

Method, Reliability & Validity, Statistics & Research:

A Comprehensive Review of Belbin Team Roles



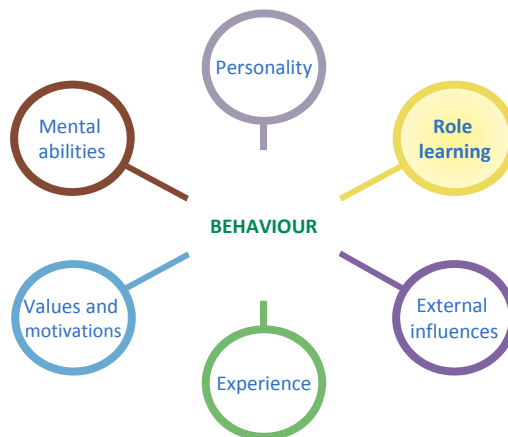
Contents

Introduction – Measuring Behaviour	1
Construction of the BTRSPI	3
Is the BTRSPI an ipsative test?	4
Administration and Scoring of the BTRSPI	6
Why use Interplace?	6
The Specialist Role	6
Observer Assessments	7
Norming	8
Team Role Advice	8
Description of the Nine Team Roles	9
Reliability	11
Test-retest reliability	11
Internal consistency	11
Validity	13
Face validity	13
Construct validity	13
Convergent validity	13
Discriminant validity	14
Concordant validity	14
The BTRSPI and other measures	15
Criterion validity	15
Belbin in practice	16
Observer Assessments	16
Assessing a job in terms of Belbin Team Roles	17
Belbin in industry	19
Testing, Bias and Fairness	20
Belbin and Gender	20
Belbin and Ethnic Origin	21
Belbin and Age	21
Belbin and Management	22
Further Reading	23
Books from Belbin	23
Other titles from Belbin	23
Research Studies	24
Recommended Articles	25
Other Articles	26

Introduction - Measuring Behaviour

The Belbin Team Role Self-Perception Inventory (BTRSPI) was designed to measure behavioural characteristics which individuals display when working in teams. Belbin Team Role Theory was devised as a result of the studies conducted by Meredith Belbin in the 1970s. For more details of this initial research, please visit the Belbin website.

Since the BTRSPI measures behaviour rather than personality, it is not considered to be a psychometric test (those which measure attributes of personality). Rather, personality is one of many factors which can influence behaviour. Other factors include internalised values and motivations, and the external working environment or “Field Constraints”:



Whilst most personality traits are acknowledged to be fairly constant, behaviour can change more readily, adapting to changes in any of those factors which influence it. As a result, Belbin expects that Team Role preferences might change over time. Whilst it is unlikely that an individual’s Team Roles will change dramatically or be reversed altogether, some alterations are expected, in line with a change of job role or work environment, or as a result of a major life change.

The BTRSPI measures behaviour because Belbin believes that this provides the most useful and verifiable information regarding an individual to a recruiter, manager or consultant, as well as to the individual concerned. Whilst it could be argued that only the individual himself knows his own personality, behaviour is observable and can be interpreted and used to predict future reactions and conduct.

The difficulty when measuring personality alone is that there may be a large discrepancy between personality and behaviour. Whilst an individual may purport to be an extravert, that person’s behaviour in the workplace may lean towards introversion. The individual’s self-perception of extraversion may indicate limited self-awareness or may reflect a personality trait the individual wishes to possess. It is arguable that identifying certain personality traits does not directly help the manager concerned with recruitment or promotion. In the case of many psychometric tests, managers expend much energy understanding the psychometric dimensions or traits, rather than applying the knowledge to improve performance.

Rather than providing information regarding individual personality traits, the BTRSPI gauges behaviour in order to identify groupings or clusters (Team Roles) which characterise an individual’s behavioural contribution to the workplace. For example, you might find a question in a personality test along the lines of: “When I’ve made a decision about something, I still keep wondering whether it’s right or wrong.” Here, the focus is on how the individual thinks and feels. By contrast, the BTRSPI asks questions like: “I can be relied upon to finish any task I undertake,” focusing on practical contributions an individual might make.

As well as self-reporting, whereby an individual identifies behaviours he believes he exhibits, the BTRSPI uses Observer Assessments (OA) to substantiate or refute these claims with the perceptions of colleagues, managers and other co-workers, to form a more rounded view of the individual's contribution. Aritzeta, Swailes & Senior write:

“

The dynamic configuration of team roles measured by the [B]TRSPI and the relative stability of traits measured by personality questionnaires leads to the conclusion that traits measured by the latter are different from those measured by the [B]TRSPI. Thus, both instruments may be tapping different but complementary constructs.

”

~ Aritzeta, Swailes & Senior, 2007¹

Many individuals and organisations are concerned that the introduction of psychometric tests might lead to “pigeon-holing” or labelling of individuals. With Belbin Team Roles, the relationship between an individual and the Team Roles they exhibit is a far more complex one. An individual does not have one Team Role, but a combination of preferred, manageable and least preferred roles. The distribution and interrelation of these roles, as shown in an individual's report, have a great influence on the way the roles will be played out in practice and experienced by others. Whilst an individual may claim to prefer or enjoy a particular role, it does not necessarily mean that they can or should play only this role. The theory of Team Roles is concerned with acknowledging strengths and weaknesses, but also with cultivating strengths to becoming a model, strong example of a given Team Role type.

¹*Belbin's Team Role Model: Development, Validity and Applications for Team Building* (Aritzeta, Swailes & Senior, 2007), p. 110.

Construction of the BTRSPI

The BTRSPI is a behavioural test designed for use in organisational and work settings. The BTRSPI was formulated by Meredith Belbin in the 1980s following on from his research at Henley Management College and inception of Team Role Theory.

The BTRSPI measures nine dimensions or Team Roles and has one scale known as Dropped Points (DP), which measures claims about oneself rather than valid Team Role contributions. The inventory consists of eight sections, with each section containing a heading and ten statements, one item per Team Role and a tenth item representing DP.

The headings give a work-based scenario or situation with which the individual can identify. This is intended to anchor the behaviours described in a familiar work context and to encourage candidates to reflect and draw on examples from their own experiences.

