The British Psychological Society
Psychological Testing Centre

Test Review

PROPHET

Overview

Test Review of PROPHET

Reviewers: Stephen Fisher & Nicole Banerji

Consultant Editor: Linda Marshall

Senior Editor: Charles Eyre

PART 1 - DESCRIPTION OF THE INSTRUMENT

General Description

Test Name: PROPHET

Date of current review: April 2017

Date of previous review: n/a

Original test name: n/a

Authors of the original test: Colin Sheppard, Ben Stones, Stephanie Wyman

Authors of the local adaptation: n/a

Local test distributor/publisher: PROPHET Profiling

Publisher of the original version of the test: n/a

Date of publication of current revision/edition: July 2016

Date of publication of adaptation for local use: n/a

Date of publication of original test: 2014

ISBN: 9781854336729

General description of the instrument

PROPHET is a business-focused tool to be used for development purposes with employees especially those in executive roles. It looks at aspects of personal approach and style, and links these to the broad roles that should come easily to a person in terms of the business cycle. Not a traditional personality measure per se, it is based on business relationship theory and derives measures that link into the developer's model of the business cycle and the roles important within this.

PROPHET is an ipsative tool, presenting 112 statements in groups of four which load onto eight base scales of Motivation and Decision Making style. The four Motivation scales are: Driving, Influencing, Team and Order. The four Decision Making scales are: Inspired, Experimental, Systematic and Analytical. The computer generated output provides analysis and description as to where someone comes on each of these. It also provides interpretation for two sets of derived scales that come from these in terms of Inclinations and Roles. The Inclinations are 4 bi-polar scales of: People – Things, Managing – Leading, Present – Future and Structured – Flexible. The Roles are derived from the Inclinations and consist of 16 separate roles that have a distinct value in different parts of the business life cycle. The roles are: Entrepreneur, Evangelist, Marketer, Inventor, Architect, Advisor; Team Builder; Planner; Operations Director; Organiser; Co-ordinator; Operator; Trouble Shooter, Motivator, Change Manager and Fixer.

PROPHET is for use in development-focused interventions either with individuals or with teams to improve working relationships and company effectiveness mindful of the requirements of the business cycle. It is not designed to be used to make decisions about people as in selection, promotion, redundancy. Those who have experience and skills in using other Level 2 instruments are eligible to be trained to use PROPHET.

Classification

Content domains:

- Motivation
- Personality – Trait

Intended or main area(s) of use:

- Advice, guidance and career choice
- Work and Occupational

Description of the populations for which the test is intended

Adults at work, with specific relevance for managers, executives and leaders.

Number of scales and brief description of the variables) measured by the instrument

There are two sets of base scales (labelled Motivation and Decision style) with 4 scales in each set. These are labelled as follows:

Motivation (base scales)

- Driving
- Influencing
- Team
- Order

Decision Making Style (base scales)

- Inspired
- Experimental
- Analytical
- Systematic

There are then four pairs of 'inclinations' and sixteen 'role' scales. These are labelled as follows:

Inclinations (derived scales)

- Present – Future
- Flexible – Structured
- Leading – Managing
- Things – People

Roles (derived scales)

- Entrepreneur
- Evangelist
- Marketer
- Inventor
- Architect
- Advisor
- Team Builder
- Planner
- Operations Director
- Organiser
- Coordinator
- Operator
- Trouble Shooter
- Motivator
- Change Manager
- Fixer

Response mode

- Computerised

Demands on the test taker:

Manual capabilities

- necessary information given

Handedness

- Irrelevant / not necessary

Vision

- necessary information given

Hearing

- Irrelevant / not necessary

Command of test language

- necessary information given

Reading

- necessary information given

Writing

- Irrelevant / not necessary

Items format

- Multiple choice (mixed scale alternatives) Number of alternatives: 4

Ipsativity:

- Yes, multiple choice mixed scale alternatives resulting in partially or fully ipsative scores

Total number of test items and number of items per scale or subtest

There are 112 statements arranged into quadrads. This means that the test taker is presented with 28 quadrads where they are invited to choose between the four items the one that is Most and then Least like them. Hence there are 14 quadrads for each of the two sets of base scales.

