

Northumbria Research Link

Citation: Önkal, Dilek, Gonul, Sinan and De Baets, Shari (2019) Trusting Forecasts. Futures and Foresight Science, 1 (3-4). e19. ISSN 2573-5152

Published by: Wiley-Blackwell

URL: <https://doi.org/10.1002/ffo2.19> <<https://doi.org/10.1002/ffo2.19>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/39381/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



Northumbria
University
NEWCASTLE



UniversityLibrary



TRUSTING FORECASTS

Journal:	<i>Futures & Foresight Science</i>
Manuscript ID	FFS-19-0001.R1
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Onkal, Dilek; Northumbria University, Marketing, Operations and Systems Gonul, Sinan; Northumbria University, Marketing, Operations and Systems De Baets, Shari; Universiteit Gent, Department of Economics and Business Administration
Keywords:	trust, forecast, judgment, algorithm aversion

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

TRUSTING FORECASTS

Dilek Önkal^a, M. Sinan Gönül^a and Shari De Baets^b

^aNewcastle Business School, Northumbria University, Newcastle, UK

^bFaculty of Economics and Business Administration, Ghent University, Ghent, Belgium

Running title: Trusting forecasts

TRUSTING FORECASTS

Abstract

Accurate forecasting is necessary to remain competitive in today's business environment. Forecast support systems are designed to aid forecasters in achieving high accuracy. However, studies have shown that people are distrustful of automated forecasters. This has recently been dubbed 'algorithm aversion'. In this study, we explore the relationship between trust and forecasts, and if trust can be boosted in order to achieve a higher acceptance rate of system forecasts and lessen the occurrence of damaging adjustments. In a survey with 134 executives, we ask them to rate the determinants of trust in forecasts, what trust in forecasting means to them and how trust in forecasts can be increased. The findings point to four main factors that play a role in trusting forecasts: (1) the forecast bundle, (2) forecaster competence, (3) combination of forecasts, and (4) knowledge. Implications of these factors for designing effective forecast support and future-focused management processes are discussed.

Keywords: trust, forecast, judgment, algorithm aversion

Word Count: 7592

Introduction

Accurate forecasting is indispensable for organisations to remain competitive. Forecast support systems are designed to aid decision makers in attaining acceptable forecasts. However, a plethora of studies have shown that people are sceptical of predictions handed to them by forecast support systems or by other forecasters (e.g., Dietvorst, Simmons, & Massey, 2015; Fildes, Goodwin, Lawrence, & Nikolopoulos, 2009; Önkal, Gönül, Goodwin, Thomson, & Öz, 2017; W. Wang & Benbasat, 2005). The literature on advice taking argues that such distrust may be viewed as a form of advice discounting. Forecasters overweigh their own judgment and continuously adjust forecasts provided to them. This judgmental adjustment of statistical forecasts is the prevalent practice in business forecasting (Fildes & Petropoulos, 2015). However, field studies have found that such adjustment is not always warranted: forecasters adjust the output they receive from statistical forecasting models ('model advice'), often when there is no reason for such adjustment and it in fact damages forecast accuracy (e.g. Baecke, De Baets, & Vanderheyden, 2017; Fildes et al., 2009; Van den Broeke, De Baets, Vereecke, Baecke, & Vanderheyden, 2018). One reason that has been put forward recently is referred to as 'algorithm aversion': people are averse to using advice from algorithms and are unforgiving towards any errors made by the algorithm (Dietvorst et al., 2015; Prah & Van Swol, 2017). While the combination of algorithms and human decision makers has the potential to lead to improved accuracy, in practice, the advice of the algorithm is often discounted by the human member of the human-algorithm dyad. One way to alleviate this so-called advice discounting, is to boost trust in the source of advice. Trust is a vital moderator in the relationship between the reliance on computer-assisted forecasts (Dzindolet, Pierce, Beck, & Dawe, 2002; Madhavan & Wiegmann, 2007) and the heuristics a forecaster uses in producing the final predictions (Alvarado-Valencia & Barrero, 2014). Increasing trust may therefore lead to more acceptance and to less damaging adjustments. However, how this enhancement should be achieved is not yet clear. Goodwin, Gönül, and Önkal (2013) lay the groundwork for this workstream and find that trust is influenced by the complexity of the time series

1
2
3 that are being forecast. We build further on their work and employ a survey to investigate explicitly
4 what forecasters deem important for trusting forecasts. In doing so, this paper presents the findings
5 from a trust survey and reveals four main clusters affecting trust: (1) the forecast bundle, (2) forecaster
6 competence, (3) combination, and (4) knowledge. This is supplemented by analysing the responses to
7 two qualitative questions focusing on (i) what trusting forecasts means to decision-makers, and (ii)
8 how it can be improved.
9

10
11
12
13
14
15
16 Importantly, literature has been looking for a way to stop forecasters from making harmful
17 adjustments to forecasts provided to them. However, directly intervening in the forecast support
18 systems has thus far been unsuccessful (e.g., Goodwin, Fildes, Lawrence, & Stephens, 2011). Indirectly,
19 enhancing trust may lead to a decline in the detrimental adjustments made by forecasters. The 'trust
20 in forecasting' survey used in this study can serve as a diagnostic tool for further research in this field.
21
22
23
24
25
26
27
28
29

