the test (the first six tasks) within (online) screens 3 through 32 of the first part's 53 screens. The ten Factor 2 marker items all came from screens 12 through 35 of the second part's 57 screens (seven tasks). This reflected a likely context effect (rather than solely a position effect) given that the TOPI 5 administration was counterbalanced by first and second parts of the test.

In a cross-check for context effects using the TOPI 4, we found that nine of its 10 marker items of Factor 1 were on the first third of the test, whereas all 10 items of the Factor 2 extended through the last two thirds of the test (we used that portion of the Archive A sample on which the model originally was constructed).

## Analyses of Clarification: Discussion

### *Differences due to TOPI test forms*

In Study 3, we determined that the difference in model fit across data sets was due to the version of the TOPI in which our 58 common items were embedded. When the items were embedded in the TOPI 4th generation tests, the two-factor TOPI 4 model fit best. When the *same 58 items* were embedded in the TOPI 5th generation tests, the two-factor TOPI 5 model fit best. No effect was found for whether the college samples were chiefly military or civilian in status.

***Why do test versions affect the two factors?***

We found some evidence that the placement of the items within the broader test could be of issue: perhaps lead-in items create priming, practice, or other similar effects. Because we had neither intentionally nor systematically reordered items to examine this issue, however, the exact nature of the context effect is uncertain. (We revisit these issues in the General Discussion). The more pressing question at this point in our research program was *what to do next?* in terms of advancing the measure of personal intelligence.

**buffer text contiguous with table**

Table 6.3

Study 4: Model Fits of the Factor Analyses[a] of the One-Factor Models of the TOPI 5G and 5E Across Relevant Archives and Samples

| Archive and Source | *N* | Items/ Item Splits | Variables/ Free Parameters | Chi-2 | df | RMSEA | CFI | TLI |
|---|---|---|---|---|---|---|---|---|
| TOPI One-Factor Exploratory Factor Analysis for the 58 Common Items | | | | | | | | |
| *Archives of Military and Civilian Test-Takers on the TOPI 4 and 4R* | | | | | | | | |
| A[b], military | 5,174 | 58 | 58/116 | 3565.35 | 1595 | .015 | .939 | .937 |
| B, military | 8,459 | 58 | 58/116 | 5506.22 | 1595 | .017 | .946 | .944 |
| C, military | 4,922 | 58 | 58/116 | 4143 | 1595 | .018 | .945 | .943 |
| D, civilian | 1,072 | 58 | 58/116 | 2248.73 | 1595 | .020 | .938 | .936 |
| *Study 1 and 2 Samples Taking the TOPI 5 and 5R* | | | | | | | | |
| TOPI 5 Sample | 961 | 58 | 58/116 | 3687.62 | 1595 | .037 | .879 | .874 |
| TOPI 5R Sample | 548 | 58 | 58/116 | 2355.40 | 1595 | .029 | .903 | .900 |
| TOPI-5G Confirmatory Factor Analyses | | | | | | | | |
| *Archives of Military and Civilian Test-Takers Taking the TOPI 4 and 4R* | | | | | | | | |
| A[b], military | 5,174 | 47 | 47/94 | 2034.57 | 1034 | .014 | .962 | .960 |
| B, military | 8,459 | 47 | 47/94 | 3096.99 | 1034 | .015 | .965 | .963 |
| C, military | 4,922 | 47 | 47/94 | 2308.46 | 1034 | .016 | .965 | .964 |
| D, civilian | 1,072 | 47 | 47/94 | 1397.20 | 1034 | .018 | .960 | .958 |
| *Study 1 and 2 Samples Taking the TOPI 5 and 5R* | | | | | | | | |
| TOPI 15 Sample | 961 | 47 | 47/94 | 1828.27 | 1034 | .028 | .943 | .941 |
| TOPI 17 Sample | 548 | 47 | 47/94 | 1491.87 | 1034 | .028 | .939 | .937 |
| TOPI 5E Confirmatory Factor Analyses | | | | | | | | |
| TOPI 15 Sample | 961 | 66 | 66/132 | 4086.01 | 2079 | .032 | .940 | .938 |
| TOPI 17 Sample | 548 | 56 | 56/112 | 2085.00 | 1484 | .027 | .952 | .951 |

[a]Note that for 1-factor models only, EFA and CFA fit statistics are the same.

[b]The even half, a.k.a., cross-check sample of Archive A is reported here for purposes of direct comparison with the statistics reported in Study 1 of Mayer et al., 2017 (the odd-numbered participants' data was used for model construction).

**buffer text contiguous with table**

## Invariance Tests of the Six Data Sets

Post-hoc we also carried out some key invariance tests, although we were aware of a strong argument against doing so: Although by most standards our Study 1 and 2 samples were large, at *N* = 961 and 548 post-screening, even together (*N* = 1509), they were considerably outnumbered by the TOPI 4 test-takers (*N* = 24668); roughly speaking, the ratio was roughly 16:1. In such instances, good fits can be obtained even if invariance exists because the larger group outweighs the smaller so substantially (see Yoon & Lai, 2018).

Even so, in many instances, the invariance findings largely mirrored those of the sample-by-sample breakdowns shown in Tables 6.2 and 6.3, with the exception that we could "override" misfits by using the larger group to set parameter estimates. An example will illustrate our meaning.

## Selected Examples of Invariance Tests for the Two- and One-Factor Models 4 and 5

Table 6.4 shows the fit of a few key invariance tests on our data. For these analyses, the comparison was between samples who took the TOPI 4 generation and TOPI 5 generation tests.

**buffer text contiguous with table**

Table 6.4

Selected Invariance Tests: 1. Tests for Metric Invariance of the TOPI 4 Original 2-Factor Model across those who took the TOPI 4 generation and TOPI 5 generation tests.

| Archive and Source | | *N*[a] | Items/ Item Splits | Variables/ Free Parameters | Chi-2 | df | RMSEA | CFI | TLI | $r_{12}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| TOPI 4 Test-Takers (*N* = 24668) versus TOPI 5 Test-Takers (*N* = 1509) for the 58 Common Items | | | | | | | | | | |
| *Archives of Military and Civilian Test-Takers on the TOPI 4 and 4R* | | | | | | | | | | |
| Metric Inv. 1[b] | Overall | 26280 | 27/31 | 176 | 15939.41 | 3246 | .017 | .942 | .941 | -- |
| | Group 4 | 24771 | 27/31 | 176 | 7494.76 | -- | -- | -- | -- | $r_{12}$ =.832 |
| | Group 5 | 1509 | 27/31 | 176 | 8444.65 | -- | -- | -- | -- | $r_{12}$ = 1.07 |
| Metric Inv. 2[c] | Overall | 26280 | 27/31 | 175 | 16227.31 | 3247 | .017 | .940 | .939 | -- |
| | Group 4 | 24771 | 27/31 | 175 | 7397.41 | -- | -- | -- | -- | $r_{12}$ = .855 |
| | Group 5 | 1509 | 27/31 | 175 | 8829.90 | -- | -- | -- | -- | $r_{12}$ = .855 |

[a] TOPI 4 Test-Takers (*N* = 10,318 + 8,459 + 4920 + 1072 = 24668) versus TOPI 5 Test-Takers (*N* = 961 +548 = 1509) for the 58 Common Items

[b]The latent variable matrix for Metric Inv. Model 1 was not positive definite: The correlation between the two factors exceeded 1.0.

[c]In the Metric model 2, we allowed Mplus to estimate the correlation—but fixed the estimate across groups.

**buffer text contiguous with table**

The above findings are in keeping with the group-by-group analyses: The 2-factor TOPI 4 model appeared to fit okay, but the correlation between factors rose to unity ($r$ = 1.09) among the TOPI 5 test-takers.

The surprise occurred when we allowed Mplus to estimate the correlation—but fixed the correlation to equal one another across groups in a second Metric Invariance Model (Metric Inv. 2). When we did that, the estimated correlation between factors across groups converged at $r$ = .85. Note that the fit barely suffered at all—a sign (as we interpreted it) that the approximately 16:1 ratio of TOPI 4 test-takers simply overwhelmed the lack of fit of the model to the TOPI 5 test-takers.