Intended mode of use:

- Controlled mode: No direct human supervision of the assessment session is involved but the test is made available only to known test-takers. Internet tests will require test-takers to obtain a logon username and password. These often are designed to operate on a one-time-only basis.

Administration mode(s):

- Computerised web-based application – unsupervised/self-assessment

Time required for administering the instrument

Preparation: 5 minutes

Administration: 10 - 15 minutes

Scoring: Instant – scored by web-based system

Analysis: 10 minutes to review automated report

Feedback: 1 hour (recommended)

Indicate whether different forms of the instrument are available and which form(s) is (are) subject of this review: Only one form is available

1. Measurement and scoring

Scoring procedure for the test:

Computer scoring with direct entry of responses by test taker

Scores:

Within an item set, a point is added to the scale score for the item chosen as 'most like' and subtracted for the scale of the item chosen 'least like'. This means that the raw scores for the base scales range from -14 to +14. Pairs of base scales are then combined to create the 4 inclinations. The 16 PROPHET roles are derived from different combinations of scales.

Scales used:

Other:

The two **base scales** have a raw score range of -14 to +14 and are then standardised with mean of 0 and SD of 5.5

The derived **Inclinations** scales are calculated to provide the same standardised scale with a mean of 0 and SD of 5.5

The derived **Role** scores are standardised with a mean of 0 and SD of 22 and a range of -56 to +56.

Score transformation for standard scores:

Not-normalised – standard scores obtained by linear transformation

Computer- Generated Reports

Are computer generated reports available with the instrument?

Yes

Name or description of report: PROPHET Individual Profile	
Media	• Integrated text and graphics •
Complexity	• Medium (A mixture of simple descriptions and some configurations of scale scores, and scale interactions)
Report structure	• Construct based – where the report is built around one or more sets of constructs (e.g. in a work setting these could be such as team types, leadership styles, or tolerance to stress; in a clinical setting these could be different kinds of psychopathology; etc.) which are linked to the original scales
Sensitivity to context	• One version for all contexts
Clinical-actuarial	• Based on clinical judgment of group of experts
Modifiability	• Not modifiable (fixed print-only output)
Degree of finish	• Publication quality
Transparency	• Mixture of clear/concealed linkage between constructs, scores and text
Style and tone	• Guidance/suggests hypotheses
Intended recipients)	• Qualified test users • Test takers • Third parties
Do distributors offer a service to modify and/or develop customised computerised reports?	• No

Supply Conditions and Costs

Documentation provided by the distributor as part of the test package:

User Manual

Technical (psychometric) manual

Methods of publication:

Paper

Internet download

Start – up costs: £1500 - for 2-day accreditation which includes the User Guide and Technical Manual

Recurrent costs: £150 per online administration and report

Prices for reports generated by user installed software: n/a

Prices for reports generated by postal/fax bureau service: n/a

Prices for reports by internet service: n/a

Prices for other bureau services: correcting or developing automatic reports: n/a

Test-related qualifications required by supplier of the test

Test specific accreditation

Accreditation in general personality and assessment: measures of typical behaviour, attitudes and preferences (equivalent to EFPA Level 2)

Professional qualifications required for use of the instrument:

EFPA Test User Qualification Level 2 or national equivalent

PART 2 - EVALUATION OF THE INSTRUMENT

Key to symbols:

[n/a]	This attribute is not applicable to this instrument
0	Not possible to rate as no, or insufficient information is provided
★	Inadequate
★★	Adequate
★★★	Good
★★★★	Excellent

Quality of the explanation of the rationale, the presentation and the information provided

Quality of the explanation of the rationale

Overall rating of the quality of the explanation of the rationale ★★

Theoretical foundation of the constructs	★★
Test development (and/or translation or adaption) procedure	★★
Thoroughness of the item analyses and item analysis model	0
Presentation of content validity	★★
Summary of relevant research	★★

Adequacy of documentation available to the user (user and technical manuals, norm supplements, etc.)