30 **Literature**

31
32 Judgmental adjustments to model forecasts are the most common way of forecasting in
33 business practice (Fildes & Petropoulos, 2015). These adjustments are often unnecessary and
34 damaging to forecasting quality (e.g., Fildes et al., 2009). So why do forecasters adjust so often? A
35 number of explanations are possible. First, it has been suggested that people adjust because of a need
36 of a sense of ownership and/or a sense of control over the forecasting process (Önkal & Gönül, 2005).
37 Second, people may discern patterns in noise where there is none (Fildes et al., 2009). Third,
38 forecasters may alter the outcomes in order to continue receiving feedback. When tasks become
39 automated, people may feel like they are at risk of becoming de-skilled by not participating in the task
40 (Bainbridge, 1983). Fourth, changes are sometimes made because of political reasons or confusion
41 with target setting (Fildes & Hastings, 1994; Goodwin, 1996). Fifth, people suffer from algorithm
42 aversion: they are distrustful of algorithms and punish them severely for any error that is detected
43 (Dietvorst et al., 2015; Prahla & Van Swol, 2017). Yet, models can generate predictions based on the
44 logical and systematic processing of information and can handle large amounts of data (Goodwin &
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Wright, 2010). The advantage of models is that their predictions have a high degree of consistency and
4
5 generate fewer errors than human judgment (Blattberg & Hoch, 1990). Models can thus improve
6
7 forecasting performance by increasing the consistency of predictions (Hoch & Schkade, 1996).
8
9 Interestingly, in an experimental study concerning the use of decision support systems, Goodwin,
10
11 Fildes, Lawrence, and Nikolopoulos (2007) found that participants ignored the “advice” of the system
12
13 on which model to use (in the form of an “optimize” button). Only in 14.1% of the forecasts examined
14
15 was the optimize button used, and only in 9.2% of the total examined forecasts was the advised
16
17 method eventually chosen. Similarly, Lim and O’Connor (1995) found a tendency among forecasters to
18
19 persist with damaging adjustments in subsequent forecasts, despite the feedback that they were
20
21 reducing accuracy. People making forecasts are in general overconfident as well in the performance of
22
23 their forecasts (Arkes, 2001; Bovi, 2009; Lawrence, Goodwin, O’Connor, & Önkal, 2006) and suffer from
24
25 self-serving attribution bias, whereby they overestimate the importance of their own judgment when
26
27 making adjustments to statistical forecasts (Hilary & Hsu, 2011; Libby & Rennekamp, 2012). Decision
28
29 makers seem to discount advice from statistical forecasts (Önkal, Goodwin, Thomson, Gönul, &
30
31 Pollock, 2009). The advice literature shows us that they tend to discount advice relative to their own
32
33 judgment (Harvey & Fischer, 1997; Lim & O’Connor, 1995), even when it has the potential to improve
34
35 their accuracy (Yaniv, 2004a). Judges weight their own judgment more heavily, and incorporate advice
36
37 only to a limited degree: only 20-30% on average of the advice is taken into account (Harvey & Fischer,
38
39 1997).
40
41
42
43
44

45 Trust can play an important role in whether advice is fully accepted or discounted (Sniezek &
46
47 Van Swol, 2001; Wang & Du, 2018) and may thus have an effect on the acceptance of forecasting
48
49 algorithms in everyday business practice. Many definitions of trust exist, but in accordance with
50
51 previous studies on trust in forecasting (Goodwin et al., 2013), we focus on the definition of Johnson-
52
53 George and Swap (1982): “a psychological state comprising the intention to accept vulnerability based
54
55 on positive expectations of the intentions or behaviour of another” (p.395). Importantly, the definition
56
57 points towards vulnerability of the advice taker: taking advice has an inherent risk to it, as one needs
58
59
60

1
2
3 to rely on the benevolence, integrity and ability of the advice giver (Mayer, Davis, & Schoorman, 1995).
4
5 One could therefore state that trust is relevant in situations which have an inherent uncertainty to
6
7 them (Mayer et al., 1995; Snizek & Van Swol, 2001). Such is the case with forecasts: when predicting
8
9 the future, uncertainty is generally high. Accepting forecasts from another, be it another person or an
10
11 algorithm, requires the acceptance of this uncertainty and the trust that the other has a certain degree
12
13 of relevant expertise. Such trust seems to be especially difficult when receiving advice from an
14
15 algorithm (Dietvorst et al., 2015; Kleinmuntz, 1990; Lim & O'Connor, 1995; Önkal et al., 2009).
16
17 According to Alvarado-Valencia and Barrero (2014), trust is a moderator in the relationship between
18
19 heuristics and reliance, and affects the final forecast behaviour. Heuristics are the strategies used to
20
21 simplify the forecasting task into something more manageable, and reliance implies the degree of
22
23 acceptance of a computer-assisted forecast (i.e., will the forecaster adjust or not). Trust then affects
24
25 heuristics and reliance in such a way that it can lead to different adjustment behaviour of forecasts or
26
27 assignment of different weights to forecasting cues (Alvarado-Valencia & Barrero, 2014; Lee & See,
28
29 2004). Trust can be a personal trait, an appraisal of the credibility of the source or a consequence of
30
31 the context (Alvarado-Valencia & Barrero, 2014). While trust as a personality trait cannot be
32
33 influenced, the other two perspectives can. Source credibility for instance, can be influenced by
34
35 highlighting previous successes of the forecaster/computer in question, via visibility of the system or
36
37 comprehensible measures of accuracy (Dou, Walden, Lee, & Lee, 2012; Edwards, Spence, Gentile,
38
39 Edwards, & Edwards, 2013). These findings stem from human-automation interaction literature. More
40
41 specific to forecasting, this credibility has been deemed as one of the most important factors to
42
43 evaluate forecasts in a sales context by Mentzer and Kahn (1995). Additionally, forecasting literature
44
45 has shown that the provision of explanations may also increase trust in system forecasts (Gönül, Önkal,
46
47 & Lawrence, 2006; Önkal, Gönul, & Lawrence, 2008).