## Metric Invariance of the One-Factor Model(s)

In the following selections, we present the Metric Invariance model findings. We note that in those cases where we went on to a more stringent scalar invariance test, the results were much the same.

**buffer text contiguous with table**

Table 6.5

Selected Invariance Tests: 1. Tests for Metric Invariance of the TOPI 4 Original 1-Factor Model across those who took the TOPI 4 generation and TOPI 5 generation tests, and the revised 1-Factor Model Using 47 Selected Items.

| Archive and Source | | $N^a$ | Items/ Item Splits | Variables/ Free Parameters | Chi-2 | df | RMSEA | CFI | TLI | $r_{12}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| TOPI 4 Test-Takers ($N$ = 24668) versus TOPI 5 Test-Takers ($N$ = 1509) for the 58 Common Items | | | | | | | | | | |
| *Archives of Military and Civilian Test-Takers on the TOPI 4 and 4R* | | | | | | | | | | |
| Common 58 | Overall | 26280 | 58 | 232 | 16774.73 | 3190 | .018 | .938 | .935 | -- |
| | Group 4 | 24771 | 58 | 232 | 11040.91 | -- | -- | -- | -- | -- |
| | Group 5 | 1509 | 58 | 232 | 5733.81 | -- | -- | -- | -- | -- |
| TOPI 5G | Overall | 26280 | 47 | 188 | 9106.04 | 2068 | .016 | .960 | .959 | -- |
| | Group 4 | 24771 | 47 | 188 | 6418.07 | -- | -- | -- | -- | -- |
| | Group 5 | 1509 | 47 | 188 | 2687.97 | -- | -- | -- | -- | -- |

[a] TOPI 4 Test-Takers ($N$ = 10,318 + 8,459 + 4819 + 1072 = 24668) versus TOPI 5 Test-Takers ($N$ = 961 +548 = 1509) for the 58 Common Items

[b]The latent variable matrix for Metric Inv. Model 1 was not positive definite: The correlation between the two factors exceeded 1.0. In the Metric model 2, we allowed Mplus to estimate the correlation—but fixed the estimate across groups.

**buffer text contiguous with table**

# Chapter 7. Study 3 Development of the TOPI 5—General Version (TOPI 5G, a.k.a. TOPI-RG47)

## Notes on the TOPI 5G (a.k.a., TOPI 1.4-RG47)

If a general one-factor model failed to fit all 58 common items, we could improve upon it by dropping some subset of items. An examination of the one-factor model indicated that, when it exhibited poor or marginal fit, the issue concerned a small number of item pairs which violated the models assumptions—as indicated by higher-than typical modification indices for the item pairs. (The modification indices, in turn, estimate the change in $\chi^2$ that result from freeing fixed parameters). We identified 11 items that either appeared in a pair with an especially high modification index or that appeared in problematic pairs repeatedly (Further details may be found in the Technical Supplement to this paper). The pairs from which they came often included items from the same task, and especially those adjacent or otherwise nearby in position, and the items we removed also were especially easy it turned out: Eight of the 11 items removed (73%) were among the easiest $1/5^{th}$ of items of the 58 items on the TOPI 5-58, with mean correct endorsements between 86% to 92%. When we examined the wording of a number of the problematic pairs of items, we noticed little apparent overlap in specific content between them, except in the case of one pair (sga1 and sga3), which both contained a distractor with the words "forceful" in them. The problematic item pairs' characteristics of item simplicity and proximity suggested that they have violated the model in their sensitivity to general (lack of) attention and similar format. We labeled the final 47-item version as the 1-factor model the Test of Personal Intelligence the Robust General Model with 47 items (TOPI-RG-47). It was referred to as robust in the hope that it would generalize reasonably well across archives and samples.

## The Removal Process for Problematic Pair Items

To create the TOPI Robust General Model, we first examined the 1-factor exploratory factor analysis results for the 58 common items in the Archive A-odd sample, the full TOPI 5 205 item solution, and TOPI 5 58-item solution. The modification indices for each one were copied into a file and we searched the highest loading MIs among item pairs for all three solutions. (Doing so capitalized on chance a bit across two data sets, but still left four entirely independent data sets as additional tests of fit: i.e., Archives B through D and the TOPI 5R sample). There was considerable consistency as to problematic items, with many arising across samples in the RVA, FMD, and SGA tasks, and somewhat fewer in GCC and FMA. To simplify matters, we focused on the 58 items in the TOPI 5 sample. The top 16 problematic pairs and their modification indices can be found in Table 7.1

**buffer text contiguous with table**

Table 7.1.

Problematic pairs of items on the TOPI 5, 58 item EFA solution

| | 58 Items, from TOPI 5 Sample | |
|---|---|---|
| Modification Index | Paired Issues | |
| | First Member | Second Member |
| 13 | rva1 | rva4 |
| 13 | rva6 | rva7 |
| 13 | fma5 | fma6 |
| 15 | rva7 | rva8 |
| 15 | sga3 | sga4 |
| 16 | sga4 | sga6 |
| 16 | fmd2 | fmd5 |
| 16 | fmd4 | fmd7 |
| 18 | rva5 | rva7 |
| 18 | rva6 | rva8 |
| 18 | sga1 | sga3 |
| 19 | rva4 | rva7 |
| 20 | rva1 | rva5 |
| 23 | rva4 | rva6 |
| 24 | gcc8 | gcc9 |
| 29 | rva4 | rva8 |
| 286 | | |

**buffer text contiguous with table**

Next, we arranged the items in a table as shown.

**buffer text contiguous with table**

Table 7.2

Problematic pairs of items on the TOPI 5, 58 item EFA solution

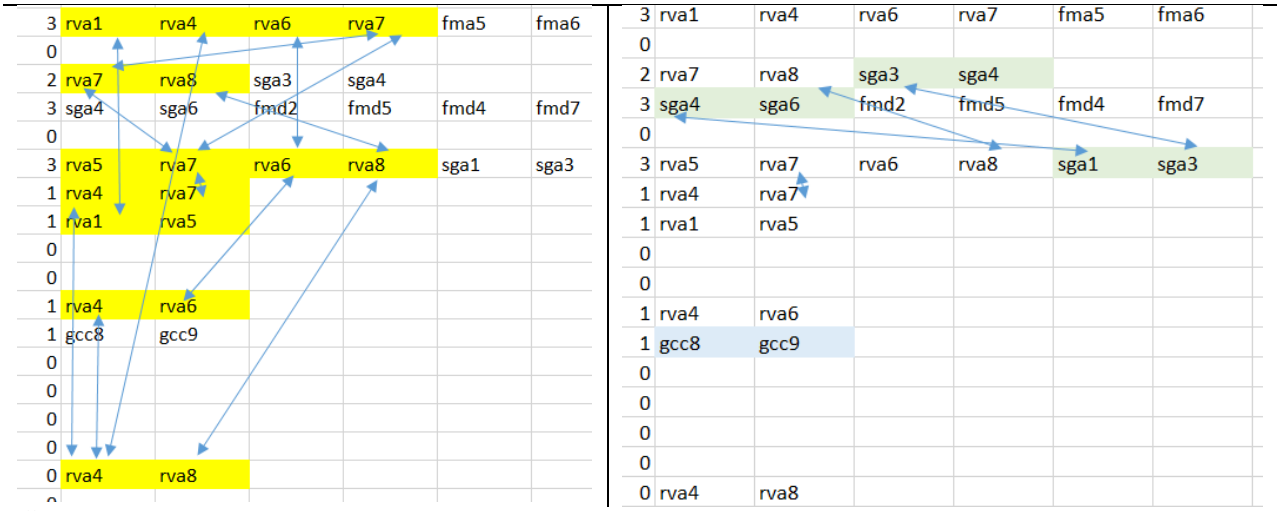
| Mod Index | # of pairs | Pair 1 | | Pair2 | | Pair3 | | Pair4 | |
|---|---|---|---|---|---|---|---|---|---|
| 12 | | | | | | | | | |
| 13 | 3 | rva1 | rva4 | rva6 | rva7 | fma5 | fma6 | | |
| 14 | 0 | | | | | | | | |
| 15 | 2 | rva7 | rva8 | sga3 | sga4 | | | | |
| 16 | 3 | sga4 | sga6 | fmd2 | fmd5 | fmd4 | fmd7 | | |
| 17 | 0 | | | | | | | | |
| 18 | 3 | rva5 | rva7 | rva6 | rva8 | sga1 | sga3 | | |
| 19 | 1 | rva4 | rva7 | | | | | | |
| 20 | 1 | rva1 | rva5 | | | | | | |
| 21 | 0 | | | | | | | | |
| 22 | 0 | | | | | | | | |
| 23 | 1 | rva4 | rva6 | | | | | | |
| 24 | 1 | gcc8 | gcc9 | | | | | | |
| 25 | 0 | | | | | | | | |
| 26 | 0 | | | | | | | | |
| 27 | 0 | | | | | | | | |
| 28 | 0 | | | | | | | | |
| 29 | 1 | rva4 | rva8 | | | | | | |
| 30 | 0 | | | | | | | | |
| 3?. | 0 | Note: We employed wildcard searches of data for MIs in the 30s and above by using the advanced MS Word option "use wildcards" and entering the strings below) | | | | | | | |
| 4?. | 0 | | | | | | | | |
| 5?. | 0 | | | | | | | | |
| 6?. | 0 | | | | | | | | |
| 7?. | 0 | | | | | | | | |
| 8?. | 0 | | | | | | | | |