Overall adequacy of documentation available to the user (user and technical manuals, norm supplements, etc.) ★★

Rationale	★★
Development	★
Development of the test through translation/adaption	n/a
Standardisation	★★★
Norms	★★
Reliability	★★
Construct validity	★★
Criterion validity	★★
Computer generated reports	★★★★

Quality of the procedural instructions provided for the user

Overall adequacy ★★★

For test administration	★★★
For test scoring	n/a
For norming	n/a
For interpretation and reporting	★★★★
For providing feedback and debriefing test takers and others	★★★★
For providing good practice issues on fairness and bias	★
Restrictions on use	★★★
Software and technical support	★★★
References and supporting material	★★★
Quality of the procedural instructions provided for the user	★★★

Reviewers' comments on the documentation

There are two main components of the PROPHET documentation; a User Guide and a Technical Manual. There is also a catalogue which details product and product training charges.

The User Guide provides guidance concerning the theoretical basis of the instrument and its use and interpretation. The scales are well described in detail with comprehensive advice on their interpretation. This is very important, as PROPHET is fundamentally a coaching or development tool and the quality of the subsequent development conversation is paramount.

There is a section on feedback which provides the reader with an excellent guide to disseminating the findings to test-takers. A feature of this section is the use of historical figures as exemplars of each of the sixteen 'Roles'. This makes for interesting reading but some might question the validity of this as we, for example, don't know how Baron Haussmann scored. It also covers likely responses on a good day, on a bad day, as a team leader, as a team member and as a stakeholder. Suggestions for working with those who are either high or low on each aspect along with feedback prompts and questions are provided. Some actuarial evidence is included: any gender or age differences on each scale are pulled out, as is any significant correlation with other scales and where caution should be used in interpreting specific scales.

However it is important to remember that this interpretation is founded primarily on the expert judgement of the authors. A fuller explanation of the theoretical basis of the instrument is missing and in particular why these scales were identified as the most relevant to assess. To some degree, the instrument has evolved through use rather than being designed initially with psychometric principles in mind, refining the instrument using these principles has come later. A possible consequence is a lack of focus on the meaning of the scales and the items that compose them. Certainly little information is provided about how the scales were identified, the criteria used to design the items and the item analyses that were carried out. The absence of this information does undermine confidence in the content validity of PROPHET.

The Technical Manual is well structured and it presents a lot of supporting information for a relatively new tool in its current format. Psychometric concepts are explained well, and the tables of statistical information, off-putting for some users, have useful explanations with them to aid understanding. Data is well presented and explained in detail.

There is currently insufficient discussion of the use of an ipsative response format. Apart from the mathematical consequence of negative correlations between scores, no further reference is made to the challenges inherent in this form of response format of which users qualified to EFPA Level 2 will be aware. This is a controversial topic but not one that can be ignored in evaluating the technical qualities of PROPHET. In terms of its development questions arise as to the social desirability of the items and whether this was taken into account when creating each quadrant of four items. Also there needs to be a rationale to justify the number of quadrants needed to establish an accurate preference hierarchy. In terms of the use of norms, clarity is required to explain the purpose of the comparison of individual preferences with those of groups and how such comparisons can add information for interpreting preference scores (Hogan, 2003). It has also been argued that the use of internal consistency is not appropriate with ipsative data as the assumption of consistent coding is not met (Brown, 2015) but again this topic is not addressed. Finally there are questions about how to evaluate evidence of validity which uses correlations as part of the analysis because the nature of ipsative data leads to spurious positive and negative correlations (Brown, 2015).

The ratings given to the documentation in this review try to reflect these issues. At one level the information provided for PROPHET is well presented and clear however the lack of information and discussion on important topics reduce its effectiveness for the user.