54 But what are the determinants of trust exactly? Human-automation literature suggests that
55
56 trusting a decision support system can be a complicated function of multiple variables, including
57
58 perception of the source, actual credibility, and perceived source credibility. Given a fundamental
59
60

1
2
3 understanding of decision support systems, decision makers filter their observations through their
4 perceptions and biases to arrive at a specific level of trust (Madhavan & Wiegmann, 2007). In the
5 forecasting field, a study by Goodwin et al. (2013) has shown that the complexity of the series that is
6 forecasted is an antecedent to trust. More specifically the level of noise is found to have an inverse
7 relationship with trust, and that the presence of trends decreased the level of trust. Equally important
8 was the formulation of the boundaries of the interval forecasts: an interval formulated as best-
9 case/worst-case led to higher trust than 'upper bound and lower bound' formulations. Thus, trust can
10 have an effect on the level acceptance of forecast advice in a number of ways. Given the continued
11 struggle of businesses to have their forecasters accept the advice of the forecasting software without
12 damaging adjustments, working on the level of trust may prove to be an important factor in
13 heightening this acceptance.

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28 The current study can be seen as a further attempt to explore the determinants of trust in
29 forecasting via a survey employed on forecasting practitioners. The details of the sample and the
30 procedure follows next:
31
32
33

34 35 36 37 **Method**

38 39 *Sample*

40
41 A survey was sent out to 138 people from the professional networks of the researchers. They
42 were all practicing forecasters with experience in making their professional forecasts as well as using
43 forecasts made by others. Participants worked in different companies across various sectors in Turkey
44 and there were no academics. One hundred and thirty-four complete surveys were returned, resulting
45 in a response rate of 97.10%.

46 47 48 49 50 51 52 *Design*

53
54 The survey questionnaire items were based on extensive discussions with executives and
55 managers participating in the Executive MBA programme at Bilkent University. These executives held
56 a minimum of five years of forecasting experience in order to be accepted to the programme.
57
58
59
60

Procedure

After their consent, participants were first asked to indicate the sector they worked in, their job experience (in years), how long they have been using other people's forecasts in their job (in years), how long they have been making their own forecasts (in years), how they would rate their own knowledge of forecasting techniques – a self-assessed rating of expertise (5-point scale from 1: None at all to 5: Very high/extensive), how much trust they put in statistical forecasts (5-point scale from 1: None at all to 5: Very high) and how much trust they put in expert's forecasts (5-point scale from 1: None at all to 5: Very high). The survey furthermore requested the participants to complete the following open-ended statements (1) trust in forecasts means the following to me and (2) the following is required to increase/boost/enhance my trust in forecasts. Next, participants were asked to rate the factors that affect their trust in a given forecast on a Likert scale ranging from 1 to 5, with 1 indicating "Not important at all" and 5 indicating "Extremely important". They were given the opportunity to add a factor themselves if they felt the list was incomplete. Participants were subsequently thanked for their participation.

Results

Survey respondents had an average job experience of about 11 years ($Mean = 11.19$; $SD = 6.53$). The most often cited sectors include the defence industry (26.5% of respondents) and the public sector (14.7%). They have been using others' forecasts for about 8 years ($Mean = 8.22$; $SD = 5.81$) and have been making their own forecasts for over 6 years ($Mean = 6.63$; $SD = 5.41$). Surprisingly, participants rated their own knowledge of forecasting techniques ($Mean = 2.78$; $SD = 1.22$) to be slightly below 3 (the midpoint between 1: no knowledge at all and 5: very high/extensive knowledge). A closer inspection reveals that 38.1% of the executives rated themselves as below the scale midpoint on this question, 33.6% provided a rating of exactly 3 (midpoint) and only 28.3% rated themselves to

1
2
3 be above the scale midpoint. The trust they place on statistical forecasts ($Mean = 3.51$; $SD = 0.89$) was
4 similar to the trust placed on other people's forecasts ($Mean = 3.54$; $SD = 0.79$) and in both cases
5 slightly above the neutral midpoint of the scale (i.e., 3).
6
7
8

9
10 Table 1 exhibits the scores and standard deviations of each factor, including a two tailed one-
11 sample t-test that compares the score with the neutral midpoint (a score of 3). With the exception of
12 'Forecasts are prepared by consulting with me', all items are significantly different from the neutral
13 middle point. Only 'Forecasts are prepared by consulting with me' scores below the middle point. A
14 one-tailed t-test shows this difference is significant at $p < .05$. Thus, consultation is the only factor that
15 is, on average, seen as less important in determining trust in forecasts. All other factors are seen as
16 important in determining trust in forecasts.
17
18
19
20
21
22
23
24
25
26
27

28 *****Table 1 About Here*****
29
30
31

32 Correlation analyses were used to explore potential associations between participants'
33 forecast experience, self-assessed expertise and their ratings of importance for each factor, Forecast
34 experience was measured by the longest time (in years) the participants reported to be either making
35 their own forecasts or using others' forecasts. For the association between forecast experience and
36 importance rankings, none of the Spearman correlation coefficients between the pairs was found to
37 be significant (all $p > .05$). It seems that the length of experience with making/using forecasts isn't
38 necessarily related to how important people believe certain factors are as determinants of trust. To
39 investigate further the existence of any association, the forecast experience was then categorized into
40 two groups ('moderate-experience' group – 1 to 9 years of forecast experience; 'high experience'
41 group – ≥ 10 years of forecast experience) and was tabulated against the percentages of high
42 importance ratings (aggregated across ratings of 4:very important and 5:extremely important) given
43 for each factor. Table 2 depicts these results.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

*****Table 2 About Here*****

The percentages presented in Table 2 also reinforce the finding that there appears to be no association between years of forecast experience and the importance people attribute to various factors as determinants of trust.

