

# Measuring preferences

## *What are the MBTI word pairs and phrase questions doing?*

John Bathurst and  
Peter Geyer

What does it mean when  
the phrase question and  
word pair sections of the  
MBTI point to opposite  
preferences?

How do the phrases and  
word pairs point towards  
actual type preferences?

Currently-available MBTI forms contain a mix of little situations with forced-choice responses (*phrase questions*), and *word pairs* of which one ‘appeals’ to the subject. Originally, however, the MBTI (or BMTI, as it was first called) comprised phrase questions only.

**Mary McCaulley** (1981) has recounted Isabel Myers’ concern that many people (predominantly extraverts) take a situational approach in answering MBTI questions, which was not her intent at all.

In searching for a solution to this problem, Myers came up with the idea of the word pairs. Her reasoning was that in the phrase questions people were asked to respond to a key word in each option anyway, so taking away the context implied by a phrase might be a solution.

This proved to be quite successful, with many of the word-pair questions proving to be better items. It also enabled Myers to offer more questions on the MBTI—something that concerned her. Word pairs first appeared on Form D and also Form E (originally designed for school populations), but were first available to the research public via Form F (1958).

As with other aspects of MBTI questions, Myers investigated similarities and differences between the phrases and word pairs. For example, what does it mean when these two sections of the Indicator point towards opposite preferences? Which is correct? How do the phrases and word pairs point towards *actual* type preferences?

These are good questions for which there do not appear to be clear answers, partly because details of Myers’ research have never been made public. No mention is made of this area of research in the 1962 *MBTI Manual*. In the 1985 *Manual*, how-

ever, Myers and McCaulley offer three hypotheses for the dynamics of the phrase questions and word pairs (p 61):

- **Hypothesis 1:** People who wish they were a different type may tend to answer the phrase questions according to *what they are*, and the word pairs according to *what they would like to be*: that is, what ‘appeals’ to them.
- **Hypothesis 2:** People who tend to feel insecure about the worth of their type may tend to reverse that process and answer the phrase questions with what seem ‘*right answers*’, and the word pairs with answers that reflect *what they really feel they are*.
- **Hypothesis 3:** People who are secure in their type but are working to develop their less-liked processes may tend to present a *mixed picture* on the word pairs versus the phrase questions.

**Otto Kroeger** has suggested an additional reason: that word pair scores that go in the opposite direction to an overall preference score can illustrate a difference between the person *at home* and *at work*. There is no published research to support this point of view, although the scoring of word pairs separately from the overall preference score, in order to elicit this information, has been taught in Otto Kroeger Associates’ MBTI qualifying workshops over two decades.

Indeed, there is almost no research to test these hypotheses. The one exception is a study by Nechworth & Carskadon (1979), who asked students who had recently taken the MBTI to answer the Indicator the way they would *like* to be. Their scores changed significantly on the phrase questions, but not on the word pairs. This study therefore supported Hypothesis 2.

## John Bathurst's study

Our hypothesis is that *the section of the MBTI with the more fully-discriminating items is the better predictor of best-fit preferences*. On that basis, from Table 1 we would expect the better predictors for each section to be:

- E-I** phrase questions
- S-N** word pairs
- T-F** word pairs
- J-P** phrase questions

To test this hypothesis, John used a convenience sample of 4929 MBTI records that had both *best-fit preferences* and *separate scores* for the word pairs and phrase questions.

These records were obtained from a wide variety of sources, such as churches, community colleges and organisations where the MBTI was used for personal or career development purposes. The administrators of these programs submitted participants' MBTI forms along with their best-fit types.

To gain an understanding of the predictive abilities of the word pairs and the phrase questions, the preferences indicated by each were compared to the best-fit preferences. The sections which better predict best-fit preferences are shown in Table 2:

- The *phrase questions* are better at predicting E-I and J-P preferences
- The *word pairs* are better at predicting F preferences, and T preferences for women.
- *Neither* the phrase questions nor the word pairs are significantly better at predicting S-N preferences, nor T preferences for men.

The hypothesis is therefore demonstrated to be *partially true* in that, as expected:

- the **E-I** and **J-P** scales are predicted by the phrase questions; and
- apart from T for males, the **T-F** scale is predicted by the word pairs.

Neither section better predicts the **S-N** scale, so for that preference the hypothesis is not proven to be true.

This lack of research has led to a recommendation against the practice of providing separate scores for the different sections of Form M (Myers, McCaulley, Quenk & Hammer, 1998)—and, to avoid inferences being made about the significance of any discrepancies, the removal of separate scores from generated reports. However, it is a trivial exercise to score the word pairs and phrase questions separately to give further usable information on type preferences.

The mix of phrase question and word pair **Form G** items may give a clue as to what is going on. Ideally there should be an even mix of preference items in both sections: but, as Table 1 shows, this is not the case.

- On the **E-I** scale there are only six word pair items. Two of these contain zero weights and are not fully discriminating, resulting in only four fully usable items.
- On the **T-F** scale, there are only seven phrase question items. Four of these contain zero weights, leading to only three fully usable items.