When completing the BTRSPI, candidates are asked to distribute ten marks in total per section of the inventory. If a candidate identifies equally with only two statements, 5 points should be allocated to each of the two statements. If four of the statements are relevant, but two more so than the other two, the allocation of points might be 3, 3, 2 and 2, as shown below:

BELBIN®

Self-Perception Inventory (SPI) Step 2 of 11

Section 1

I believe I can make positive contributions to a team because:

I am quick to see and take advantage of new opportunities.	<input type="checkbox"/>
I am seen as a natural team player.	2
I am happy to take on varied work as and when the team requires.	<input type="checkbox"/>
I can think laterally to solve problems.	3
I am good at identifying and using the potential in fellow team members.	2
I am keen to improve things by focusing on the details.	<input type="checkbox"/>
I am enthusiastic about applying my training and expertise.	<input type="checkbox"/>
I am ready to speak out in the interests of making things happen.	2
I ensure that my work is delivered on time.	<input type="checkbox"/>
I can offer reasoned and balanced judgements of different courses of action.	<input type="checkbox"/>
Total	10

← Previous Next →

BELBIN®

This is then repeated for each section of the inventory. Candidates may only allocate marks in whole numbers and are asked to avoid either extreme (allocating all 10 marks to one statement or 1 mark to each) where possible.

Is the BTRSPI an ipsative test?

The BTRSPI is designed to ascertain as much information as possible about an individual's Team Role preferences, whilst keeping the inventory manageable in terms of item length, inventory length and answering style. Many psychometric tests such as the 16PF and OPQ require the respondent to evaluate around 200 items using a Likert scale (e.g. ticking an answer along a spectrum from "Strongly agree" to "Strongly disagree", usually with a neutral option of "Don't know" or "Not sure"). In contrast, the BTRSPI asks the respondent to consider items within each section in relation to each other and to prioritise some above others. As a result, the questionnaire takes only between 15 and 20 minutes to complete, less time than many psychometric evaluations.

Since the total score achieved in the BTRSPI is always constant, it is an ipsative measure overall. This means that a respondent must express a relative preference between two or more statements measuring different characteristics, thus creating a degree of interdependence between the characteristics being measured.

However, since the items are dispersed in the 8 sections such that there is one item for each role in each section, the scores given to items for any Team Role are not fully ipsative, since they do not sum to a constant value. Whilst the scores for items in the same scale are independent of each other, they are partly dependent on the scores given to other scales. In other words, the BTRSPI is ipsative within its sections (since scores always sum to 10) but not between its sections.

Some early research studies criticised the scoring style of the BTRSPI, suggesting that it "forced choice" between statements. In 1993, Meredith Belbin defended the inventory, claiming that "some restriction of choice [was] operationally desirable" since "self-rating on independent scales yields little of value in industrial and occupational settings" (see discussion above).

In their 1998 research, Sommerville & Dalziel converted the BTRSPI to a Likert-type scale questionnaire. They found that 73% of participants had the same Team Role across both versions of the test, indicating that there is no significant difference in the prediction of Team Roles between the two versions.²

In 1993, Meredith Belbin maintained that Likert-type scales caused more frustration among candidates in industry and later, in research conducted in 2004, Aritzeta, Swailes & McIntyre-Bhatty point out that Likert-type scales bring their own set of problems:

“ There is a strong controversy in the literature about the ipsative scoring of constructs, and such discussion while very interesting is still far from clear. Ipsative data is not free from criticism as neither are normative data [sic]. Likert type scales are not free from important threats as central tendency bias, acquiescence, social desirability and falsification of responses which may be much larger from these types of scales than for ipsative scales. ”

~ Aritzeta, Swailes & McIntyre-Bhatty, 2004³

²*Project teambuilding – the applicability of Belbin's team role self-perception inventory* (Sommerville & Dalziel, 1998), pp. 166-167.

³*Further Evidence on the validity of the Belbin Team Role Self Perception Inventory and the Observer's Assessment Sheet* (Aritzeta, Swailes & McIntyre-Bhatty, 2004), p.8.

Arguably, a Likert-type scale also forces choice by asking candidates to choose the neutral response if no other applies. For example, if a bipartite statement is presented, with the candidate agreeing to one part of the item and disagreeing with the other, a neutral response could reflect the poor wording of the item rather than the candidate's true response. Additionally, the Likert-type scale assumes that the figurative distance in "preference" between "Strongly agree" and "Agree" is the same as that between "Agree" and the neutral response or between "Disagree" and "Strongly disagree". This is an assumption which aids calculation and analysis, but again, is not reflective of true response.

Having undertaken comprehensive statistical and factor analysis on data from more than 5000 candidates who have completed the BTRSPI, Swailes & Aritzeta conclude that the scoring system of the BTRSPI has no adverse effect on its construct validity and that levels of interdependency are low (please see the "Validity" section below for further discussion)⁴. Since the BTRSPI's current format confers no statistical disadvantage and confers considerable advantages from the viewpoint of the candidate, Belbin has chosen to retain the format.

⁴*Scale Properties of the Team Role Self Perception Inventory* (Swailes & Aritzeta, 2006), p.10.

Administering and Scoring of the BTRSPI

The BTRSPI is available from Belbin Associates at <http://www.belbin.com>. There are also a number of consultants and distributors who are licensed to resell Belbin internationally in English and other languages. Whilst the inventory is readily available online, an Interplace system is required to score the inventory, process the data and produce Belbin reports.

In his book, *Management Teams: Why They Succeed or Fail*, first published in 1981, Meredith Belbin included a self-scoring BTRSPI designed to provide the individual reader with a “quick fix” indicator of what their Belbin Team Roles might be. As further research was conducted, this inventory was shown to be inadequate in determining Team Roles and the Interplace system was developed to handle the norming, data analysis and complex algorithms which arise from different combinations of response to the BTRSPI. Rather than providing simply a ranked order of roles, the Interplace software produces a full feedback report, integrating information gleaned from both the Self-Perception Inventory and Observer Assessments.

Belbin Associates owns the copyright for the Self-Perception Inventory included in *Management Teams* and does not allow this questionnaire to be reproduced in any form. Individuals may purchase the book and complete the self-scoring version of the BTRSPI for their own personal development, but any copying or wider usage is an infringement of copyright and will be prosecuted. For more information, please visit <http://www.belbin.com/rte.asp?id=7>.

Why use Interplace?

The self-scoring questionnaire is now obsolete and is not a reliable way of determining Team Roles. Many research studies have suggested that Interplace provides a much more reliable and valid method of establishing Team Roles.⁵ Moreover, the old self-scoring method does not include the ninth role of Specialist, lacks the balance of observer input, is not properly normed and most importantly, does not offer any Team Role advice. These points are outlined in more detail below.

The Specialist role

Meredith Belbin’s original research in the 1970s identified eight Team Roles. After the initial research had been completed, a ninth Team Role, “Specialist” emerged. This role was discovered only after the Henley experiments had been concluded. Since the business game had been constructed to set all participants on a level playing-field in terms of knowledge and expertise, Specialist behaviours could not emerge. Whilst this premise was useful for the purposes of the experiment, it is not representative of real life. All information regarding the contribution and shortcomings of the Specialist has been gleaned from later experiences in the practical application of the theory in industry.