Quality of the test materials

Quality of the test materials of CBT and WBT

Quality of the design of the software (e.g. robustness in relation to operation when incorrect keys are pressed, internet connections fail etc.)	★★★★
Ease with which the test taker can understand the task	★★★★
Clarity and comprehensiveness of the instructions (including sample items and practice trials) for the test taker, the operation of the software and how to respond if the test is administered by computer	★★★★
Ease with which responses or answers can be made by the test taker	★★★★
Quality of the design of the user interface	★★★★
Security of the test against unauthorized access to items or to answers	0
Quality of the formulation of the items and clarity of graphical content in the case of non-verbal items	★★★★
Quality of the materials of CBT and WBT	★★★★

Reviewers' comments on quality of the materials

The instructions on how to complete the questionnaire are clear and easy to understand. All the language used is simple, everyday and conversational, which reinforces the developmental nature of this instrument and reduces the risk of the completer thinking it is a test with right and wrong answers. That the instrument should not be used to make decisions about the person e.g. selection or promotion is stated clearly, giving transparency and further comfort that the tool is purely for development purposes.

The questionnaire is easy to complete. The items are presented in sets of four and the participant has to choose which statement is most like them, and then which is least like them on a duplicate list placed beside the original. The buttons to click on are large, and the completer can pause and save their work coming back to it at a later time. They can also navigate backwards and forwards through the questionnaire to check and change their answers if necessary. On finishing the questionnaire the completer receives an email message to this effect to give reassurance that the technology worked and that their data is being processed.

Norms

Is the test norm referenced? Yes

Norm referenced interpretation

Overall Adequacy: 

Appropriateness for local use	
Appropriateness for intended applications	
Sample sizes (classical norming)	
Sample sizes continuous norming	n/a
Procedures used in sample selection	Incidental
Representativeness of the norm sample(s)	
Quality of information provided about minority/protected group differences, effects of age, gender etc.	
How old are the normative studies?	Between 2013-16
Practice effects	n/a

Is the test criterion referenced? **No**

Reviewers' comments on the norms

The standardisation was based on a large sample over 1200 respondents from a range of sectors and job roles who completed PROPHET between 2013 and 2015. The purpose of standardisation was 'to ensure the equivalence of meaning of the scores on different scales are equivalent' (p17 Technical Manual). The scores are then interpreted using a nine point scale running from a very low preference, through a balanced preference to a very strong preference.

The standardisation data for the base scales is also used to explore group differences in terms of gender, age and ethnicity. In line with other research small differences were found on a number of dimensions between genders, as men had slightly higher average preference scores for Driving and Inspired, and women had higher average preference scores for Team and Systematic. These differences seem to remain with seniority taken into account. There was no evidence that Age influenced scores. Small to moderate differences were found for Ethnicity although these were confounded by factors such as seniority and job role.

The norm sample is regularly updated and its new composition compared with the standardisation sample and reported as an appendix in the Technical manual.

The use of norms with ipsative measures can be viewed as controversial because they are referenced to the individual whereas normative scores are referenced to the norm group. However there is a history of ipsative measures using norm tables to aid interpretation of scores. (Hogan, 2003) Certainly the account of the scores in the computer generated report is helpful in this regard. Under the heading 'your preferences benchmarked' the text explains 'you are able to see how your preference compares to a 'typical' executive, taken from a large norm population of people in professional roles'. This makes it clear that it is preferences rather than the degree or amount of a characteristic that is being described. It is also very important for the user to keep this in mind and not to misinterpret the meaning of the scales.

Reliability

Overall Adequacy: 

Overall Adequacy	
Data provided about reliability	Standard error of measurement given for a number of different groups (for each scale or subscale)
Internal consistency:	
Sample size	n=341
Kind of coefficients reported (select as many as applicable)	➢ Coefficient alpha or KR-20
Size of coefficients	
Reliability coefficients are reported with samples which…..	…. match the intended test takers
Test related reliability-temporal stability:	
Sample size	Two samples used one with 110 respondents, the other with 180 respondents 
Size of coefficients	Median values for the base scales: 2 week interval 0.86; 6 month interval 0.69.
Data provided about test-re-test interval	There are two studies - one with a 2-week interval (n=110) and another with a 6-month interval (n=180)
Reliability coefficients are reported with samples which	…. match the intended test takers
Equivalence reliability:	
Sample size	Not applicable
Are the assumptions for parallelism met for the different versions of the test for which equivalence reliability is investigated?	Not applicable
Size of coefficients	Not applicable