**buffer text contiguous with table**

Using the above table, we diagrammed the pairs as indicated in Figure 7.1

**buffer text contiguous with table**

Figure 7.1. Diagram of Particularly Problematic Pairs of Items from the TOPI 5 Sample, 58-item EFA




**buffer text contiguous with table**

[At one point, the file location was]: file:///C:\Users\xxxxxxx\ACDM\RSR\Intell-PI\PI-TOPI\Artcl_T15-GenlPrecs\Phse3a-PwG-StpbyStp-Anlys-2018-06-28\STP9b-T-RG-48-ModIndx205&58Itms\ComparativeModIndices-58S1458S17Cross-Chk-2018-07-10-1401.xlsx

Based on the above, 11 items were removed (indicated in the codebook), but for the record, and using TOPI 5 nomenclature, by task, the removals were of:

fmd4r, fmd5r (2)
rva1r, rva4r, rva5r, rva6r, rva7r, rva8r, and (6)
sga1r, sga4r, sga6r (3)

It may be worth recording that 8 of the 11 items removed (73%) were among the easiest 1/5$^{th}$ of items on the TOPI 5-58 (11 items of 58), with mean correct endorsements between 86% to 92%.

## Why Did Problematic Pairs Arise?

This suggests that at least some of the pairs covaried more than expected because in addition to reflecting some real variance due to personal intelligence, they also were reflecting attention levels—i.e., they were so easy to get for most participants that missing them meant the test-taker's attention had lapsed. This argument becomes more compelling if one considers that many of the items among these problematic pairs were from the same task, and were often spaced adjacently or nearly so.

## Examining the TOPI 5G Through the Lens of Item Analysis

We further applied a 2-parameter IRT model to the item-level data for the RG47scale from both the Archive A-odd and TOPI 5 data sets (because these were very differently-behaving data sets). All 47 items exhibited *a* parameters (slopes) of .35 or higher in both data sets, indicating the items discriminated low from high scorers without exception. There were no marginal dependencies across items sufficiently problematic to consider removing any further items. The *b* parameters were nearly all in negative territory indicating that they were maximally sensitive to lower-scoring test-takers and less so for higher-scorers (largely unchanged from the TOPI 4R). The marginal reliability and the RMSEA of the fit for the RG47 in the Study 1 sample are indicated in Table 7.2. The TOPI 5G's marginal reliabilities were .78 and .86 in the two samples, and the RMSEAs were .03 and .01 (see Table 4). Overall, this indicated a good fit of the 2 parameter IRT model.

*Model fit of the TOPI-5G across samples.* A test of the 5G across samples bore out our hope that this form was robust across varying conditions, as can be seen in the middle rows of Table 5 labeled "TOPI-5G Confirmatory Factor Analysis": Not only did the fit of the model improve to a reasonable, albeit unexceptional, level for the TOPI 5 sample of .028, CFI=.943 and TLI = .941, and just marginally lower for the TOPI 5R, but the fit was excellent across all the TOPI 4/4R archives (see Table 7.2/Table 4 from manuscript). So, finally, we ended up with an acceptable common, robust model across all our samples and test forms.

**buffer text contiguous with table**

Table 7.2 (Table 4 from manuscript)

Study 4: Model Fits of the Factor Analyses[a] of the One-Factor Models of the TOPI 5G and 5E Across Relevant Archives and Samples

| Archive and Source | *N* | Items/ Item Splits | Variables/ Free Parameters | Chi-2 | df | RMSEA | CFI | TLI |
|---|---|---|---|---|---|---|---|---|
| One-Factor Exploratory Factor Analysis for the 58 Items Common to TOPI 4 and 5 Test Versions | | | | | | | | |
| *Archives of Military and Civilian Test-Takers on the TOPI 4 and 4R* | | | | | | | | |
| A[b], military | 5,174 | 58 | 58/116 | 3565.35 | 1595 | .015 | .939 | .937 |
| B, military | 8,459 | 58 | 58/116 | 5506.22 | 1595 | .017 | .946 | .944 |
| C, military | 4,922 | 58 | 58/116 | 4143 | 1595 | .018 | .945 | .943 |
| D, civilian | 1,072 | 58 | 58/116 | 2248.73 | 1595 | .020 | .938 | .936 |
| *Study 1 and 2 Samples Taking the TOPI 5 and 5R* | | | | | | | | |
| TOPI 5 Sample | 961 | 58 | 58/116 | 3687.62 | 1595 | .037 | .879 | .874 |
| TOPI 5R Sample | 548 | 58 | 58/116 | 2355.40 | 1595 | .029 | .903 | .900 |
| TOPI-5G Confirmatory Factor Analyses | | | | | | | | |
| *Archives of Military and Civilian Test-Takers Taking the TOPI 4 and 4R* | | | | | | | | |
| A[b], military | 5,174 | 47 | 47/94 | 2034.57 | 1034 | .014 | .962 | .960 |
| B, military | 8,459 | 47 | 47/94 | 3096.99 | 1034 | .015 | .965 | .963 |
| C, military | 4,922 | 47 | 47/94 | 2308.46 | 1034 | .016 | .965 | .964 |
| D, civilian | 1,072 | 47 | 47/94 | 1397.20 | 1034 | .018 | .960 | .958 |
| *Study 1 and 2 Samples Taking the TOPI 5 and 5R* | | | | | | | | |
| TOPI 15 Sample | 961 | 47 | 47/94 | 1828.27 | 1034 | .028 | .943 | .941 |
| TOPI 17 Sample | 548 | 47 | 47/94 | 1491.87 | 1034 | .028 | .939 | .937 |
| TOPI 5E Confirmatory Factor Analysis | | | | | | | | |
| TOPI 15 Sample | 961 | 66 | 66/132 | 4086.01 | 2079 | .032 | .940 | .938 |
| TOPI 17 Sample | 548 | 56 | 56/112 | 2085.00 | 1484 | .027 | .952 | .951 |

[a]Note that for 1-factor models only, EFA and CFA fit statistics are the same.

[b]The even half, a.k.a., cross-check sample of Archive A is reported here for purposes of direct comparison with the statistics reported in Study 1 of Mayer et al., 2017 (the odd-numbered participants' data was used for model construction).