⁵Please see the “Further Reading” section for more details on research conducted using the Interplace system.

Observer Assessments

Belbin strongly recommends the use of Observer Assessments or OA (our own integrated form of 360-degree feedback) to qualify the individual's self-perception. Whilst many psychometric tests rely entirely on self-reporting, Belbin points to the limitations of this approach. An individual may have little self-awareness, especially if he or she has not been working for very long. Meredith Belbin argues that the need for such corroboration arose from a demand for a more robust way of assessing the potential Team Role contribution of individuals:

“

Line managers were usually wary of using self-reporting measures when reaching crucial decisions about people. That reservation is seldom connected with technical issues of test construction but more with the recognition that people are subject to illusions about the self and are also tempted to distort their responses once they believe that their answers affect job and career prospects. Line managers place a greater emphasis on observations of others, believing that such material has greater validity, in terms of effective decision making, providing it is properly gathered.

”

~ Meredith Belbin, 1993⁶

As well as validating an individual's self-perception with observations of “real-world” behaviour, Observer Assessments provide learning and personal development opportunities. For example, where individuals indicate different Team Role preferences than are identified by their team, discussion may arise as to whether the individual in question is able to achieve full potential or is asked to play other roles for the benefit of the team. For more information on the use of Observer Assessments, please view the “Criterion Validity” section below.

⁶A reply to the Belbin Team-Role Self-Perception Inventory by Furnham, Steele and Pendleton (Belbin, 1993), *Journal of Occupational and Organizational Psychology* (1993), 66, p.259.

Norming

Belbin recognises that different groups and cultures may produce different Team Role balances. Since Belbin is sold and used internationally, the Belbin Interplace system makes provision for this. Individual users of the system can create their own norms for a particular organisation or other defined group, or can choose to refer to UK norms, if preferred. In his article on Belbin and culture, Chris Morison writes:

“

Without norming, it would be impossible to compare scores between roles. Without that comparison, the selection of natural, managed and avoided roles would be impossible. Secondly, norming filters out cultural impacts. This makes the role selection relative to the norms used.”

”

~ Chris Morison, 2008⁷

Team Role Advice

Belbin Interplace uses the data gained from the BTRSPI to produce a full feedback report, interpreting the individual's Team Role preferences in textual and graphical forms. The six-page Self-Perception report includes:










- Team Role Overview
- Your Team Role Preferences
- Team Role Feedback
- Maximizing your Potential
- Feedback and Development Suggestions
- Suggested Work Styles

With the addition of Observer Assessments, an extended ten-page report analysing individual and observer feedback can be provided. For more information and to view sample Belbin Team Role reports, please visit <http://www.belbin.com/rte.asp?id=10>.

⁷*An Investigation of Belbin Team Roles as a Measure of Business Culture* (Chris Morison, 2008), p.32.

Description of the Nine Team Roles

The nine Belbin Team Roles are shown below, along with the strengths and allowable weaknesses for the particular role. According to Belbin Team Role theory, as each Team Role makes an individual contribution to the team, so each Team Role has an allowable weakness which is the flipside of the strength.

Team Role	Contribution	Allowable Weaknesses
Plant 	Creative, imaginative, free-thinking. Generates ideas and solves difficult problems.	Ignores incidentals. Too preoccupied to communicate effectively.
Resource Investigator 	Outgoing, enthusiastic, communicative. Explores opportunities and develops contacts.	Over-optimistic. Loses interest once initial enthusiasm has passed.
Co-ordinator 	Mature, confident, identifies talent. Clarifies goals. Delegates effectively.	Can be seen as manipulative. Offloads own share of the work.
Shaper 	Challenging, dynamic, thrives on pressure. Has the drive and courage to overcome obstacles.	Prone to provocation. Offends people's feelings.
Monitor Evaluator 	Sober, strategic and discerning. Sees all options and judges accurately.	Lacks drive and ability to inspire others. Can be overly critical.
Teamworker 	Co-operative, perceptive and diplomatic. Listens and averts friction.	Indecisive in crunch situations. Avoids confrontation.
Implementer 	Practical, reliable, efficient. Turns ideas into actions and organises work that needs to be done.	Somewhat inflexible. Slow to respond to new possibilities.
Completer Finisher 	Painstaking, conscientious, anxious. Searches out errors. Polishes and perfects.	Inclined to worry unduly. Reluctant to delegate.
Specialist 	Single-minded, self-starting, dedicated. Provides knowledge and skills in rare supply.	Contributes only on a narrow front. Dwells on technicalities.

Team Roles are clusters of behaviour, rather than individual traits or characteristics. As mentioned above, it is envisaged that a candidate will have more than one preferred Team Role. In the report, 'Your Team Role Preferences', an individual's Team Roles are analysed in three categories:

- **Preferred Roles** – those roles which the individual is comfortable playing and which come naturally.
- **Manageable Roles** – those roles which an individual can play if required for the benefit of the team. These may be cultivated to broaden the individual's teamworking experience.
- **Least Preferred Roles** – those roles which the individual does not naturally or comfortably assume. It is generally recommended that the individual avoids contributing in these areas, lest the pitfalls of the behaviour outweigh the strengths.

The nine Team Roles may also be categorised as Action, Social and Thinking roles:

- **Action** – Completer Finisher (CF); Implementer (IMP); Shaper (SH)
- **Social** – Co-ordinator (CO); Resource Investigator (RI); Teamworker (TW)
- **Thinking** – Monitor Evaluator (ME); Plant (PL); Specialist (SP)



As mentioned above, no individual Team Role or characteristic should be considered in isolation. Rather, the specific combination and interaction of an individual's Team Roles – along with observer input and style of response to the BTRSPI – help to shape and inform the Team Role report, with the intricate interplay of Team Roles handled by the Belbin Interplace system. Whilst Team Roles are not likely to change dramatically, individuals who are new to a job or to the world of work may discover more fluctuation in their preferences than those who have been working for much longer. Some individuals may find that only two or three roles come into play, whilst others may find that the variety of their job – or a "Jack of all trades" disposition – calls upon four or five different roles, which can be played as the situation demands.

Reliability

In psychometric testing, a reliable test is one which will produce consistent results when the same individual is tested on different occasions (often referred to as ‘test-retest’ reliability). Often, psychologists and statisticians analyse the internal consistency of a test to verify that different parts of a test are all measuring the same quality or trait (in this case, they are measuring ‘internal consistency’). When evaluating a test, reliability is generally measured before validity, since the reliability of a test places an upper limit on its validity – in other words, a test cannot measure what it purports to measure unless it is stable and consistent in its measurement.