Reliability coefficients are reported with samples which….	Not applicable
IRT based method:	
Sample size	Not applicable
Kind of coefficients reported (select as many as applicable)	Not applicable
Size of coefficients (based on the final test length)	Not applicable
Inter-rater reliability:	
Sample size	Not applicable
Kind of coefficients reported (select as many as applicable)	Not applicable
Size of coefficients	Not applicable
Other methods of reliability estimation:	
Sample size	Not applicable
Results	Not applicable

Reviewers' comments on reliability

Three aspects of reliability were described in the technical manual: internal consistency, short term test-retest and longer term test-retest stability. As an ipsative measure there are issues with putting too much weight on internal consistency data, although this was presented and was adequate.

Internal consistency (Cronbach's α) is reported as being above 0.70 for all the base scales. Inclination scale reliabilities are similar with only the 'Things' scale dropping below 0.70 at 0.68. Role scale internal consistencies are more variable, with 6 of the 16 scales (Inventor 0.34, Advisor 0.63, Fixer 0.63, Operations Director 0.67 and Marketer 0.69) falling below 0.70. Caution should be exercised when using these scales.

The 6 month test-retest data was adequate, demonstrating adequate stability of outcome over a reasonable time frame. However, as this is a development focused tool, one might like to hope that those completing it had embarked on a course of personal change and thus some change in response would be expected over this time frame.

Thus for this instrument the most useful aspect of reliability is likely to be the short term test-retest study. This was very strong with 28 of the 32 aspects (both base and derived scales) scoring above 0.8 and several above 0.9. This indicates that over a short time period people are responding in similar ways to the questionnaire and similar outcomes will be achieved.

It can be argued that a more appropriate measure of consistency for ipsative measures is the similarity of the profile shape (i.e. maintenance of the relative preference hierarchy over time).

The authors have derived an approach to address this by presenting a measure of 'allocation'. This quantifies the percentage of people who are allocated to the same inclination or role on both occasions (using the 2-week study of 110 respondents). Whilst these percentages appear high (71% to 91%) the categorisation is relatively crude (i.e. remain below or above the mid-point) which means that it will not be particularly sensitive.

Validity

Overall Adequacy: ★★

Construct validity:	
Design used (select as many as are applicable)	➢ (Corrected) item-test correlations ➢ Difference between groups ➢ Correlations with other instruments and performance criteria
Do the results of (exploratory or confirmatory) factor analysis support the structure of the test?	0
Do the items correlate sufficiently well with the (sub) test score?	★★★
Is the factor structure invariant across groups and/or is the test free of item-bias (DIF)?	0
Are the differences in mean scores between relevant groups as expected?	★★★
Median and range of the correlations between the test and tests measuring similar constructs	★★
Do the correlations with other instruments show good discriminant validity with respect to constructs that the test is not supposed to measure?	★★
If a Multi-Trait-Method design is used, do the results support the construct validity of the test (does it really measure what it is supposed to measure and not something else)?	0
Other, e.g. IRT-methodology, (quasi-) experimental designs (describe):	0
Sample sizes	With NEO n=135; with MBTI n=113; with WAVE n=135.
Quality of instruments as criteria or markers	★★★
How are old are validity studies?	Studies between 2013 -2015. Number of years: up to 4 years at time of the review

Construct validity – Overall adequacy	★★
Criterion – related validity:	
Type of criterion study or studies	**Concurrent**
Sample sizes	n=195 and n=97
Quality of criterion measures	★★
Strength of the relation between test and criteria	★
Criterion – related validity – overall adequacy	★
How old are the validity studies	Studies carried out between 2013 - 2015. Number of years: up to 4 years at time of the review

Reviewers' comments on validity

For a relatively new instrument there is an impressive range of validity evidence for PROPHET including construct as well as criterion-related evidence. What is harder to evaluate is the value of the evidence owing to the problems of interpreting correlations based on ipsative data. One important consideration is how the instrument will be used. It can be argued that instruments used purely for development purposes do not need as much evidence as those for decision making purposes however it still needs to be demonstrated that the instrument measures a real-life quality that is useful.