We can see that the E-I score comes predominantly from the phrase question items, and the T-F score from the word pair items. When looking at the differences between preferences derived from word pairs and from phrase questions, it would seem that, following the greater number of fully-scored items, the more reliable predictors of preference would be the phrase questions result for E-I, and the word pairs result for T-F.

Table 1: Numbers of items used to measure preferences in the phrase question (PQ) and word pair (WP) sections of the MBTI Form G

	Total items		Zero weighted		Fully discriminating	
	PQ	WP	PQ	WP	PQ	WP
<b>E-I</b>	15	6	3	2	12	4
<b>S-N</b>	11	15	2	3	9	12
<b>T-F</b>	7	16	4	3	3	13
<b>J-P</b>	16	8	4	1	12	7

*Zero-weighted* items have a zero weight for one of the preference pairs. *Fully-discriminating* items have non-zero weights for both preferences.



**John Bathurst** (ISTP) looks after the New Zealand MBTI database, which holds records from the last 18 years of MBTI use. His interest is in the performance of the MBTI and the factors that influence its accuracy. He has also supported numerous graduate theses that drew on type preferences for groups under study.

John is a senior lecturer at The Open Polytechnic of New Zealand, where he teaches organisational psychology, and assists with the delivery of courses in the psychology and management domains.

idic@paradise.net.nz

Table 2: Comparative accuracy of word pairs and phrase questions in predicting best-fit type preferences (MBTI Form G)

Best-fit preference	n	Accuracy		Better predictor
		Phrase Questions	Word Pairs	
E	2257	87.4%	72.5%	PQ
I	2672	87.8%	79.2%	PQ
S	2456	84.9%	84.3%	-
N	2473	84.2%	86.5%	-
T				
female	1038	69.7%	85.8%	WP
male	1004	88.1%	89.2%	-
F				
female	2225	78.7%	83.6%	WP
male	662	58.5%	78.9%	WP
J	2791	88.8%	79.7%	PQ
P	2183	82.0%	75.6%	PQ

This study points towards a purely psychometric driver for the differences between the phrase questions and word pairs. Where does this leave the interpretation of differences in preferences indicated by the different parts of the Indicator? There seem to be two situations that arise: firstly, where *low* or *tied* scores result from a balancing of the word pair and phrase question scores; and secondly, where the preference score is *moderate*, and the two sections indicate opposite preferences.

With the *low*- or *tied*-score situation, the individual is likely to be unsure of their preference and looking for clues as to what it might be. They will probably go along with the preference indicated by the overall score (Bathurst, 2004).

More useful information could be gained from the *moderate* score situation, where the differences between the two sections may be saying something usable about the person's true preferences. Perhaps one of the three hypotheses posed by Myers and McCaulley (1985) then comes into play. Unfortunately, this study does not throw any light on this, other than suggesting that the differences may mean nothing at all.

## Peter Geyer's practice

I've taught people how to score the MBTI after the Otto Kroeger fashion for 10 years. Outside qualifying workshop settings, I've also used it in my MBTI consulting work as part of my preparation and interpretation.

The general principle of the method is that some extra information about the person completing the MBTI might be gathered in this way. As John points out, word pair scores in the opposite direction to the overall score occur predominantly when the preference clarity is slight anyway, so the information provided is not crucial.

I usually interpret that sort of result as an indication that the respondent might have part of their life where they use the word-pair indicated preference, not the overall preference. I rarely find that the word pair preference is the person's overall preference. For me, the best context seems to be the difference between home and work. I have found no support for the hypotheses suggested by Myers and McCaulley, neither from my questioning, nor from the person volunteering information.

The rules and ethics of MBTI feedback, as I understand and experience them, suggest that a slight preference on a scale requires you to offer an alternative to the reported type for the respondent to consider. Written evidence for this kind of method, however, is scanty. The MBTI manuals as a whole give broad guidelines rather than a method. Elsewhere, Sandra Hirsh (1985; 1991) has suggested a similar method, also implied by Gordon Lawrence and Charles Martin (2000). So this process is not necessarily universally practised.

Suggesting an alternative to the reported type is intended to provide a stimulus to arriving at a best-fit type, which is the purpose of the MBTI in any case. The purpose is *not* to justify MBTI results. Otto Kroeger also states that more than one slight preference in the results still requires the offering for consideration of only one alternative (undated; 1995; 1996). If one option is suggested from three possibilities (or four, including the original result), the purpose is not to arrive at the 'right' answer, but to set up a cognitive



**Peter Geyer** (INTP) has taught people how to understand and use the MBTI since 1993, and; since 1995 in the Otto Kroeger method of interpreting word pairs. He has investigated presentation and feedback principles from acknowledged type experts, developed his own methods, and written about them for others to use.

Peter also uses with confidence the Kroeger method of interpreting un-scored items on Form G, and has developed a method of interpreting MBTI Step II results using type dynamics.

alchymia@ozemail.com.au  
www.petergeyer.com.au

process whereby the person arrives at their *own* answer, which is the purpose of feedback. Professional judgement comes into play here as far as suggesting an effective alternative goes, and this involves looking at all the information available, including the demographic information provided by the respondent on the answer sheet.