**buffer text contiguous with table**

# Chapter 8. Study 3 Development of the TOPI 5—Extended Version (TOPI 5E)

In Study 4, in the section "Were the TOPI 5 items more challenging…", we present a t-test regarding the relative difficulty levels of items on the TOPI 4R and the new items of the TOPI 5, and showing that the TOPI 5 items were a bit more difficult than those on the 1.4R. Here are those analyses are described in a bit greater detail.

## Difficulty Level and Reduced Skew of the TOPI 5

We had hoped that the TOPI 5 items would be somewhat more difficult than those on the TOPI 4, as reflected in both a lower percentage correct on the test, and in a reduced skew of the scales. To undertake these analyses we asked two questions: (a) How did the 58 items that overlapped with the TOPI 4 compare with the TOPI 4 as a whole, and, then, (b) how did the 58 items compare to the TOPI 5 overall.

## Comparison of the TOPI 4R as a Whole to the 58 Common Items

Regarding the comparison of the 58 items to the TOPI 4R, they had a similar mean to the test as a whole. Using the TOPI 4 standardization sample of 10,318 participants, test-takers passed the the TOPI 4 items as a whole at the rate of 79.9% compared to a rate of 80.8% on the 58 common items; the respective means were $M_{14R}$ = .799 and $M_{58items}$= .807, for a small but significant difference of .008 ($t_{paired}(10317) = -37.9$, $p < .001$).

## Comparison of the TOPI 5 With the 58 Items

The present sample ($N$ = 961) performed a bit less well than the Archive A sample. The present sample answered the 58 common items correctly at a rate of 76.4% compared to the earlier sample's 80.7% success rate. Even given the difference, however, the TOPI 5 was more difficult than the 58 common items carried forward from the TOPI $M_{TOPI1.5(961)}$ = .734 versus the $M_{TOPI1.5\text{-}58items(961)}$= .764; this was statistically significant with a difference of .03 $t_{\text{paired}}(961) = 19.69$, $p < .000$. Because this difference was more than three times that of the common items relative to the overall TOPI 4R, it appears likely that the TOPI 5 is genuinely a bit more difficult overall. These results are reported in Table 8.1.

Table 8.1 indicates that the overall proportion correct on the TOPI 5 was 73% (.734 proportion correct). The comparable figure for the complete Archive A sample on the 58 common items of the TOPI 4 had been 80.7% (see Table 8.1 for details).

**buffer text contiguous with table**

Table 8.1

Mean Correct Scores on the 58 Items Shared Across Tests Compared to the Overall Mean Correct of the TOPI 4 and 1.5

| | | TOPI 4R Standardiz. Sample (*N* = 10,318) | | | TOPI 5 (Present Sample) (*N* = 961) | | |
|---|---|---|---|---|---|---|---|
| | | TOPI 4R, 58 items *checked* | TOPI 4R *checked* | Mean Difference; paired *t* *checked* | TOPI 5, 58 items *checked* | TOPI 5[b] *checked* | Mean Difference; paired *t* |
| Factor 1 | Mean | .839 | .831 | .008; *t* = 23.743*** | .803 | .728 | .075; *t* = 23.25*** |
| | Std. Dev. | .118 | .117 | -- | .172 | .187 | |
| | Skew | -1.58 | -1.58 | -- | -.141 | -1.13 | |
| | Std. Err | .024 | .024 | -- | .079 | .079 | |
| Factor 2 | Mean | .775 | .769 | .006; *t* = .27.8*** | .724 | .740 | -.016; *t* = -5.63*** |
| | Std. Dev. | .150 | .146 | -- | .191 | .204 | |
| | Skew | -1.32 | -1.27 | -- | -.912 | -1.14 | |
| | Std. Err | .024 | .024 | -- | .079 | .079 | |
| Overall | Mean | .807 | .800 | .008[a]; *t* = 37.88*** | .764 | .734 | .030; *t* = 19.69*** |
| | Std. Dev. | .120 | .119 | -- | .171 | .176 | |
| | Skew | -.154 | -1.52 | -- | -1.41 | -.998 | |
| | Std. Err | .024 | .024 | -- | .079 | .079 | |

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

[a]The .008 is rounded up from -.0075

[b]The overall TOPI statistics are presented for the 145-item version of the test (employing all items that met initial criteria; that is, loading > .35 on the initial factor and < .25 on any alternative factor; it includes the small numbers of items from tasks FME, FMF, and SGE).

**buffer text contiguous with table**

## Notes on the TOPI 5E in (Slightly) Greater Detail (a.k.a., TOPI 15-G66)

The new TOPI 5 items were more challenging than the TOPI 4R items, on the whole. To test this hypothesis, we divided the test into the 62 items that we carried forward from the TOPI 4R and 153 new items and compared their difficulty levels. The means, standard deviations, and skew of the item sets are recorded in Table 8.2. Tests of the TOPI 5 indicated that new items had a $M$ = .63, significantly harder than overlapping items with the TOPI 4R $M$ = .76, $t$ = 56.76, $p$ < .001.

***[Text and tables continue on next page]***

**buffer text contiguous with table**

Table 8.2

Studies 1 and 2. Mean Correct Scores on the 58 Items Shared by the TOPI 4R, 5, 5R in Comparison with Newly-Added Items

| | Study 1. TOPI 5 Sample (*N* = 961) | | | | | Study 2. TOPI 5R Sample (*N* = 548) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Skew | SE skew | | Mean | SD | Skew | SE Skew |
| 58 Overlapping Items | .760 | .171 | -1.13 | .079 | | .732 | .166 | -.845 | .104 |
| New items | .628 | .154 | -.695 | .079 | | .696 | .184 | -.734 | .104 |
| Difference | .132 | .072 | .435 | -- | | .036 | .072 | .102 | -- |
| *t-test of difference* | 56.76[a] | -- | 5.51[a] | -- | | 11.61[a] | -- | n.s. | -- |
| TOPI 5 Overall | .734 | .176 | -.998 | .079 | | .711 | .174 | -.781 | .104 |

a. $p < .001$

**buffer text contiguous with table**

To create the TOPI 5E, a one factor scale that would take advantage of the best items of the TOPI 5, we carried out the following steps. We first took the 170 items of the TOPI 5 that loaded greater than .35 on the first factor, and analyzed those in a 2-parameter IRT model so as to obtain estimates of each items' *a* and *b* parameters. The *a* parameter indicates the slope of the item curve at its steepest point, which reflects how well the item assesses its target variable of personal intelligence, or *theta* (theta is analogous to the classical test theory concept of true score). The b parameter reflects the level of ability at which the item most optimally assesses personal intelligence (i.e., theta).

Next, we divided items into six intervals based on the level at which they were best discriminating: that is, based on the value of their *b* parameters: (a) .1 and above, (b) -.01 to -.50, (c) -.51 to -1.00, (d) -1.51 to -2.00, (e) -2.01 to -2.50, and (f) -2.51 and below. We set a target to include 11 items in the highest range and 9 items thereafter, and then ranked the items within an interval according to their *a* parameters, but stopped inclusion of items from the same task after they reached five in number. This led to 56 variables. Finally, we supplemented tasks that had representation of fewer than five items, adding back in 10, resulting in a 66-item test we labeled the TOPI 5E (for "Extended" range of measurement).

The 5E ended up including 16 items from the 5G, and correlates with it $r(961) = .93$, as is indicated in Table 8.3. Adjusting for reliabilities of the two tests (.90 and .93, see Table 8.4), the estimated correlation between their true scores is 1.02 (the value above 1.0 likely an artifact of the shared items): The tests measure exactly the same attribute. That said, the G66 is slightly more difficult and exhibits markedly less skew than the RG47. We could further test an abridged version of the scale in the TOPI 5R sample (the shorter TOPI 5R omitted 10 items on the T15-66). This "5ER" scale performed similarly to the T15-66 in most regards, but we do not recommend its use as the abridged version lacks at least one key item that discriminates among high ability test-takers.

*Examining the TOPI-5G and 5E through the lens of item analysis.* The *a* and *b* parameters for the T15-66 likewise seemed good in the IRT model. The matrix was too sparse to estimate the RMSEA for the scale normally, but we employed the $M_2$ statistic to estimate RMSEAs of .05 for the TOPI 5 sample and .04 for the abridged test in the 1.7 sample.