In terms of construct validity a number of studies are presented.

There are inter-correlation matrices of scales for Base, Inclination and Role measures using the 1218 standardisation sample. As the authors acknowledge, because of the ipsative nature of PROPHET and its relatively few scales, caution needs to be exercised in interpreting high positive and negative correlations which occur in forced choice questionnaires, nevertheless correlations are usually in the predicted direction.

Work has also been carried out to demonstrate construct validity by job function, the hypothesis being that people with particular preferences are drawn to, are selected into and succeed more readily in different types of role. Those in different job functions in the standardisation sample were separated out to see if their profiles, en masse, were significantly different to those in other job functions in a predictable way. Most hypotheses were confirmed to a reasonable degree. A similar study was also carried out in respect to seniority, the hypothesis being that 'driving', 'influencing' and 'inspired' would be more apparent in a senior population than in a junior one. The hypotheses were confirmed.

PROPHET is compared to three other instruments widely used in an occupational setting: Saville Wave, MBTI and the Five Factor Model. Hypotheses were generated as to points of congruence between instruments and in many instances these were as anticipated and reached statistical significance although not the 0.55 level required by the EFPA standards for

congruent constructs. Although there are similarities between some of the scales across the various instruments they are not the same and thus strong relationships would not be expected. The rationale for linking these instruments to PROPHET is not made explicit, and further expansion on this point would help in interpreting these findings.

In more detail firstly, a questionnaire designed to measure the Big Five Model (a public domain 120 item questionnaire known as the IPIP-NEO) was administered to 135 employed individuals who also completed PROPHET. The authors predicted six 'expected' correlations with four of the Big Five Factors while 5 of the 6 show statistically significant results, results are also provided with Big Five Factors at the more detailed facet level.

A sample of 113 employed people completed the MBTI and PROPHET. The authors make a number of hypotheses which are then largely confirmed. It could be argued that the 16 role scores mirror the MBTI 16 types but this link is not made explicit.

The justification for using Saville Wave would appear to be that both are business oriented tests. The results of a study of 135 employed individuals conducted in 2015 supports the relationships hypothesised by the authors, with additional negative correlations explained as a consequence of the ipsative scoring used by PROPHET.

Two concurrent validity studies are presented. The first is of a sample of 193 area sales managers using ‘pounds’ and ‘sales rank’ criteria with mixed results, with the authors admitting there is only ‘weak support’ for PROPHET as a predictor of performance in this context.

The second smaller study of 98 professional staff from a large firm of architects used managers’ ratings of competency, based on the Great Eight Competency model (Kurz & Bartram, 2002). Some hypotheses were confirmed, others were not, and other non-hypothesised correlations were found. Again the evidence was inconclusive. This is not surprising, as PROPHET does not attempt to measure competence or self assessment of competence (unlike some personality measures that frame questions in terms of ‘how good’ someone thinks they are). It only measures style and approach e.g. in measuring decision making style (inspired, experimental, systematic or analytical) there is no attempt to get people to assess their decisiveness or their accuracy of decision making. The authors treat the results with caution suggesting more research will be needed before the relationship between PROPHET and competency is clarified.

Quality of computer generated reports

Overall adequacy of computer generated reports: 

PROPHET Individual Profile

Scope or coverage	
Reliability	
Relevance or validity	
Fairness, or freedom from systematic bias	
Acceptability	
Length	

Reviewers' comments on computer generated reports:

The computer generated reports are well presented with easily understood narrative and graphics giving a wealth of information likely to be of use in development and coaching contexts. An individual's report introduces the test-taker to the underlying principles of the test and goes on to give a detailed account of the scales and introduces the idea of the 'Heat Map' to present preferences.

Each section of the report is well laid out covering general information, a diagrammatic representation of the person's preferences, and a page of well written text that has the feel of an expert interpretation rather than a mechanistic repetition of scale descriptors. The text is accessible and conversational. It covers aspects that are of direct interest to people in their working situations and is highly suited to the target population.