A moderate preference also requires that no alternative be given, even if the word pair scores go in the opposite direction to the overall score. In that sort of situation, which is rare, I would put some confidence in asking the person when they would use the opposite preference.

Two examples used in teaching this scoring method (OKA 1998) illustrate the practice and interpretation:

**Example 1:** Reported type ENFP, with word pairs in the opposite direction on S-N.

Points		Result	Clarity
E 11 <sub>4</sub>	I 10 <sub>2</sub>	<b>E</b>	Slight
S 11 <sub>10</sub>	N 15 <sub>7</sub>	<b>N</b> ☿	Slight
T 8 <sub>6</sub>	F 16 <sub>12</sub>	<b>F</b>	Moderate
J 8 <sub>1</sub>	P 14 <sub>5</sub>	<b>P</b>	Moderate

Word pair scores are shown in subscript to indicate they are part of the overall score, not an adjunct or superior to it. A ☿ symbol is placed next to any word-pair scores that go in the opposite direction to the overall score. (This symbol is simply a reminder to the interpreter; it is not shown to the respondent.)

In *Example 1* the word pair anomaly is on a scale (S–N) where there is a *slight* preference. Another scale (E–I) also indicates a slight preference, so it’s part of general information, rather than something that stands alone.

As one option only is required to help the person arrive at a best-fit type, a choice has to be made among three alternatives: INFP, ISFP, ESFP. My practice is to suggest the option *most different* from the reported type, so I would suggest ISFP, simply to get the person reflecting on all the options. The word pairs here might represent a specific activity—home, or work—or nothing.

**Example 2:** Reported type ENFJ, with word pairs in the opposite direction on J-P.

Points		Result	Clarity
E 21 <sub>6</sub>	I 0 <sub>0</sub>	<b>E</b>	Very clear
S 2 <sub>1</sub>	N 24 <sub>16</sub>	<b>N</b>	Clear
T 0 <sub>0</sub>	F 24 <sub>18</sub>	<b>F</b>	Very clear
J 14 <sub>0</sub>	P 8 <sub>6</sub>	<b>J</b> ☿	Moderate

With *Example 2*, the word pair anomaly is on J–P, with a moderate preference for J. In this case there would be no point offering an alternative, as the moderate results give a high level of confidence that J is in fact the person’s preference. But I would be looking for something in the feedback session that might indicate a component of this person’s life where they prefer P. It could be scheduling spontaneity on a weekend, for instance.

Their other results give a high level of confidence that the person will manifest E, N, and F in their behaviours. And when they walk in the room there may be dress and language expression that indicates J or P.

Essentially, the idea here is to use information to enable a more rounded and natural portrait of the respondent’s personality as a whole—not just whether they prefer one thing over another, which sometimes isn’t very informative at all to anyone.

The dilemma is that people without these word pair anomalies might also be able to describe when they use the other preference. I certainly find this in Step II interpretation where a person may say that they use the other pole of their preferred subscale or facet, notwithstanding that they have not answered any questions in that direction.

Whatever the truth of the matter, it’s important to remember that the MBTI is a sorting device. It asks what you *prefer*, but it doesn’t imply that you have no acquaintance with the opposite preference.

Or, as Isabel Myers said, ‘It’s not asking you what you *can do*.’ ❖

## References

Bathurst, John 2004, ‘Breaking ties: Interpreting very low MBTI scores’, *Australian Psychological Type Review* 6:3, 43-45.

Geyer, Peter 2003, ‘MBTI word pairs and their interpretation’, *An MBTI qualifying workshop course workbook*, Warrnambool: Otto Kroeger Associates/ Peter Geyer.

Hirsh, Sandra K 1985, *Using the Myers-Briggs Type Indicator in organizations: A resource book*, Palo Alto, CA: CPP.

Hirsh, Sandra K 1991, *Using the Myers-Briggs Type Indicator in organizations* (2nd edition), Palo Alto, CA: CPP.

McCaulley, Mary H 1981, ‘Interpreting scores on the MBTI’ (audio tape), MBTI / APT IV Conference, Stanford University, Palo Alto, CA.

Myers, Isabel B (undated, 1962), *MBTI manual*, Palo Alto, CA: CPP.

Myers, Isabel B, and Mary H McCaulley 1985, *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*, Palo Alto, CA: CPP.

Myers, Isabel B, and Mary H McCaulley, Naomi L Quenk, and Allen L Hammer 1998, *MBTI Manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd edition), Palo Alto, CA: CPP.

Nechworth, J A, and T G Carskadon 1979, ‘Experimental validation of an assumption underlying the clinical interpretation of discrepancies between Myers-Briggs Type Indicator scores computed separately from word-pair and phrased question items’, *Research in Psychological Type* 2, 56-59.

Otto Kroeger Associates (undated), *Typewatching qualifying workshop: Trainers manual*, Fairfax, VA.

Otto Kroeger Associates 1995, *An MBTI qualifying workshop course workbook*, Fairfax, VA.

Otto Kroeger Associates 1996, *OKA trainer’s manual* (draft), Fairfax, VA.

Otto Kroeger Associates 1998, *Teaching materials for scoring MBTI Form M*, Fairfax, VA.