**buffer text contiguous with table**

Table 8. 3

Characteristics of the TOPI 4 58 Items Common to Model 4 and 5, the TOPI 5G, and TOPI 5E and 5ER, in Studies 1 and 2

| | Items | M | S | Skew | S.E. Skew | Relia-bility | TOPI 4-58 correlations | | | Information from IRT (where tested) | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Alpha in SPSS | Total | Factor 1 CC | Factor 2 DA | Marginal reliab. | RMSEA |
| | | Study 1 Sample *N* = 961 (TOPI 5) | | | | | | | | | |
| | *TOPI Scales Employing Common Items Only* | | | | | | | | | | |
| TOPI 4 and 5 58 items | 58 | .763 | .171 | -1.15 | .079 | .92 | 1.00 | .94 | .95 | – | – |
| TOPI 5G | 47 | .738 | .181 | -1.01 | .079 | .90 | .98 | .88 | .97 | .86(.78[a]) | .03 (.01[c]) |
| | *New TOPI 5-based Scales* | | | | | | | | | | |
| TOPI 5E | 66 | .700 | .188 | -0.79 | .079 | .93 | .93 | .85 | .90 | .91 | .05[b] |
| TOPI 5ER | 56 | .725 | .188 | -.92 | .079 | .93 | .93 | .86 | .90 | .90 | .04[b] |
| | Study 2 Sample *N* = 548 (TOPI 5R) | | | | | | | | | | |
| | *TOPI Scales Employing Common Items Only* | | | | | | | | | | |
| TOPI 4 model items | 58 | .735 | .165 | -.86 | .104 | .90 | 1.00 | .93 | .95 | – | – |
| TOPI 5G | 47 | .706 | .179 | -.71 | .104 | .89 | .99 | .89 | .96 | – | – |
| | *New TOPI 1.5-Based Scale* | | | | | | | | | | |
| TOPI 5ER | 56 | .699 | .184 | -.65 | .104 | .92 | .94 | .86 | .90 | – | – |

a. Not analyzed

b. Matrix too sparse for computation of a $\chi^2$, so the RMSEA is based on an $M_2$ statistics using 1- and 2-way marginal tables.

c. Comparative values from Archive A-odd sample

**buffer text contiguous with table**

## Corrections for Attenuation due to Unreliability

**buffer text contiguous with table**

Table 8.4

By hand computations for correction for attenuation using Study 1 data

| Pair of tests | Test 1 reliability | Test 2 reliability | Test 1 & 2 corr | 1st term denominator | 2nd term denominator | denominator | Result |
|---|---|---|---|---|---|---|---|
| $r_{4w5G}$ | .92 | .90 | .98 | .959 | .949 | .91 | 1.08 |
| $r_{4w5E}$ | .92 | .93 | .93 | .959 | .964 | .924 | 1.01 |
| $r_{5Gw5E}$ | .90 | .93 | | .949 | .964 | .915 | |
| | | | | | | | |

*Note:* "4" refers to the 58-item version of the TOPI 4R; "5G" refers to the 47-item 5G form
"5E" refers to the 66-item 5E form

**buffer text contiguous with table**

## Data for the Test Characteristic Curves for the TOPI 5G and 5E for the TOPI 5 Sample *N* = 961 (Figure)

Table 8.5 below shows the data relevant to the following analysis. It includes Theta—the level of the participants' abilities, and next to that, estimates of test-takers' tau or true score on a T-Scale.

**buffer text contiguous with table**

Table 8.5

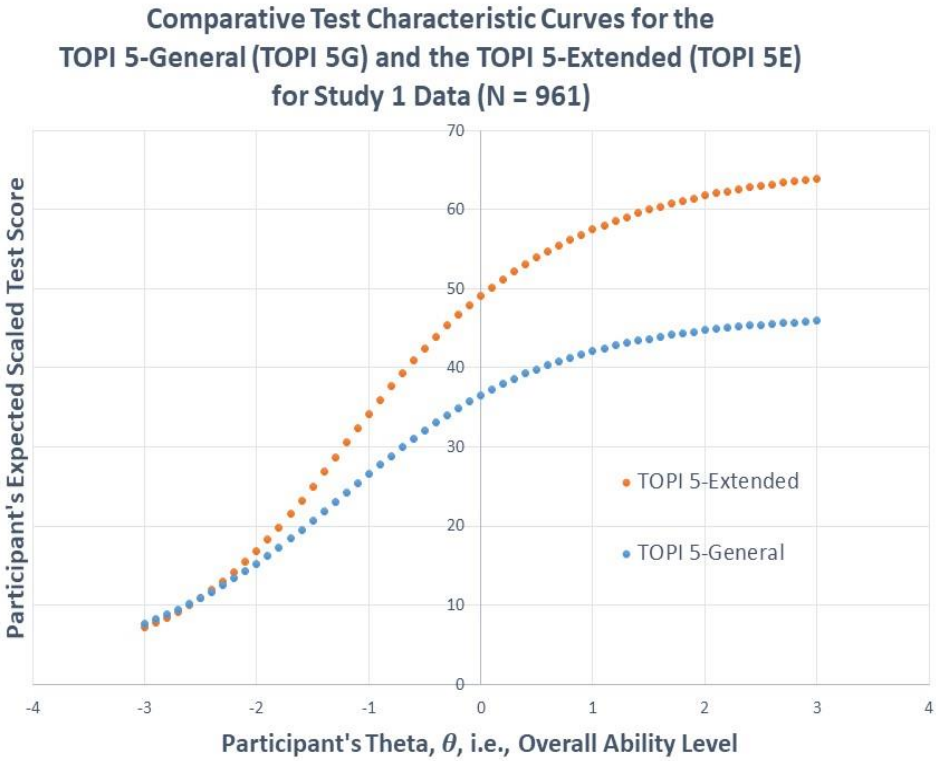
Comparative Values for the TOPI 5G and 5E Characteristic Curves

| **Theta** | **TOPI 5G, Total Inf.** | **TOPI 5E, Total Inf.** |
|---|---|---|
| -3 | 7.6515 | 7.1407 |
| -2.9 | 8.2113 | 7.7627 |
| -2.8 | 8.8111 | 8.4462 |
| -2.7 | 9.453 | 9.1971 |
| -2.6 | 10.1389 | 10.0216 |
| -2.5 | 10.8706 | 10.9256 |
| -2.4 | 11.6494 | 11.9151 |
| -2.3 | 12.4763 | 12.9952 |
| -2.2 | 13.3514 | 14.1703 |
| -2.1 | 14.2742 | 15.4435 |
| -2 | 15.2436 | 16.8157 |
| -1.9 | 16.2573 | 18.2857 |
| -1.8 | 17.3123 | 19.8492 |
| -1.7 | 18.4045 | 21.4985 |
| -1.6 | 19.5292 | 23.2224 |
| -1.5 | 20.6804 | 25.0066 |
| -1.4 | 21.8518 | 26.8338 |
| -1.3 | 23.0361 | 28.6855 |
| -1.2 | 24.2259 | 30.5428 |
| -1.1 | 25.4135 | 32.3878 |
| -1 | 26.591 | 34.2047 |
| -0.9 | 27.7511 | 35.9804 |
| -0.8 | 28.887 | 37.7045 |
| -0.7 | 29.9925 | 39.369 |
| -0.6 | 31.0623 | 40.9684 |
| -0.5 | 32.0921 | 42.4986 |
| -0.4 | 33.0783 | 43.9575 |
| -0.3 | 34.0185 | 45.344 |
| -0.2 | 34.9111 | 46.6578 |
| -0.1 | 35.7552 | 47.8999 |
| 0 | 36.5507 | 49.0715 |