In terms of the association between self-assessed expertise and importance of factors, Spearman correlations showed that self-assessed level of knowledge/expertise was significantly related to the importance given to (i) knowledge about the statistical techniques used to generate forecasts ($p = 0.001$), and (ii) the forecast format (point forecasts, interval forecasts, etc.) ($p = 0.025$). To shed more light on this association, a similar analysis to the one presented in Table 2 was performed. The self-assessed forecast expertise was categorized into two groups ('moderate-expertise' group – ratings of 1, 2 and 3 ; the 'high-expertise' group – ratings of 4 and 5) and was tabulated against the percentages of high importance ratings (ratings of 4 and 5) given for each factor. These results are displayed in Table 3.

*****Table 3 About Here*****

The correlation analysis and the scores presented in Table 3 suggest that the forecasters who believe they have higher levels of expertise/knowledge give more importance to knowledge of statistical techniques used to generate the forecasts than those with moderate self-assessed expertise. There appear to be no significant differences for any of the other items (all $p > .05$).

In the next step of analysis, a principal axis factor analysis was conducted on the 15 features (excluding the last open-ended items) – i.e., importance ratings, that may affect a decision maker's trust in a given forecast. The chosen rotation method was the oblique rotation (direct oblimin) in order to accommodate for the case that extracted factors may not be independent and may correlate with each other. The Kaiser-Meyer-Olkin measure (Kaiser, 1970) designated that the sample was adequate

1
2
3 for the analysis, KMO = 0.765 (middling to meritorious according to Hutcheson & Sofroniou, 1999).
4
5 Bartlett's test of sphericity (Bartlett, 1950) also showed that the original correlation matrix is
6
7 significantly different from an identity matrix ($p < 0.0001$) further indicating the sampling adequacy.
8
9

10 An inspection of the scree plot (given in Figure 1) has revealed that inflection occurred
11
12 between 3 to 5 factors. In agreement with this, Kaiser's criterion (Kaiser, 1960) also designated that
13
14 there were four factors with eigenvalues strictly larger than 1.00 and a fifth factor just on the border
15
16 with an eigenvalue of exactly 1.00 (1.005 to be precise). Therefore, for brevity, it was chosen to retain
17
18 four factors in the analysis. These four factors together explained 59.98% of the total variation in the
19
20 data. The loadings of these factors after the rotation are exhibited in Table 4.
21
22
23
24
25
26
27

28 ***Figure 1 About Here***
29
30
31

32 ***Table 4 About Here***
33
34
35
36
37
38

39 This loading structure suggests that trust placed in a forecast seem to depend on four factors.
40
41 Factor 1 represents the **forecast bundle**, i.e., supportive features and tools provided in accompaniment
42
43 along with forecasts. As such, items loading high on this factor include scenarios, explanations, graphic
44
45 tools etc. Factor 2 can be interpreted as **forecaster competence**, i.e., perceived expertise and past
46
47 experience of the decision maker with the people who generate the forecasts and their knowledge
48
49 about the data used in the generation of forecasts. Factor 3 represents whether forecasts are
50
51 generated by a consensus or more than one forecaster and with multiple methods. This factor thus
52
53 refers to forecast **combination**. Lastly, Factor 4 designates the **knowledge** of statistical techniques and
54
55 forecasting methods. Factor 1, the forecast bundle, averages an importance rating for trust in forecasts
56
57 of 3.71 ($SD = .51$). Factor 2, forecaster competence, scores 3.96 ($SD = 0.14$) and factor 3, combination,
58
59
60

1
2
3 averages 4.05 ($SD = .03$). Factor 4, knowledge scores 3.60 ($SD = .92$). Factor 2 and 3 are the largest, but
4
5 are not significantly different from each other ($t(133) = -.52, p = .606$). They are however, both larger
6
7 than factor 1 (both at $t(133) = -3.99(f1-2); -5.23(f1-3), p < .001$), which in turn is not significantly
8
9 different from factor 4 ($t(133) = 1.32, p = .189$).

10
11
12 To verify the choice of the rotation method (i.e. the oblique rotation), the factor correlation
13
14 matrix was produced. All the off-diagonal correlation numbers among the factors are found to be quite
15
16 large in this matrix, therefore the factors, supportively, cannot be assumed as independent.

17
18
19 Looking at the scores of the respondents (the last column in Table 1), we can observe high
20
21 overall ratings for the items that refer to the combination of forecasts. Thus, it seems that combining
22
23 forecasts (from different techniques or different judges) is important to the users of the forecast. The
24
25 two highest scoring items are “My knowledge about expertise of persons making the forecasts” and
26
27 “Explanations accompanying the forecasts are given”, with each having 81% of the scores being a 4 or
28
29 5 on the scale. On the other end of the spectrum, we find that consultation with the end user of the
30
31 forecast does not seem to matter. Surprisingly, less than 50% scored ‘knowledge of the statistical
32
33 method’ as being important to place their trust on forecasts.

34
35
36 The answers to the open-ended questions were analysed in an iterative process, where every
37
38 answer was coded according to a response category. These response categories were identified by
39
40 reading the comments and structuring groupings around the keywords raised in the comments to allow
41
42 for the identification of unexpected themes. For the first open-ended question, i.e., “trust in forecasts
43
44 means the following to me”, we identified the categories *support*, *trusting people*, *trusting statistics*,
45
46 *source credibility*, *risk*, *error*, *based on history* and *combination of methods*. The first two categories,
47
48 namely *support* and *trusting people* were mentioned 3 – 5 times as often as the other categories. First
49
50 and foremost, the most often used category shows that forecasts are seen as a form of *support*.
51
52 Importantly, this indicates that trust may serve as a factor in the acceptance of software forecasts,
53
54 without employing damaging adjustments. Participants often mention how trusting forecasts can
55
56 provide guidance in doing one’s work, planning a project, and having a reliable source to work with.
57
58
59
60