Test-retest reliability

As discussed in “Measuring Behaviour” above, Belbin is not a psychometric test, since it measures behaviour rather than personality. Whilst personality may remain fairly consistent, we would expect behaviour to change, along with a change in job role, for example. Indeed, it is desirable that an individual’s should adapt to suit the demands of the job. Although it is unusual for a Team Role order to become entirely reversed, it is highly likely that preferred and manageable roles may move around during an individual’s career.

Internal consistency

Internal consistency is also known as scale homogeneity, in other words, the ability of items in a scale to measure the same construct or trait. Belbin measures nine Team Roles, which are clusters or constellations of behavioural characteristics, rather than individual traits. As a result, Belbin does not repeat items but rather looks to identify strata of different Team Role behaviours displayed by an individual.

Since Belbin differs from psychometric tests in both its underlying theory and its format and scoring mechanism, measuring its reliability has presented a challenge to researchers over the years. The most common measure of internal consistency is Cronbach’s Alpha (α), which, when applied to the BTRSPI during early research, produced highly variable results. In their research papers, “The Reliability of the (Belbin) Team Role Self-Perception Inventory: Cronbach’s alpha and ipsative scales”⁸ and “Uses and Abuses of Reliability Estimates: The Case of the Belbin TRSPI”⁹, Stephen Swailes and Tim McIntyre-Bhatty explored limitations to Cronbach’s alpha in general and explained why existing measures of reliability were inadequate to evaluate the BTRSPI.

Firstly, the researchers point out a generic limitation of Cronbach’s alpha (that simply increasing the length of the inventory (i.e. the number of items) can increase α .) Also, inter-item correlations (the link between items) can remain low, but can achieve a high α value, so long as they are consistently low. As regards the BTRSPI in particular, it is neither fully ipsative nor non-ipsative, so whilst the total score achieved is always the same, the score for each Team Role can vary. Secondly, respondents do not allocate a value to every item in the inventory. In the past, researchers have assigned zero to all items without a response, rather than assigning a null value, thereby contaminating the results. Researchers have also made use of the obsolete self-scoring version of the inventory and have used small or inappropriate sample sizes when analysing data and drawing conclusions.

⁸*The Reliability of the (Belbin) Team Role Self-Perception Inventory: Cronbach’s alpha and ipsative scales* (McIntyre-Bhatty & Swailes, 2000).

⁹*Uses and Abuses of Reliability Estimates: The Case of the Belbin TRSPI* (Swailes & McIntyre-Bhatty, 2001).

Having recognised these limitations to Cronbach’s Alpha and its applications to the BTRSPI, Swailes et al. formulated and proposed a new measure of reliability, IR α , which offers a weighted mean of average inter-item correlation scores. Using a large dataset of respondents, they calculated α by contaminating null responses with zeros, as earlier researchers had done, for the sake of comparison. Next, they chose only those responses for which every Team Role received a score across the 7 sections of the inventory, so as to avoid the problem of assigning values of zero to null responses. α was calculated again, without contamination of null and using weighted inter-item correlation. In June 2012, these tests were repeated for the new, eight-section version of the inventory. The results are shown in Table 1 below:

	Team Roles								
Team Role	CF	CO	IMP	ME	PL	RI	SH	SP	TW
Cronbach α (with zeros)	.46	.41	.23	.40	.54	.44	.66	.48	.53
Cronbach α (7 scored items)	<u>.67</u>	<u>.79</u>	<u>.56</u>	<u>.60</u>	<u>.72</u>	<u>.78</u>	<u>.68</u>	<u>.65</u>	<u>.76</u>
Cronbach α (8 scored items)	<u>.77</u>	<u>.75</u>	<u>.80</u>	<u>.77</u>	<u>.84</u>	<u>.71</u>	<u>.77</u>	<u>.67</u>	<u>.72</u>

A score of 0.6 or above is considered good or acceptable. Whilst traditional use of Cronbach’s α shows poor results for the BTRSPI, a vast improvement is found for α using Swailes et al.’s new formula for calculating reliability. This study helps to explain why earlier research produced poor results for the BTRSPI’s reliability. In summary, earlier studies made erroneous use of α in relation to the BTRSPI, used much smaller and inappropriate samples and contaminated data by assigning zeros to null responses.

Using weighted inter-item correlation to calculate reliability via Cronbach’s Alpha (α), Belbin Team Roles as measured by the BTRSPI show good or acceptable reliability overall.

Validity

A test is said to be valid if it measures what it claims to measure. There are different kinds of validity: face validity, construct validity and criterion validity. These are explored below in relation to the BTRSPI.

Face validity

Face validity concerns the appearance of the test. If the test appears to be measuring appropriately and correctly, then respondents are more inclined to give the test their full attention and answer it honestly and openly. In their 1996 study, Balderson & Broderick state that the BTRSPI has “very high face validity” and comment on the “acceptability of the measures” or Team Roles “particularly using the Interplace system expert reports”.¹⁰

Construct validity

Construct validity assesses whether the characteristic which a test is actually measuring is meaningful and consistent with what the test is intended to measure overall. **Convergent validity** is concerned with whether a test is similar to those to which it should theoretically be similar. **Discriminant validity** is the extent to which a given scale can be distinguished from other scales which are measuring different concepts or traits. Concordant validity has to do with the level of agreement (or concordance) between one test and another, in this instance, the BTRSPI and Observer Assessment (OA).

Much research has been conducted into the construct validity of the BTRSPI. A comprehensive overview of these studies is provided in a paper by Aritzeta, Swailes & McIntyre-Bhatty.¹¹

Convergent validity

In their 2007 paper, Aritzeta et al. comment that the Team Role model has acceptable convergent validity. They explain that previous empirical studies which have found less favourable results have fallen victim to effects caused by small or inappropriate sample sizes. They write:

“ [This] leads us to conclude that the team role model has acceptable convergent validity. Factor structures for the TRSPI are coherent in its ipsative and normative forms as well as with personality measures. ”

~ Aritzeta, Swailes & McIntyre-Bhatty, 2004¹²

¹⁰*Behaviour in Teams: Exploring Occupational and Gender Differences* (Balderson & Broderick, 1996), p.33.

¹¹*Further Evidence on the Validity of the Belbin Team Role Self Perception Inventory and the Observer's Assessment Sheet* (Aritzeta, Swailes & McIntyre-Bhatty, 2004), pp.3-6.

¹²*Belbin's Team Role Model: Development, Validity and Applications for Team Building* (Aritzeta, Swailes & Senior, 2007), p.111.

Discriminant validity

Fisher, Hunter & Macrosson's 2001 paper provided overall support for discriminant validity in the BTRSPI. The researchers were highly successful in their prediction of the frequency of occurrence of various Team Roles solely from Belbin constructs and stated that this added weight to claims of construct validity for the BTRSPI.¹³

Concordant validity

In their 2004 paper, Aritzeta, Swailes & McIntyre-Bhatty found that the Observer Assessment (OA) showed concordant validity since, out of 3351 observations, 66.4% showed significant "Team Role agreement" between observers. When analysing the BTRSPI alongside the OA, 75% of the correlations could be considered from moderate to strong.