| | | |
|---|---|---|
| 0.1 | 37.298 | 50.1746 |
| 0.2 | 37.9982 | 51.2116 |
| 0.3 | 38.6527 | 52.1851 |
| 0.4 | 39.2632 | 53.0981 |
| 0.5 | 39.8316 | 53.9535 |
| 0.6 | 40.36 | 54.7544 |
| 0.7 | 40.8507 | 55.5039 |
| 0.8 | 41.3059 | 56.2049 |
| 0.9 | 41.7277 | 56.8604 |
| 1 | 42.1184 | 57.4732 |
| 1.1 | 42.4801 | 58.0459 |
| 1.2 | 42.8148 | 58.581 |
| 1.3 | 43.1245 | 59.081 |
| 1.4 | 43.4109 | 59.5479 |
| 1.5 | 43.6758 | 59.984 |
| 1.6 | 43.9208 | 60.3913 |
| 1.7 | 44.1473 | 60.7714 |
| 1.8 | 44.3569 | 61.1262 |
| 1.9 | 44.5507 | 61.4574 |
| 2 | 44.7299 | 61.7663 |
| 2.1 | 44.8958 | 62.0544 |
| 2.2 | 45.0492 | 62.3231 |
| 2.3 | 45.1912 | 62.5737 |
| 2.4 | 45.3226 | 62.8072 |
| 2.5 | 45.4442 | 63.0248 |
| 2.6 | 45.5568 | 63.2276 |
| 2.7 | 45.6611 | 63.4166 |
| 2.8 | 45.7576 | 63.5926 |
| 2.9 | 45.847 | 63.7565 |
| 3 | 45.9298 | 63.9091 |

**buffer text contiguous with table**

Figure 8.1 uses the data from Table 8.5 to indicate the total test characteristic curves for the TOPI 5G and 5E on the Study 1 sample of $N = 961$. The Y-axis indicates the participants' true-score values on a T score scale ($M = 50$ and $S = 10$). The X-axis indicates the participants' ability level in standard deviations around the mean. The slopes of the lines indicate the tests' ability at distinguishing among test-takers. As can be seen, despite its overall reliability, the TOPI 5G "tops out" when measuring true scores of near (but still below) average. By comparison, The TOPI 5E exhibits effective measurement of true scores up to 1 standard deviation above the mean. That represents a substantial improvement, although it still falls short of a test that assesses ability further up the continuum.




Note that the average expected scaled score tops out for the TOPI 5-General at about a score of 47. That said, IRTPRO provides scoring intervals for the test up to 2.5 S above the mean—a T-scale score of 75 using the Study 1 data (N = 961). That is comparable to the TOPI 1.4R values for its original standardization sample.

The scaled scores on the TOPI 5E on the same Study 1 data (N = 961) rise to 2.5 standard deviations above the mean. In other words, IRTPRO appears to be generating estimated scores in both cases well above the expected scale scores for a given theta, and in addition, it recognizes that the TOPI 5E generates higher scores than the TOPI 5G, albeit there is no strictly linear relation between the two.

By comparison, the scoring went up to about 72.5 for the 1-factor model of the TOPI 1.4R:

```
if (t14Rtot=58) t14Rtot_t=54.4.
if (t14Rtot=59) t14Rtot_t=55.9.
if (t14Rtot=60) t14Rtot_t=57.6.
if (t14Rtot=61) t14Rtot_t=59.3.
if (t14Rtot=62) t14Rtot_t=61.1.
if (t14Rtot=63) t14Rtot_t=63.1.
if (t14Rtot=64) t14Rtot_t=65.2.
if (t14Rtot=65) t14Rtot_t=67.4.
if (t14Rtot=66) t14Rtot_t=69.9.
if (t14Rtot=67) t14Rtot_t=72.5.
```

IRTPRO generated a top score of about 71 for the 47 items of the 5G N=961 and of 76 for the top scorers of the 66 items of the 5E N =961, as well as about a 72.5 for the 1-factor model of the TOPI 1.4R on N = 5174. We are not sure how the algorithm "decides" when to stop scoring at the higher range, but these all are rather similar (and, assuming the algorithm takes precision into account, in a direction that suggests the precision of tests is TOPI 5G < TOPI14R < TOPI 5E, which is what we are seeing).

It appears that even though the scores go *up to* 70 or 75, the lower average expected test score for test-takers of high theta appears to reflect the idea that although some of those folks will indeed earn the highest possible scores, other among the high ability (theta) folks score lower, because of the imprecision of the test at the higher reaches, and that lowers the expected values (which are averages) for high-scorer on the less-precise test.

| | | | |
|---|---|---|---|
| 37 | 50.962 | 3.493 | 0.0503470 |
| 38 | 52.236 | 3.639 | 0.0535925 |
| 39 | 53.604 | 3.807 | 0.0564295 |
| 40 | 55.083 | 3.997 | 0.0584976 |
| 41 | 56.691 | 4.215 | 0.0592837 |
| 42 | 58.451 | 4.463 | 0.0581024 |
| 43 | 60.389 | 4.744 | 0.0541338 |
| 44 | 62.536 | 5.060 | 0.0466023 |
| 45 | 64.924 | 5.413 | 0.0352364 |
| 46 | 67.590 | 5.801 | 0.0211218 |
| 47 | 70.569 | 6.220 | 0.0076555 |

**Marginal reliability of the scaled scores for summed scores = 0.85830**

**Summary of the Data and Control Parameters** (Back to TOC)

| | |
|---|---|
| Sample Size | 961 |
| Number of Items | 47 |
| Number of Dimensions | 1 |

**buffer text contiguous with table**

| | | | |
|---|---|---|---|
| 49 | 50.264 | 2.736 | 0.0337650 |
| 50 | 51.117 | 2.820 | 0.0353483 |
| 51 | 52.011 | 2.912 | 0.0368278 |
| 52 | 52.950 | 3.011 | 0.0381412 |
| 53 | 53.938 | 3.119 | 0.0392103 |
| 54 | 54.982 | 3.237 | 0.0399399 |
| 55 | 56.088 | 3.365 | 0.0402173 |
| 56 | 57.264 | 3.506 | 0.0399147 |
| 57 | 58.517 | 3.659 | 0.0388953 |
| 58 | 59.858 | 3.828 | 0.0370258 |
| 59 | 61.297 | 4.012 | 0.0341974 |
| 60 | 62.849 | 4.215 | 0.0303581 |
| 61 | 64.528 | 4.439 | 0.0255570 |
| 62 | 66.352 | 4.685 | 0.0199983 |
| 63 | 68.342 | 4.954 | 0.0140921 |
| 64 | 70.518 | 5.249 | 0.0084699 |
| 65 | 72.904 | 5.569 | 0.0039032 |
| 66 | 75.520 | 5.909 | 0.0010466 |

Marginal reliability of the scaled scores for summed scores = 0.90703

**buffer text contiguous with table**

## Data for the TOPI 5E and 5G for the Test Information Curves and Standard Errors

**buffer text contiguous with table**

Table 8.6

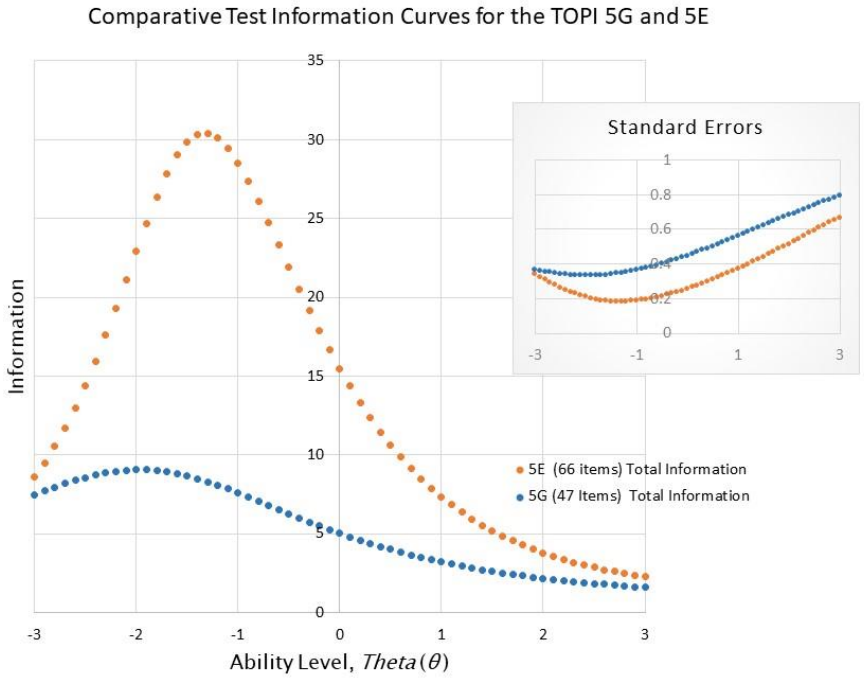
Data for the TOPI 5E-66 and 5G-47 for the Test Information Curves and Standard Errors