1
2
3 As one participant so eloquently put it, being able to trust your forecasts means “*a tree branch to hold*
4 *on to in a stormy sea of uncertainty*”. It allows one to make plans for the future and shape strategies.
5
6
7 Importantly, it means being able to make decisions based on the data that is received. The second
8
9
10 distinct category is that of *trusting people*: trusting the forecast is most often perceived as trusting the
11
12 person behind the numbers, more so than trusting the underlying statistics: one person summarizes it
13
14 succinctly by stating that trusting forecasts means “*trusting other people's experience, data processing*
15 *and related skills*”. Vital to the setup of this study is the confirmation by one participant that “*trusting*
16 *forecasts means taking the forecasts made by others and using them in my analysis exactly as they're*
17 *given or with very little adjustment*”. Indeed, it seems that trust has a direct impact on the level of
18
19 adjustment for the participants. Thus, it is important to look into the factors that can heighten trust in
20
21 forecasts, thereby heightening acceptance of the forecasts.
22
23
24
25
26
27

28 With regard to the second open ended question, i.e., “the following is required to
29
30 increase/boost/enhance my trust in forecasts”, the following categories were identified: *accuracy,*
31 *appropriate data/parameters/methods, experience/expertise, source credibility, combination of*
32 *methods, and explanations*. Three categories were more prominent than others: *accuracy, appropriate*
33 *data/parameters/methods, and experience/expertise*. : First, a proven track record of accurate
34
35 forecasts is necessary to gain the trust of the forecast user. Most often mentioned is the accuracy of
36
37 the forecaster “*Forecaster has a high hit rate/accuracy*”, and “*It is important that the forecaster is an*
38 *expert in what s/he is doing and communicates his/her info well*”. Notably, it is not just the hit rate or
39
40 accuracy that is vital in boosting trust. Communication seems to be a vital factor, as well as honesty.
41
42
43 Second, the user wants to see that the appropriate data and methods have been used in creating the
44
45 forecast: “*Forecast should involve statistical procedures/analysis or based on a sound logical*
46 *foundation*”. The method needs to be reliable, as well as the data. Again, communication plays a role:
47
48 “*Forecast methods and analysis used should also be shared along with the forecasts.*”. Third, the
49
50 experience and expertise of the forecaster who makes the forecast needs to be high: Boosting trust is
51
52
53
54
55
56
57
58
59
60

possibly by “an *increase of experience and expertise of forecasters*” and “*Increase in experience, comparisons of forecasts with realized values*”.

Discussion

In this study, we set out to investigate the determinants of trust in forecasting. Using a forecasting survey, we found four factors that play a role. The first factor is the forecast bundle, i.e. supportive features and tools, such as explanations or graphical representations. As such, this result links back to the finding of Goodwin et al. (2013), who found explanations to have a significant effect on stated trust. In the more general advice taking literature we see reference to this finding as well: people take advice less because they do not know the rationale used by the advisor (Yaniv, 2004a, 2004b). As such, providing an explanation, a reasoning, behind the forecasts, should increase user’s acceptance of the forecast and lessen the judgmental adjustment. Gönül, Önkal, and Goodwin (2009) find confirmation of this in their study: one reason for ‘*not adjusting*’ reported to them by their respondents is that “the explanations provided with the acquired forecast are persuasive” (p. 28). Another reason cited by their respondents is “the presentation and the style of language used in the acquired forecasts are persuasive”, referring again to the supportive features of the forecast.

The second factor relates back to the competence of the forecaster (both the initial forecaster (be it human or algorithm) and the user of the forecast who can potentially adjust): what is their perceived experience and past expertise? This factor can be seen as the credibility of the forecaster. Credibility plays an important role in the trustworthiness of the advisor (Önkal et al., 2017). Sources who are seen as unbiased, objective and are well-known are motivation for not adjusting the forecast (Gönül et al., 2009). Additionally, if one feels that their own expertise outweighs that of the initial forecasters, this can be a sufficient reason to adjust (Gönül et al., 2009).

The third factor describes forecast combination, be it either a group of judgmental forecasters, or the combination of judgment and algorithms. Combining forecasts has long been known to outperform single method forecasts (Clemen, 1989), whether it is a combination of multiple judges

1
2
3 (Dalrymple, 1987; Mentzer & Cox, 1984), multiple algorithms (e.g., Makridakis, Hibon, & Moser, 1979)
4
5 or algorithms with judges (e.g., Blattberg & Hoch, 1990). Combining methods can reduce error as the
6
7 errors and potentials biases have a chance to average out.
8

9
10 The fourth and last factor refers to the knowledge of the decision maker on statistical
11
12 techniques and forecasting methods. Forecast training has the potential of future savings and process
13
14 improvement: Merrick, Hardin, and Walker (2006) found investments in forecast training to pay
15
16 themselves back quickly and many times over.
17

18 *Practical implications, limitations and future research*

19
20 This study suggests that there are a number of factors we can focus on to heighten trust in
21
22 forecasting, which can be used as a proxy for heightening acceptance of forecast advice, thereby
23
24 reducing harmful adjustments. While the large number of professional forecasters participating in the
25
26 survey is a main strength, it also brings limitations; e.g., the current sample does not allow comparisons
27
28 with non-experts. There are also potential limitations due to the validity and reliability of the survey
29
30 instrument, which may restrict generalizability of findings. This study was focused on professionals
31
32 with moderate to high knowledge of forecasting who had moderate to high experience in
33
34 using/making forecasts. Given that the participants were professionals with forecasting experience
35
36 (and not students or Amazon Mechanical Turk participants, for example) enhances the likelihood that
37
38 phrases like 'forecast format' are interpreted as intended. But still there could be unanticipated issues,
39
40 as common with all such surveys.
41
42
43
44