Whilst these findings are useful and positive, it is important to recognise that, owing to the nature of Belbin Team Role theory, it is possible that observers might identify very different behaviours in the same individual. Team Role patterns on reports can be regarded as coherent, compatible, discordant or confused, depending on the level and type of agreement between self-perception and observer assessments. In cases where self-perception differs greatly from observer input, a number of factors can be responsible for this outcome:

- Limited self-awareness – this is more likely to cause disagreement between self and observers, or between observers, if someone is new in a particular job role, or new to the work environment in general.
- The individual in question may be playing a different role than he or she desires to play or would be best at playing. Whilst the observer assessments may reflect the current Team Role contribution, the self-perception might reveal an aspiration to play a different role.
- An individual may undergo a change in values, which might not be immediately obvious to colleagues, but may influence the individual's outlook and behaviour.
- The role of observers – it is important to establish the exact working relationship between an individual and those observing them. An observer may not know the person well enough to comment on their working style, especially if they have not worked together for long. Additionally, the observers' own Team Roles may come into play when answering Observer Assessments.

It is likely that any dataset will contain a mix of these different kinds of patterns. When significant agreement is found between self-perception and observer assessments, this is used along with other statistical factors, to determine whether or not an individual can be considered a strong example of a particular Team Role type.¹⁴

¹³*The distribution of Belbin team roles among UK managers* (Fisher, Hunter & Macrosson, 1998), p.132.

¹⁴In-depth information on analysing and providing feedback on Belbin reports is available via our Belbin

The BTRSPI and other measures

The BTRSPI has been analysed alongside other measures, such as the 16PF and OPQ. In his 2007 study, Aritzeta, Swailes & Senior commented that:

“

Taking the empirical studies together, there is sufficient evidence that definitions of team roles are valid and that independently of the instrument used to measure team roles, results are consistent with other theoretical models. The team role model shows evidence for validity that cannot be disregarded [...]. Knowing the type of association that a team role shows with individual cognitive styles, conflict managing behaviour and the other areas explored will help to better understand team dynamics and facilitate team building behaviours.

”

~ Aritzeta, Swailes & Senior, 2007¹⁵

For more information on validity studies which analyse the BTRSPI alongside other measures, please consult the “Further Reading” section below.

Criterion validity

Criterion validity is concerned with the applicability of a test – it is proved by demonstrating that the results of a given test relate in a meaningful way to an external criterion – for example, job performance. This is often easier to judge when dealing with ability tests, but nevertheless, many organisations across the world can attest that Belbin Team Role theory has been proven to make teams more successful.

In 1995, Dulewicz investigated the association between Team Role and responsibility or status.¹⁶ In order to do so, he measured salary, total remuneration, and the total number of staff and total expenditure budgets, for whom the individual had responsibility. He discovered that Team Roles seemed to be independent of measures of salary and job responsibilities.

Belbin Team Role theory does not specify that one Team Role or another is required for a managerial role, or indeed, a role attracting a higher salary. Rather, Shi & Tang’s study, published in 1997, makes the useful comment that a given environment may promote the rise of particular Team Roles. For example, a threatening or political climate may promote a sober and discerning individual to a managerial position, who is shrewd in judging their environment and others around them – in other words, a Monitor Evaluator. In an organisation stuck in a rut and lacking ideas and resources, a Plant or Resource Investigator is likely to stand out; whilst in a stable and industrious environment, the characteristics of an Implementer or Completer Finisher might be more highly valued.¹⁷

There are two measures which can be introduced to ensure that an individual’s Team Roles are a good match with their behaviour in a real-world scenario. These are i) Observer Assessments and ii) Job profiling.

¹⁵*Belbin Team Role Model: Development, Validity and Applications for Team Building* (Aritzeta, Swailes & Senior, 2007), p.108.

¹⁶*A Validation of Belbin’s Team Roles from 16PF and OPQ using Bosses’ Ratings of Competence* (Dulewicz, 1995)

¹⁷*Team role behaviour and task environment: An exploratory study of five organizations and their managers* (Shi & Tang, 1997), p.93.

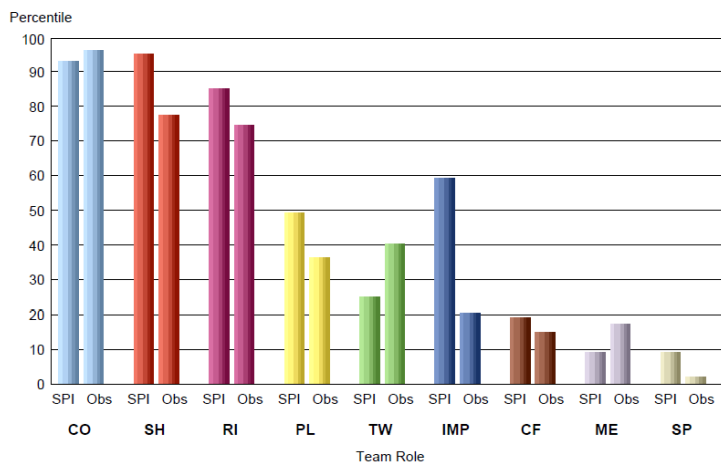
i) Observer Assessments

Observer Assessments can be added to a BTRSPI to ensure that an individual’s report is not dependent on self-awareness and -understanding, but rather is grounded in the perceptions of those who work with the person in question.

Observers are recommended because:

- Self-perception is subjective.
- An individual may have limited self-insight.
- An individual may answer regarding how they wish to behave or be perceived, rather than how they really are.
- Responses from self-perception are isolated rather than democratic.
- Since only one point of view is provided, there are limited or no opportunities for learning and self-development.

When Observer Assessments are added to Self-Perception responses, a fuller report is produced, integrating the 360-degree feedback. For example a bar graph presents the differences in Team Role preferences between an individual’s own perception and that provided by colleagues. Any discrepancies between self-perception and observer input can lead to fruitful discussion as to how the individual is perceived in the team, helping to unearth discomfort with a current role, or to discover any conflict between the individual’s existing job role and the role they would like to play.