| Theta | Group 1, Total Inf. | Standard Error | Group 1, Total Inf. | Standard Error |
|---|---|---|---|---|
| | TOPI 5G | | TOPI 5E | |
| -3 | 7.4689 | 0.3659 | 8.5792 | 0.3414 |
| -2.9 | 7.7125 | 0.3601 | 9.5101 | 0.3243 |
| -2.8 | 7.9473 | 0.3547 | 10.5526 | 0.3078 |
| -2.7 | 8.1698 | 0.3499 | 11.7142 | 0.2922 |
| -2.6 | 8.3762 | 0.3455 | 12.9998 | 0.2774 |
| -2.5 | 8.5626 | 0.3417 | 14.4108 | 0.2634 |
| -2.4 | 8.7249 | 0.3385 | 15.9431 | 0.2504 |
| -2.3 | 8.8593 | 0.336 | 17.586 | 0.2385 |
| -2.2 | 8.9625 | 0.334 | 19.3195 | 0.2275 |
| -2.1 | 9.0315 | 0.3328 | 21.113 | 0.2176 |
| -2 | 9.0642 | 0.3322 | 22.9231 | 0.2089 |
| -1.9 | 9.0594 | 0.3322 | 24.6926 | 0.2012 |
| -1.8 | 9.017 | 0.333 | 26.3517 | 0.1948 |
| -1.7 | 8.9377 | 0.3345 | 27.8214 | 0.1896 |
| -1.6 | 8.8234 | 0.3367 | 29.0214 | 0.1856 |
| -1.5 | 8.6766 | 0.3395 | 29.8812 | 0.1829 |

| | | | | |
|---|---|---|---|---|
| -1.4 | 8.5007 | 0.343 | 30.3527 | 0.1815 |
| -1.3 | 8.2995 | 0.3471 | 30.4194 | 0.1813 |
| -1.2 | 8.077 | 0.3519 | 30.0992 | 0.1823 |
| -1.1 | 7.8375 | 0.3572 | 29.4388 | 0.1843 |
| -1 | 7.5851 | 0.3631 | 28.503 | 0.1873 |
| -0.9 | 7.3236 | 0.3695 | 27.3621 | 0.1912 |
| -0.8 | 7.0567 | 0.3764 | 26.0828 | 0.1958 |
| -0.7 | 6.7875 | 0.3838 | 24.7213 | 0.2011 |
| -0.6 | 6.5188 | 0.3917 | 23.3222 | 0.2071 |
| -0.5 | 6.2529 | 0.3999 | 21.9189 | 0.2136 |
| -0.4 | 5.9916 | 0.4085 | 20.5357 | 0.2207 |
| -0.3 | 5.7366 | 0.4175 | 19.1896 | 0.2283 |
| -0.2 | 5.4888 | 0.4268 | 17.893 | 0.2364 |
| -0.1 | 5.2492 | 0.4365 | 16.6542 | 0.245 |
| 0 | 5.0184 | 0.4464 | 15.4788 | 0.2542 |
| 0.1 | 4.7967 | 0.4566 | 14.3705 | 0.2638 |
| 0.2 | 4.5844 | 0.467 | 13.331 | 0.2739 |
| 0.3 | 4.3814 | 0.4777 | 12.3605 | 0.2844 |
| 0.4 | 4.1877 | 0.4887 | 11.4583 | 0.2954 |
| 0.5 | 4.0033 | 0.4998 | 10.6224 | 0.3068 |
| 0.6 | 3.828 | 0.5111 | 9.8502 | 0.3186 |
| 0.7 | 3.6614 | 0.5226 | 9.1385 | 0.3308 |
| 0.8 | 3.5034 | 0.5343 | 8.4839 | 0.3433 |
| 0.9 | 3.3537 | 0.5461 | 7.8825 | 0.3562 |
| 1 | 3.2119 | 0.558 | 7.3304 | 0.3693 |
| 1.1 | 3.0778 | 0.57 | 6.8241 | 0.3828 |
| 1.2 | 2.951 | 0.5821 | 6.3597 | 0.3965 |
| 1.3 | 2.8313 | 0.5943 | 5.9337 | 0.4105 |
| 1.4 | 2.7183 | 0.6065 | 5.5431 | 0.4247 |
| 1.5 | 2.6118 | 0.6188 | 5.1846 | 0.4392 |
| 1.6 | 2.5113 | 0.631 | 4.8555 | 0.4538 |
| 1.7 | 2.4167 | 0.6433 | 4.5533 | 0.4686 |
| 1.8 | 2.3276 | 0.6555 | 4.2756 | 0.4836 |
| 1.9 | 2.2438 | 0.6676 | 4.0203 | 0.4987 |
| 2 | 2.1651 | 0.6796 | 3.7854 | 0.514 |
| 2.1 | 2.091 | 0.6915 | 3.5694 | 0.5293 |
| 2.2 | 2.0215 | 0.7033 | 3.3705 | 0.5447 |
| 2.3 | 1.9562 | 0.715 | 3.1873 | 0.5601 |
| 2.4 | 1.8949 | 0.7264 | 3.0186 | 0.5756 |
| 2.5 | 1.8375 | 0.7377 | 2.8632 | 0.591 |
| 2.6 | 1.7836 | 0.7488 | 2.7199 | 0.6063 |
| 2.7 | 1.7331 | 0.7596 | 2.5879 | 0.6216 |
| 2.8 | 1.6858 | 0.7702 | 2.4661 | 0.6368 |

| | | | | |
|---|---|---|---|---|
| 2.9 | 1.6415 | 0.7805 | 2.3539 | 0.6518 |
| 3 | 1.6 | 0.7906 | 2.2504 | 0.6666 |

**buffer text contiguous with table**




## Interpreting (and Checking) the Test Information Curve and Conditional Standard Errors

The TOPI 5E appears far more powerful than the TOPI 5G in terms of the information it returns. A couple of further thoughts can, however, qualify that conclusion. First, Embretson and Reise (2000, p. 270) have asked:

> "How much information is high enough? Well, if the conditional information is around 10, then the conditional standard error is about 0.31. To place this in more conventional terms, a standard error of 0.31 corresponds to a reliability coefficient of 0.90."

From that perspective, the TOPI 5G curve is consistent with our other findings that it is a highly informative, reliable scale, at least through a theta of zero (but less so above that point).

Second, we might inquire, "Could the information curve for the 5E possibly be as high as it is (above 30), between -2 and -1 of theta?" To double-check, we used equation 7A.6 (from Embretson & Reise's appendix to Chapter 7, that indicates the relationship between the Test

Information Curve and the Conditional Standard Error of the Mean. Reformulated slightly here to fit on one line (i.e., to remove the square-root sign in the denominator), the equation reads:

$$SE(\theta) = 1 / (TI(\theta))^{½} \quad (7A.6)$$

Where SE(θ) is the conditional standard error and *TI*(θ) is the conditional test information (Embretson & Reise, 2000, pp. 184–185).

To ensure that the calculations performed by IRTPRO conformed to that relationship, we entered the values from the tables above in a few instances to double check the relationships, and worked through the formula in Table 8.7. As can be seen, all the values check.