It is part of the terms of supply that the respondent will not receive their report directly and they receive it only after they have had a feedback session with an accredited user, to go through the report and explore the output. As such there is space in the report for the respondent to make notes. In addition recommendations and questions to consider are included which are likely to be of use in feedback and coaching. An appendix to the report gives further descriptions of all 16 PROPHET roles.

Although the manual does not make specific reference to the algorithms for generating reports, it is clear how the blocks of text have been chosen.

Final Evaluation

Evaluative report of the test:

PROPHET has been designed to be a tool to aid the personal development of people in the world of work, particularly those at senior and executive levels. It is not being postulated as a tool for making decisions about people. As such it has been developed in a highly professional way, acknowledging the importance of psychometrics and actuarial evidence to support measurement, unlike many development tools that see themselves as purely as an invitation to have a conversation.

More importantly, significant efforts have been made by the test publishers to ensure that PROPHET is used as it has been intended. Training is compulsory, and is focused on those who not only have experience and skill in using other Level 2 instruments, but also who are skilled in coaching, feedback or organisational development.

PROPHET is a test which is aimed at business leaders wishing to gain insight into their senior staff. It is a business focussed profiling tool based on business relationship theory (BRT). Despite this foundation in BRT, only one page is devoted to the theoretical background in the 'User Guide'. A much-expanded explanation of, and justification for, the use of the theory could be important in fully informing a potential user as to the purpose of the test. More information to link the measures yielded by PROPHET and theory is also needed. In summary, users may be more reassured in their use of the instrument if firmer theoretical foundations were apparent.

The instrument is supported by an excellent user manual that gives comprehensive, and sophisticated, advice and guidance on interpretation. In virtually all the documentation, for users and those who complete the questionnaire, clear messages are given as to the developmental nature of PROPHET and how it should and should not be used.

There is an impressive range of technical data for a relatively new instrument, often based on quite large samples, which is well presented and explained. Unfortunately, the controversial issues that arise from the use of ipsative scoring are not discussed, and evaluating the evidence presented is consequently problematic. The following evaluations are necessarily tentative. The standardised score data is useful for interpretation and the intention to update it regularly as norms is encouraging. The reliability data is more problematic with most weight falling on the short term test retest results, which reach an acceptable level. The validity data is extensive and suggestive that the instrument is able to measure the characteristics claimed, although there is less support for their direct relevance in the workplace.

With regards to the test-taker and user experience, the web based administration and profiling is straightforward and the output very well presented and easy to interpret. The quality of the profiles is one of the strong points of PROPHET. There is a balance between narrative, graphical and numerical data, which although comprehensive is easy to understand and even interesting to read. The wealth of information available is likely to be of great interest to those wishing to gain an insight into individuals in their teams. As with any such instrument, there is a concern that leaders could misuse the information and use it for placement, redundancy or termination purposes, but the authors make it very clear that this is not the intended use and they insist on appropriate training which is as much as can be expected to mitigate this issue.

Conclusions:

In summary PROPHET is an interesting model, well presented and easy to use. It is likely to add value in the development contexts it has targeted. For a relatively new instrument it already has an extensive evidence base which could usefully be extended by addressing the measurement issues inherent in ipsativity.

Recommendations:

Suitable for use in the area(s) of application defined by the distributor, by test users who meet the distributor's specific qualification requirements (at least EFPA User Qualification Level 2)

Bibliography:

Brown, A. (2015) Personality Assessment, Forced-Choice. In: Wright, J.D., ed. International Encyclopedia of the Social and Behavioural Sciences, 2nd Edition. Elsevier.

Hogan, T.P. (2003) Psychological Testing: A Practical Introduction. New York: J.Wiley, Inc.

Kurz, R. & Bartram, D. (2002) Competency and Individual Performance: Modelling the world of work. In Robertson, I.T., Callinan, M. & Bartram, D. (Eds.) Organisational effectiveness: The Role of Psychology. Wiley. pp 227-258.