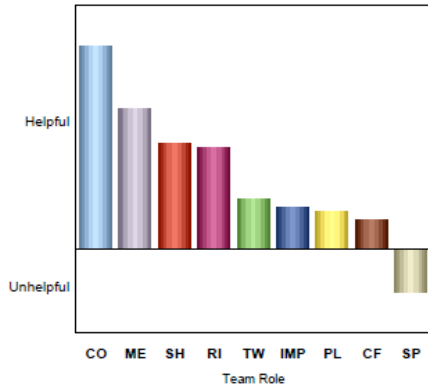


The use of Belbin reports promotes discussion as colleagues can be encouraged towards open discussion of the behaviours they display or observe. Since behaviour is evidential, the claims the report makes about an individual can be corroborated or refuted with real, everyday examples. Whilst encouraging openness and honesty, the language of Team Roles helps to diffuse conflict by using constructive, non-confrontational language to explore strengths and weaknesses.

ii) Job Profiling

It is increasingly difficult for an employer to judge an individual's potential success in a job from qualifications alone. Whether for recruitment or development purposes, it is important to ensure a good match between someone's Team Role composition and the Team Role requirements of a job.

This report is based on A Person's Job Requirements Inventory and 4 Job Observations.



The job as it is specified requires someone who has the capacity for drawing the best out of others, encouraging them to contribute and to identify personally with group objectives. Such a person should be able to conduct meetings in a mature fashion and in a way that allows participants to feel satisfied when proceedings have been concluded.

The work entails a fair degree of planning and thinking combined with level-headed judgement.

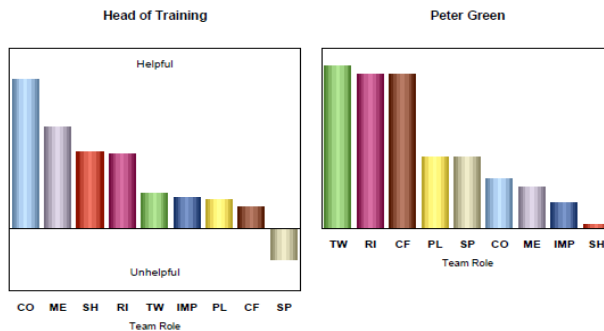
On this specification general managerial ability is more important than the possession of any particular knowledge or experience.

Belbin identifies the difference between eligibility and suitability. Eligibility refers to past experience, e.g. qualifications, training, cultivated skills. Suitability refers to future potential: the degree of fit between an individual's behavioural characteristics and the job he or she does.

In addition to assessing individual's Team Roles, Belbin can also be used to analyse a job, in other words, to define a job in terms of Team Roles. The person responsible for the job (for example, the line manager) is asked to complete the Job Requirements Inventory, which Investigates the various characteristics which are required for the job.

Job observations can also be added so that those who come into contact with, or have a close understanding of, the job concerned (e.g. colleagues, managers, clients) can contribute to defining the role.

The job definition can then be directly compared to an individual's Self-Perception responses and Observer Assessments to judge an individual's suitability and compatibility for the job in question.



Peter has a fair amount of aptitude for liaison work and a readiness to work with others to achieve group objectives.

He may lack the dispassionate analytical qualities needed for this type of work, but this reservation is less important if you have independent test material that shows he has a good critical thinking ability.

He is unlikely to have the tough-mindedness needed in certain aspects of the job.

Comparing Peter's Observer Assessments with Job Observations:

Peter is seen as having certain behavioural qualities that are useful to this job. These are being seized opportunities, and encouraging of others. However, he is not observed as being broad in outlook, conscious of priorities, challenging, or persuasive, which would have been helpful.

In any situation where matching an individual to a jobs is important, this report can be a useful tool to indicate potential success in the job. For the purposes of individual review and development, this process can help to establish where there might be discrepancies between a manager's expectations and an individual's potential, by addressing the differences through the language of Team Roles:

“ Belbin's theories on their own, are not a panacea to [sic] the industry's recruitment, management and leadership problems. However, they are a readily available and excellent aid to preventing disastrous recruitment and selection mistakes. This in turn prevents adding further to the pool of autocratic managers in our industry

~ Helen Bennett, 2001¹⁸ ”

There is no Team Role “formula” for a given functional role (e.g. marketing manager or accountant) since this may change from one organisational culture to another. However, the use of Observer Assessments and job profiling helps to anchor an individual's self-awareness in a specific working environment. For example, an individual starting a new job might learn about the kind of behaviour that job requires, whilst the manager may gain insight into how the candidate might fulfil that job given his or her Team Role propensities. The use of job reporting can provide manager and candidate alike with “behavioural expectations” as to what the job requires. It may be that someone's individual working style is a good fit with the job from the outset. Where there is not an immediate fit and the jobholder is very proficient in the role, it could be that the job boundaries and requirements are slightly different than the manager perceives, or indeed, that the individual has adapted to the requirements of the role. Job reporting can be used to provide gap analysis and as a focal point for discussion.

In *Management Teams: Why They Succeed or Fail*, Jana Krajcarova's case study describes how Belbin's individual and job analysis was used to resolve conflict between a quality control manager and the CEO. Whilst self-perception reports allowed the individuals concerned to appreciate their two distinct approaches, a Team Role definition of the job identified the real source of the conflict: the fact that the CEO envisaged the quality control manager role as requiring Co-ordinator and Shaper characteristics, whilst the present incumbent was a strong Monitor-Evaluator. The process of defining the job led towards a new job specification and a personal development plan for the manager. The author comments, “...we managed to solve this sensitive problem not only without any personal frustrations and animosity, but also with a significant increase in the motivation of both managers”. For more information on Meredith Belbin's book, *Management Teams: Why They Succeed or Fail*, please visit: <http://www.belbin.com/rte.asp?id=28>.

Please note that Belbin can be used as a recruitment and selection tool. Team Roles provide an insight into working and team relationships, but the reports should not be used as the sole basis for making recruitment decisions.

¹⁸*Finding round pegs for round holes* (Helen Bennett, The Hospitality Review, July 2001, p.19).

Belbin in practice

There is much evidence of the use of Team Role theory in general to promote self- and mutual understanding, more effective management and even significant culture change, which, in turn, translates into higher performance and evidential commercial and organisational success.

More information on the application of Team Role theory with positive results can be found in the case studies in Meredith Belbin's book, *Management Teams: Why They Succeed or Fail*. From tackling language barriers in international firms, to addressing change strategy, to personal development of young people, these case studies contain actual scenarios and outcomes from using the Belbin model.

All quotations below are taken from the case studies mentioned above. Please visit our website at <http://www.belbin.com/rte.asp?id=28> for more details and to order *Management Teams: Why They Succeed or Fail*.

“

Here Belbin's methods played a significant role in revealing the presence or absence of certain attributes and increasing team members' awareness of each other. [...] This change has without exception resulted in better teams and better management.

”

~ Asbjorn Aanesen, Linjegods AS

Using Team Roles in change strategy, one organisation credits their use of Belbin with:

“

The creation and delivery of a credible and robust strategy, which delivered 55 per cent profit growth in less than three years” and “The development of a rich, diverse team culture, which developed its skills in the delivery of a high-quality service and brands in the consumer marketing arena.