45
46 Lee and See (2004) provide a number of recommendations to increase trust in automation,
47
48 which may be transferable to the world of forecasting: amongst others, showing past performance,
49
50 showing the process and results in a comprehensible manner and showing the purpose of the
51
52 automation and how it relates to the user's goals. This relates back to the concept of communication,
53
54 which was often mentioned in the open ended questions on how to improve trust in forecasts. Thus,
55
56 providing explanations and supporting tools, stressing the expertise of the initial forecaster, combining
57
58
59
60

1
2
3 forecasts and knowledge can all contribute to heightened trust in forecasting, less forecast
4 adjustments and thus has the potential for improved forecasting accuracy.
5
6

7 Forecast support systems should be designed in a way to facilitate trust in forecasts by the
8 user, as suggested by the first factor, the supportive bundle. Features should include graphical versus
9 tabular representations and an option to provide explanations for forecasts. The latter will increase
10 trust when data, forecasts and adjustments are shared amongst forecasters. When people do not have
11 access to the rationale behind a forecast, they are more likely to adhere to their own judgment rather
12 than trusting the forecast given to them (Yaniv, 2004b). Importantly, the third factor suggests that an
13 option to combine forecasts is necessary when designing support systems: be it either different
14 statistical methods or the combination of as statistical system and judgment. This combining of
15 forecasts has long been known in research to outperform single forecasting methods (Fischer &
16 Harvey, 1999; Goodwin, 2002). This study confirms that practitioners recognize the improved forecast
17 accuracy from combining forecasts.
18
19
20
21
22
23
24
25
26
27
28
29
30
31

32 The second and fourth factor have implications for management practices such as hiring and
33 team guidance. Hired forecasters will be trusted more by their colleagues if they have a proven track
34 record and credibility. They will need to display the necessary forecasting knowledge and knowledge
35 about the product/service if their forecasts are to be trusted and their forecasts not adjusted
36 unnecessarily. At the same time, these external forecasters can effectively engage in forecast training
37 sessions where the knowledge base of the resident forecasters and decision makers could be
38 augmented. These efforts will all contribute to team building and synergy formation exercises directed
39 towards future planning and strategy development. After all, future-focused management is a
40 multiplayer effort and rests heavily on the shoulders of an effective team.
41
42
43
44
45
46
47
48
49
50
51

52 A potential limitation where the second factor, forecaster competence, is concerned, is that
53 people are generally not that good at assessing a forecaster's abilities (Gönül, Goodwin, & Önkül,
54 2012). We are heavily influenced by one-time, salient occurrences: one big error in an otherwise good
55 track record lowers our trust more than warranted – even more so for algorithms (Dietvorst et al.,
56
57
58
59
60

1
2
3 2015; Önkal et al., 2009; Prahla & Van Swol, 2017), while one accurate forecast of an unexpected event
4
5 in the middle of a bad track record raises our expectations disproportionately (Denrell & Fang, 2010).
6
7 One can also question if trust really does counter algorithm aversion. While we can derive from the
8
9 forecasting literature and advice literature that this is probably the case, future research would benefit
10
11 from structurally testing this in an experimental setup. One could manipulate trust by providing
12
13 information on past performance or on the credibility of the source. Trust is an evolving concept rather
14
15 than a static one: being confronted with accuracy of subsequent forecasts will lead to calibration of
16
17 trust in the forecaster/algorithm (the agreement between the level of trust in automation and the
18
19 actual performance of the automation; Lee & See, 2004). Importantly, trust does not develop in itself:
20
21 it is an interplay between an individual's propensity to trust, experiences with the algorithm and the
22
23 wider context in which the forecasting occurs: trust is an attitude that evolves over time (Rempel,
24
25 Holmdes, & Zanna, 1985). A careful experimental paradigm should take this feedback loop into account
26
27 to dig into the true effects of trust on algorithm acceptance.
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Alvarado-Valencia, J., & Barrero, L. H. (2014). Reliance, trust and heuristics in judgmental forecasting. *Computers in Human Behavior*, *36*, 102 - 113.
- Arkes, H. R. (2001). Overconfidence in judgmental forecasting. In J. S. Armstrong (Ed.), *Principles of forecasting*. Boston: Kluwer Academic Publishers.
- Baecke, P., De Baets, S., & Vanderheyden, K. (2017). Investigating the added value of integrating human judgement into statistical demand forecasting systems. *International Journal of Production Economics*, *191*, 85-96.
- Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*, 778-779.
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, *3*(2), 77-85.
- Blattberg, R. C., & Hoch, S. J. (1990). Database models and managerial intuition: 50% model + 50% manager. *Management Science*, *36*(8), 887 - 899.
- Bovi, M. (2009). Economic versus psychological forecasting. Evidence from consumer confidence surveys. *Journal of Economic Psychology*, *30*, 563 - 574.
- Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*(4), 559-583.
- Dalrymple, D. J. (1987). Sales forecasting practices: results from a United States survey. *International Journal of Forecasting*, *3*, 379 - 391.
- Denrell, J., & Fang, C. (2010). Predicting the next big thing: Success as a signal of poor judgment. *Management Science*, *56*, 1653-1667.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114. *Journal of Experimental Psychology: General*, *144*(1), 114-126.
- Dou, X., Walden, J. A., Lee, S., & Lee, J. Y. (2012). Does source matter? Examining source effects in online product reviews. *Computers in Human Behavior*, *28*(5), 1555-1563.
- Dzindolet, M. T., Pierce, L. G., Beck, H. P., & Dawe, L. A. (2002). The Perceived Utility of Human and Automated Aids in a Visual Detection Task. *Human Factors*, *44*(1), 79-94. doi:10.1518/0018720024494856
- Edwards, C., Spence, P. R., Gentile, C. J., Edwards, A., & Edwards, A. (2013). How much Klout do you have...A test of system generated cues on source credibility. *Computers in Human Behavior*, *29*(5), A12-A16.
- Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. *International Journal of Forecasting*, *25*(1), 3 - 23.
- Fildes, R., & Hastings, R. (1994). The Organization and Improvement of Market Forecasting. *Journal of the Operational Research Society*, *45*(1), 1-16. doi:10.1057/jors.1994.1
- Fildes, R., & Petropoulos, F. (2015). Improving Forecast Quality in Practice. *Foresight: The International Journal of Applied Forecasting*, Winter, 5 - 12.
- Fischer, I., & Harvey, N. (1999). Combining forecasts: what information do judges need to outperform the simple average? *International Journal of Forecasting*, *15*, 227 - 246.
- Gönül, S., Goodwin, P., & Önkal, D. (2012). Why should I trust your forecasts? *Foresight: The International Journal of Applied Forecasting*, *27*, 5-9.
- Gönül, S., Önkal, D., & Goodwin, P. (2009). Expectations, use and judgmental adjustment of external financial and economic forecasts: an empirical investigation. *Journal of Forecasting*, *28*(1), 19 - 37.
- Gönül, S., Önkal, D., & Lawrence, M. (2006). The effects of structural characteristics of explanations on use of a DSS. *Decision Support Systems*, *42*, 1481-1493.