**buffer text contiguous with table**

Table 8.7

Embretson and Reise's Formula 7A.6 Applied as a Check to Figure 1 Using Values from Table 8.6

| | SE(θ) From IRTPRO Table | TI(θ) From IRTPRO Table | $TI(\theta)^{½}$ | $1/ TI(\theta)^{½}$ SE(θ) Calc. value | |
|---|---|---|---|---|---|
| TOPI 5G @ -2 | .33 | 9.02 | 3.00 | .33 | okay |
| TOPI 5G @ +2 | .68 | 2.16 | 1.47 | .68 | okay |
| TOPI 5E @ -2 | .21 | 22.9 | 4.79 | .21 | okay |
| TOPI 5E @ +2 | .51 | 3.79 | 1.95 | .51 | okay |
| | | | | | |

**buffer text contiguous with table**

As a further check (actually conducted first), we informally read the values from the curves of Figure 1 itself. The results were approximately the same, as can be seen in Table 8.8, although the worked-through values are only approximately the same given the distortions involved in using the smooth curve and then reading them from a figure while blind to the actual values.

**buffer text contiguous with table**

Table 8.8

Embretson and Reise's Formula 7A.6 Applied as a Check to Figure 1 Using Approximated Values Read Informally from Figure 8.1

| | SE(θ) read from graph | TI(θ) read from graph | $TI(\theta)^{½}$ | $1/ TI(\theta)^{½}$ SE(θ) Calc. value | |
|---|---|---|---|---|---|
| TOPI 5G @ -2 | .36 | 9.5 | 3.08 | .33 | Approx. okay |
| TOPI 5G @ +2 | .70 | 2.5 | 1.58 | .63 | Approx. okay |
| TOPI 5E @ -2 | .21 | 23.5 | 4.85 | .21 | Approx. okay |
| TOPI 5E @ +2 | .65 | 4.5 | 2.12 | .47 | Approx. okay |
| | | | | | |

**buffer text contiguous with table**

The TOPI 5E is, then, far *more* reliable than needed at values of theta between -2.5 and .5, allowing for the option of trimming some items that function best in that range sometime in

the future. After a theta of about .5, however, its discrimination declines in a familiar and less-than-desirable fashion—although still outperforming the earlier-developed TOPI 5G.

# Chapter 9. Reanalysis of the TOPI-at-Work Data Using the TOPI 5G (Table)

Here are the selected results from the reanalysis of the TOPI-at-work data using the TOPI 5G.

**buffer text contiguous with table**

Table 9.1

Comparative Predictive Validity of the TOPI 5G and TOPI 4R scales with Selected Variables from Mayer et al. (2018). "Employees with High Personal Intelligence…" (*N* = 481)

| | TOPI5G | TOPI 4R Mean [T4Rmeanscl] | Factor 1 [T14Rfa_t] | Factor 2 [T14Rfb_t] |
|---|---|---|---|---|
| Number of items | 47 | 67 | 29 | 31 |
| Mean | 49.34 | 49.23 | 49.24 | 49.22 |
| S | 14.02 | 10.90 | 11.59 | 11.41 |
| Skew | -.861 | -1.12 | -1.12 | -.66 |
| Alpha Reliability | .92 | .94 | .90 | .89 |
| **Correlations with** | | | | |
| TRG47 | 1.00 | .98 | .91 | .96 |
| Age | .18 | .20 | .18 | .19 |
| Gender | .20 | .20 | .19 | .19 |
| CWB Score | | | | |
| Overall | -.22 | -.24 | -.26 | -.19 |
| Abuse | -.22 | -.23 | -.24 | -.20 |
| Sabotage | -.27 | -.30 | -.30 | -.25 |
| Theft | -.24 | -.27 | -..22 | -.22 |
| WDQ | | | | |
| Social Support | .15 | .15 | .14 | .14 |
| Vocabulary | .50 | .49 | .43 | .50 |

**buffer text contiguous with table**

# Chapter 10. Further Thoughts on Two Factors versus One

## Does Personal Intelligence Consist of Two Mental Abilities or One? Revised Thoughts

***The two-factor model at face value.*** The initial look at our findings here and in earlier research certainly indicate that two mental abilities make up personal intelligence: Both exploratory and confirmatory factor analyses of the TOPI 4, 4R, 5, and 5R all indicate that two-factor representations of the TOPI tests offered marked improvements in model fit relative to a one-factor approach. Such findings strongly argue for the TOPI tests as two-factor instruments—that is, that people employ two mental abilities to problem-solve in the area—at least if one takes such findings at face value.

We argue against taking the results at face value, however: Several considerations, explained below, argue for a one factor model rather than two.

***Consideration 1. More factors always fit better.*** Factor models always fit better as more factors are added, given that the technique's algorithms better approximate data when estimating a larger number of parameters; this occurs almost always when additional factors are added. Ergo, the mere appearance of a better-fitting two-factor model is not conclusive of additional substantive factors, in-and-of itself.

***Consideration 2: The two-factor solutions were challenging to interpret.*** Distinct mental abilities, should they exist, can typically be distinguished according to the kinds of problem-solving they facilitate. We regarded the two mental abilities of the TOPI 4 and 4R as challenging to characterize, but ultimately, we believed we had arrived at approximate themes for each one. Interpretation of the two mental abilities identified for the TOPI 5 proved so elusive, however, that we ultimately decided it was most prudent to abandon any effort to characterize them.

Moreover, in our examination of marker items that represented factors one and two of the TOPI 5, we found evidence that the covariance among them might have arisen from context effects (i.e., the placement of items in the test), perhaps due to priming, practice, or simply fatigue effects generated by other items that led up to the factor-defining items themselves.

***Consideration 3: The two-factor model was unstable.*** If there really were two meaningful mental abilities in the area, they should be reflected by the same groups of test items, in test after test. Here, however, the precise nature of the two factors shifted depending upon the test form employed: A model created for 4$^{th}$ generation TOPI tests failed to fit the data generated by the 5$^{th}$ generation of TOPI tests, and vice versa.

***Consideration 4: The two-factor models indicated the two mental abilities were very highly correlated.*** The numerous times that tests of our two-factor models generated estimated correlations between the two dimensions of $r$ = .92 or higher spoke perceptibly to us of the similarity—if not identity—between the two purported mental abilities.

***Consideration 5: The one-factor model was better than it looked initially.*** The apparently less-than-stellar fits of the one-factor models to the TOPI appeared due to a relatively minor issue: the over-similarities of a few pairs of individual items—unsurprising given the

2,211 possible pairs of the 67 items of the TOPI 4R, and the 20,910 possible pairs of the 205 items of the TOPI 5. Aberrations in fit are common in tests with many items (Little, Cunningham, Shahar, & Widaman, 2002). Indeed, in Study 4, we could improve the fit of the one-factor models to conventionally-acceptable standards simply by removing about a dozen single members of the most problematic item pairs.

***Consideration 6: The single-factor model generalized over all samples and test versions.*** Unlike the two-factor models, we succeeded in fitting a one factor model consistently and reasonably well across all six data sets examined here: across college students who were civilians or in the military, and across two different generations of TOPI tests. We could not duplicate that performance with two-factor models.

# References

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, N.J: L. Erlbaum Associates.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*(2), 151–173. https://doi.org/10.1207/S15328007SEM0902_1

Scott, B. (2018). Readability Formulas. *My Byline Media*. Retrieved from www.readabilityformulas.com

UNH Admissions. (2016). ROTC at UNH. Retrieved from https://admissions.unh.edu/rotc-unh

U.S. Department of Defense (Author). (2013). "Ethnic breakdown of active-duty officers." Figure 2.20, p. 21 in 2013 Demographics profile of the military community. Office of the Deputy Assistant Secretary of Defense and ICF International. Retrieved from Downloaded from http://download.militaryonesource.mil/12038/MOS/Reports/2013-Demographics-Report.pdf.

Yoon, M., & Lai, M. H. C. (2018). Testing factorial invariance with unbalanced samples. *Structural Equation Modeling*, *25*(2), 201–213. https://doi.org/10.1080/10705511.2017.1387859

# End of Technical Supplement