~ Paul Wielgus, Social Innovation UK and Chris Zanetti, Allied Domecq, Germany

”

Commenting on the use of Interplace in a school, Delphine Rushton comments:

“

Whether it be a young teacher considering their first move from classroom operator into a managerial post, or a senior manager wishing to pursue further leadership roles, it provides feedback on team strengths and career direction. Unconnected with assessment of performance linked with pay, it helps staff evaluate their current responsibilities and contributions. Staff have not had access to this kind of feedback before and they have found it to be both illuminating and productive.

”

~ Delphine Rushton, Richmond School

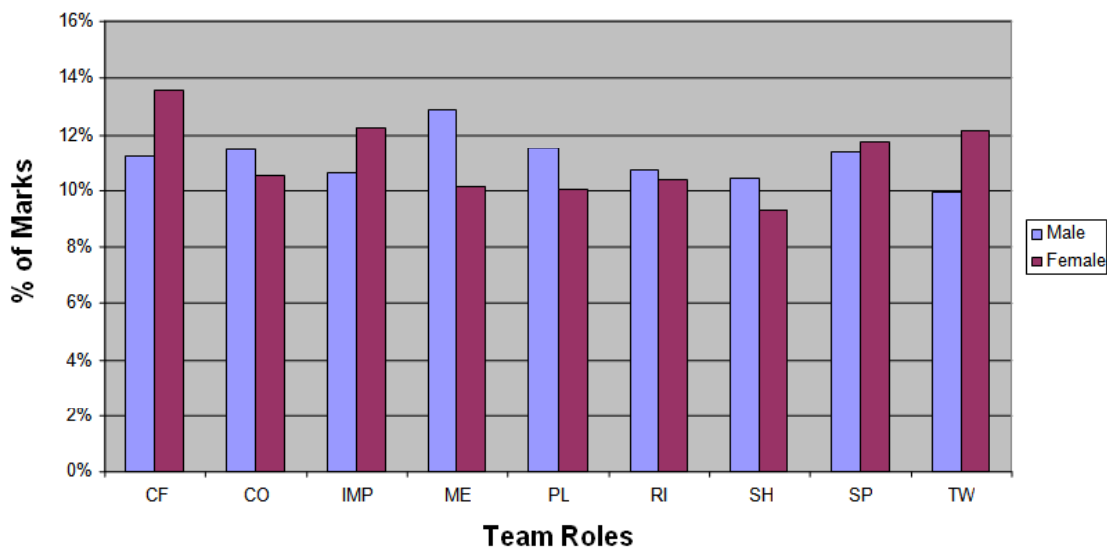
Testing, Bias and Fairness

In testing, it is important to ensure that there are not significant differences in test scores for certain groups. A test can be said to be biased, for example, if females consistently scored significantly higher than did males for a particular trait; or if members of different ethnic groups felt discouraged to answer certain items for cultural reasons.

Belbin and Gender

The Belbin Interplace system collects information regarding the gender of respondents in order to script reports correctly. This information is not used in any part of the calculation process and does not affect the data collected.

Belbin has conducted analysis using a database of more than 20,000 respondents, whose data was processed through the Belbin Interplace system. With this large dataset and up-to-date, valid version of the BTRSPI, the effect of Type I and Type II errors were significantly reduced (for more information on Type I and Type II errors, please see the “Further Reading” section below).



The graph demonstrates that there is very little gender difference for any Team Role (in this case, only Self-Perception data was used). The greatest difference (2.6%) between male and female respondents was found for Monitor Evaluator (ME) and even this figure represents only a small difference of 1.8 “ME marks” per SPI.

There are two published research studies which have been conducted concerning Belbin and gender, with varied results. However, it should be noted that both studies used considerably smaller sample sizes of 185 and 390 respondents respectively. (For more information on the statistical errors caused by using a small or inappropriate sample size, please see the “Further Reading” section below.) Moreover, the findings of both studies were compromised by the use of the obsolete self-scoring BTRSPI to produce results. (For more information on why this version should **not** be used, please see the “Administration and Scoring” section above).

In 1996, Balderson & Broderick used clusters of Team Roles and reported that there was no statistically significant difference between membership of clusters for men and women. However, they also reported that, in a separate analysis, ME and PL were found to discriminate between gender, with women scoring higher on both Team Roles than men.¹⁹ In turn, Anderson & Sleaf's 2004 study claimed to find a bias in favour of CO and SH for men, and TW for women. However, the researchers admit that their results are far from definitive.²⁰

Given that the results of the two studies appear to contradict one another, and in view of Belbin's findings from a dataset over 100 times larger, Belbin concludes that there is no significant gender bias for any Team Role and that the two research studies demonstrate natural variation between sample groups, highlighting only those individual differences which Belbin celebrates.

Belbin and Ethnic Origin

Belbin Team Roles is an international language, used all over the world. The BTRSPI has been translated into many different languages by professionals familiar with the Belbin Team Role theory, who take great care to ensure that the nuances of language in the BTRSPI are captured in different languages. For more details on our agents and distributors outside the UK, please visit <http://www.belbin.com/rte.asp?id=93>. The Belbin Interplace system allows an individual user to create norms which reflect Team Role preferences for a given culture, whether organisational or national. For more information on norming, please see the "Norming" section on page 8.

Belbin and Age

To date, no data have been gathered regarding age and respondents, since this is not required by Interplace for the profiling process. Whilst there are no known differences in terms of preferred Team Roles, it is sometimes observed that those who have been in the work environment for a longer period of time are likely to have more defined or polarised Team Role preferences, since a longer experience of a working environment tends to increase awareness of one's strengths and weaknesses.

If an individual is less sure of their contribution and their Team Roles do not seem so well-defined, it is best practice to add Observer Assessments. As well as overcoming the limitations of self-reporting, this gives someone new to the work environment some hints as to where their colleagues acknowledge their strengths and talents to lie, thereby maximising the opportunity for self-development and cultivation of their stronger Team Roles. For more information on the limitations of self-reporting and the advantages of adding Observer Assessments, please see the "Concordant Validity" section on page 14.

¹⁹*Behaviour in Teams: Exploring Occupational and Gender Differences* (Balderson & Broderick, 1996).

²⁰*An Evaluation of Gender Differences on the Belbin Team Role Self-Perception Inventory* (Anderson & Sleaf, 2004).

Belbin and Management

The original research conducted by Meredith Belbin in the 1970s was conducted with a sample of top managers. However, research has demonstrated that the BTRSPI can be used at all operational levels. **In other words, Team Roles are relevant for anyone in the work environment who interacts with others and wishes to learn more about his or her own behaviour, and that of colleagues and managers.** The Observer Assessment encourages feedback throughout the strata of an organisation and is intended to provide a universal language to address difficult issues which might otherwise cause conflict within teams.