- 1
2
3 Goodwin, P. (1996). Statistical correction of judgmental point forecasts and decisions. *Omega*,
4 *International Journal of Management Science*, 24(5), 551 - 559.
- 5 Goodwin, P. (2002). Integrating management judgment and statistical methods to improve short-
6 term forecasts. *Omega*, 30(2), 127-135. doi:[https://doi.org/10.1016/S0305-0483\(01\)00062-7](https://doi.org/10.1016/S0305-0483(01)00062-7)
- 7 Goodwin, P., Fildes, R., Lawrence, M., & Nikolopoulos, K. (2007). The process of using a forecasting
8 support system. *International Journal of Forecasting*, 23(3), 391 - 404.
- 9 Goodwin, P., Fildes, R., Lawrence, M., & Stephens, G. (2011). Restrictiveness and guidance in support
10 systems. *Omega : The International Journal of Management Science*, 39(3), 242 - 253.
- 11 Goodwin, P., Gönül, S., & Önkal, D. (2013). Antecedents and effects of trust in forecasting advice.
12 *International Journal of Forecasting*, 29, 354-366.
- 13 Goodwin, P., & Wright, G. (2010). The limits of forecasting methods in anticipating rare
14 events. *Technological forecasting and social change*, 77(3), 355-368.
- 15 Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing
16 responsibility. *Organizational Behavior & Human Decision Processes*, 70(2), 117-133.
- 17 Hilary, G., & Hsu, C. (2011). Endogenous overconfidence in managerial forecasts. *Journal of*
18 *Accounting and Economics*, 51(3), 300 - 313.
- 19 Hoch, S. J., & Schkade, D. A. (1996). A Psychological Approach to Decision Support Systems.
20 *Management Science*, 42(1), 51 - 64.
- 21 Hutcheson, G. D., & Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics*
22 *using generalized linear models*: Sage.
- 23 Johnson-George, C., & Swap, W. C. (1982). Measurement of specific interpersonal trust: Construction
24 and validation of a scale to assess trust in a specific other. *Journal of personality and social*
25 *psychology*, 43(6), 1306-1317.
- 26 Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and*
27 *psychological measurement*, 20(1), 141-151.
- 28 Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, 35(4), 401-415.
- 29 Kleinmuntz, D. (1990). Why we still use our heads instead of formulas: toward an integrative
30 approach. *Psychological Bulletin*, 107(3), 296 - 310.
- 31 Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). Judgmental forecasting: A review of
32 progress over the last 25years. *International Journal of Forecasting*, 22, 493 - 518.
- 33 Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human*
34 *Factors*, 46(1), 50-80. doi:10.1518/hfes.46.1.50_30392
- 35 Libby, R., & Rennekamp, K. (2012). Self-Serving Attribution Bias, Overconfidence, and the Issuance of
36 Management Forecasts. *Journal of Accounting Research*, 50(1), 197 - 231.
- 37 Lim, J. S., & O'Connor, M. (1995). Judgmental adjustment of initial forecasts: its effectiveness and
38 biases. *Journal of Behavioral Decision Making*, 8, 149 - 168.
- 39 Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human-human and
40 human-automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*,
41 8(4), 277-301. doi:10.1080/14639220500337708
- 42 Makridakis, S., Hibon, M., & Moser, C. (1979). Accuracy of Forecasting: An Empirical Investigation.
43 *Journal of the Royal Statistical Society. Series A (General)*, 142(2), 97-145.
44 doi:10.2307/2345077
- 45 Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integration model of organizational trust.
46 *Academy of Management Review*, 20(3), 709-734.
- 47 Mentzer, J. T., & Cox, J. E. (1984). Familiarity, application and performance of sales forecasting
48 techniques. *Journal of Forecasting*, 3, 27 - 36.
- 49 Mentzer, J. T., & Kahn, K. (1995). Forecasting Technique Familiarity: Satisfaction, Usage and
50 Application. *Journal of Forecasting*, 14(5), 465-476.
- 51 Merrick, J. R., Hardin, J. R., & Walker, R. (2006). Partnerships in Training. *Interfaces*, 36, 359 - 370.
- 52 Önkal, D., & Gönül, S. (2005). Judgmental adjustment: A challenge for providers and users of
53 forecasts. *Foresight: The International Journal of Applied Forecasting*, 1(1), 13-17.
- 54
55
56
57
58
59
60