Having examined the distribution of Team Roles among UK Managers in their 1998 study, Fisher, Hunter & Macrosson went on to investigate the use of Belbin for non-managers. They employed a team working exercise in their experiment, concluding that Belbin Team Role theory could be equally applicable to non-managerial as well as managerial teams.²¹ Please see the “Further Reading” section for more information regarding Belbin and management.

²¹*Belbin’s Team Role Theory: For Non-Managers Also?* (Fisher, Hunter & Macrosson, 2001).

Research Studies

Owing to its popular acclaim and success in the workplace, the BTRSPI has been the subject of much research from 1993 to the present day, claiming the interest of psychologists and academics, as well as those working with teams. However, research into the BTRSPI has been characterised by a number of erroneous practices:

The BTRSPI measures behaviour and produces output in terms of Team Roles. Team Roles are not personality traits, but clusters of behaviour. Since the BTRSPI is a behavioural, not a psychometric, analysis, research into its psychometric properties produced variable results. Although 23 out of 32 studies showed positive evidence supporting the BTRSPI, only 4 out of 9 studies on the psychometrical properties of the inventory produced supportive evidence, because the inventory was being tested for characteristics it was not intended to possess.

Many studies have been conducted using the now obsolete, self-scoring version of the BTRSPI, which is missing a Team Role, is not properly normed, is not substantiated by Observer Assessments and does not produce detailed feedback. For more information, please see the “Administration and Scoring” section above.

Some studies created, and experimented with, a normative version of the BTRSPI. Please note that Belbin owns the copyright to the BTRSPI and that reproduction and/or alteration of the BTRSPI is prohibited by Belbin. For more information, please visit <http://www.belbin.com/rte.asp?id=7>.

The majority of early studies were conducted with small or inappropriate sample sizes, which magnified the chance of Type I and Type II errors. These are statistical errors which are more likely to occur when the sample size is small:

In Type I errors (also called “rejecting null when null is true” or “false positive”), researchers mistakenly think that a statistical difference exists when, in truth, there is no statistical difference (in other words, the null hypothesis is true but was mistakenly rejected).

In Type 2 errors (also called “retaining null when null is false” or “false negative”), researchers fail to reject the null hypothesis, even though the alternative hypothesis is true.

When reading research studies conducted on the BTRSPI and analysing the findings, it is useful to bear these frequently encountered shortcomings in mind.

Interlace Research Articles

The following studies were conducted using Interlace data. The authors asked Belbin Associates for access to this data, which is fully-normed and consists of more than 5,000 records.

Stephen Swailes et al. are independent researchers based at the University of Hull. His studies demonstrate that the BTRSPI has good reliability and validity (please see the “Reliability” and “Validity” sections above) and which takes into account the unique ipsative and non-ipsative characteristics of the BTRSPI. The following articles in particular are recommended (in date order).

Aritzeta, Ayestaran & Swailes, *Team Role Preference and Conflict Management Styles* (2005)

Aritzeta, Senior & Swailes, *Belbin Team Role Preference and Cognitive Styles: A Convergent Validity Study* (2004)

Aritzeta, Senior & Swailes, *Belbin’s Team Role Model: Development, Validity and Applications for Team Building* (2007)

Aritzeta, Swailes & McIntyre-Bhatty, *Further Evidence on the Validity of the Belbin Team Role Self Perception Inventory and the Observer’s Assessment Sheet* (2004) *

Aritzeta, Swailes & Senior, *Team Roles: Psychometric Evidence, Construct Validity and Team Building* (2005)

Beck, Fisch & Bergander, *Functional Roles in Work Groups – An Empirical Approach to the Study of Group Role Diversity* (1999)

McIntyre-Bhatty & Swailes, *The Reliability of the (Belbin) Team Role Self-Perception Inventory: Cronbach’s alpha and ipsative scales* (2000)

Morison, Chris, *An Investigation of Belbin Team Roles as a Measure of Business Culture* (2008)

Swailes & Aritzeta, *Scale Properties of the Team Role Self-Perception Inventory* (2006)

Swailes & McIntyre-Bhatty, *Uses and Abuses of Reliability Estimates: The Case of the Belbin TRSPI* (2001)

van Dierendonck & Groen, *Belbin Revisited: The Construct Validity of the Interlace II Team Role Instrument* (2008)

* = Contains a comprehensive literature review of the research to date on the BTRSPI.

Other Articles

The following studies were conducted using the obsolete self-scoring version of the BTRSPI and without approaching Belbin Associates to request the use of a large dataset. Please note that evidence is largely affected by the use of a non-sanctioned BTRSPI and smaller datasets, as outlined above.

Anderson & Sleep, *An Evaluation of Gender Differences on the Belbin Team Role Self-Perception Inventory* (2004)

Arroba & Wedgwood-Oppenheim, *Do Senior Managers Differ in the Public and Private Sector?: An Examination of Team Role Preferences* (1994)

Balderson & Broderick, *Behaviour in Teams: Exploring Occupational and Gender Differences* (1996)

Blignaut & Venter, *Teamwork: can it equip university science students with more than rigid subject knowledge?* (1998)

Dulewicz, Victor, *A Validation of Belbin's Team Roles from 16PF and OPQ using Bosses' Ratings of Competence* (1995)

Dulewicz & Higgs, *Can emotional intelligence be measured and developed?* (1999)

Fisher, Hunter & Macrosson, *The Distribution of Belbin Team Roles Among UK Managers* (1998)

Fisher, Hunter & Macrosson, *Belbin's Team Role Theory: For Non-Managers Also?* (2001)

Fisher, Macrosson & Semple, *Control and Belbin's team roles* (2000)

Fisher, Macrosson & Sharp, *Further Evidence Concerning the Belbin Team Role Self-Perception Inventory* (1996)

Fisher, Macrosson & Wong, *Cognitive Style and Team Role Preference* (1998)

Henry & Stevens, *Using Belbin's leadership role to improve team effectiveness: An empirical investigation* (1999)

Lessem & Baruch, *Testing the SMT and Belbin inventories in top management teams* (1999)

Macrosson & Hemphill, *Machiavellianism in Belbin team roles* (2000)

Prichard & Stanton, *Testing Belbin's team role theory of effective groups* (1999)

Rushmer, *Is Belbin's shaper really TMS's thruster-organizer? An empirical investigation into the correspondence between the Belbin and TMS team role models* (1996)

Shi & Tang, *Team role behaviour and task environment: An exploratory study of five organizations and their managers* (1997)

Sommerville & Dalziel, *Project teambuilding – the applicability of Belbin's team-role self-perception inventory* (1998)