- 1
2
3 Önkal, D., Gönül, S., Goodwin, P., Thomson, M., & Öz, E. (2017). Evaluating expert advice in
4 forecasting: Users' reactions to presumed vs. experienced credibility. *International Journal of*
5 *Forecasting*, 33(1), 280-297.
- 6 Önkal, D., Gönül, S., & Lawrence, M. (2008). Judgmental adjustments of previously adjusted
7 forecasts. *Decision Sciences*, 39(2), 213 - 238.
- 8 Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice
9 from human experts and statistical methods on forecast adjustments. *Journal of Behavioral*
10 *Decision Making*, 22, 390 - 409.
- 11 Prah, A., & Van Swol, L. M. (2017). Towards an understanding of algorithm aversion: Why do
12 decision-makers discount advice from automation. *Journal of Forecasting*. *Journal of*
13 *Forecasting*, 36(6), 691-702.
- 14 Rempel, J. K., Holmde, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of personality*
15 *and social psychology*, 49(1), 95 - 112.
- 16 Snizek, J. A., & Van Swol, L. M. (2001). Trust, confidence, and expertise in a judge-advisor system.
17 *Organizational Behavior & Human Decision Processes*, 84(2), 288-307.
- 18 Van den Broeke, M., De Baets, S., Vereecke, A., Baecke, P., & Vanderheyden, K. (2018). Judgmental
19 forecast adjustments over different time horizons. *Omega*, *In Press*.
- 20 Wang, W., & Benbasat, I. (2005). Trust and adoption of online recommendation agents. *Journal of the*
21 *Association for Information Systems*, 6, 72-101.
- 22 Wang, X., & Du, X. (2018). Why Does Advice Discounting Occur? The Combined Roles of Confidence
23 and Trust. *Frontiers in Psychology*, 9(2381), 1-8.
- 24 Yaniv, I. (2004a). The benefit of additional opinions. *Current Directions in Psychological Science*, 13,
25 75 - 78.
- 26 Yaniv, I. (2004b). Receiving other people's advice: Influence and benefit. *Organizational Behavior &*
27 *Human Decision Processes*, 93, 1 - 13.
- 28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Mean scores and standard deviations for each factor, including comparisons with neutral midpoint=3

(df for all t-tests = 133)

Factor	Mean	SD	t-value (test value = 3)	p-value (2-tailed)	Percentage of 4 and 5 ratings on all items
Explanations accompanying the forecasts are given	4.24	.91	15.60	< .001	81%
My knowledge about the expertise of the person(s) making the forecasts	4.11	.95	13.59	< .001	81%
Forecasts are elicited from more than one forecaster	4.08	.92	13.64	< .001	78%
Multiple techniques are used in generating forecasts	4.04	.98	12.20	< .001	78%
Consensus forecasts is given (by more than one forecaster)	4.02	.92	12.96	< .001	78%
My past experiences with the forecaster(s) making the forecasts	4.02	.99	11.93	< .001	75%
Presented forecasts are supported by scenarios	4.02	.86	13.72	< .001	73%
My knowledge about the data used to generate the forecasts	4.01	.98	11.98	< .001	75%
Information that the forecasts will be revised in known intervals	3.97	1.03	10.87	< .001	72%
Information about the person(s) making the forecasts	3.88	1.00	10.15	< .001	72%
My past experiences with technique(s) used to generate the forecasts	3.75	1.16	7.52	< .001	65%
Forecast format (point forecasts, interval forecasts, etc.)	3.63	1.02	7.23	< .001	57%
Forecasts are presented using graphic tools/plots	3.59	1.11	6.14	< .001	60%
Knowledge about the statistical techniques used to generate the forecasts	3.45	1.04	5.00	< .001	49%
Forecasts are prepared by consulting with me	2.81	1.21	-1.85	.067	30%

Table 2. Percentages of high importance ratings (4 and 5) for each factor as determinant of trust : Breakdown across forecast experience (counts/category total in parenthesis)

Percentage of 4 and 5 ratings on the factors	High Forecast Experience (>10 years)	Moderate Forecast Experience (1 to 9 years)	Is the difference significant?
Knowledge about the statistical techniques used to generate the forecasts	48.39% (30/62)	48.61% (35/72)	N.S.
Multiple techniques are used in generating forecasts	77.42% (48/62)	79.17% (57/72)	N.S.
My past experiences with technique(s) used to generate the forecasts	66.13% (41/62)	63.89% (46/72)	N.S.
Information about the person(s) making the forecasts	77.42% (48/62)	68.06% (49/72)	N.S.
Forecasts are elicited from more than one forecaster	82.26% (51/62)	75.00% (54/72)	N.S.
Consensus forecasts is given (by more than one forecaster)	77.42% (48/62)	79.17% (57/72)	N.S.
My past experiences with the forecaster(s) making the forecasts	74.19% (46/62)	75.00% (54/72)	N.S.
My knowledge about the expertise of the person(s) making the forecasts	80.65% (50/62)	81.94% (59/72)	N.S.
My knowledge about the data used to generate the forecasts	69.35% (43/62)	79.17% (57/72)	N.S.
Forecasts are prepared by consulting with me	33.87% (21/62)	26.39% (19/72)	N.S.
Explanations accompanying the forecasts are given	83.87% (52/62)	79.17% (57/72)	N.S.
Forecasts are presented using graphic tools/plots	58.06% (44/62)	61.11% (44/72)	N.S.
Forecast format (point forecasts, interval forecasts, etc.)	50.00% (31/62)	63.89% (46/72)	N.S.
Information that the forecasts will be revised in known intervals	72.58% (45/62)	70.83% (51/72)	N.S.
Presented forecasts are supported by scenarios	74.19% (46/62)	72.22% (52/72)	N.S.

Table 3. Percentages of high importance ratings (4 and 5) for each factor as determinant of trust: Breakdown across self-assessed expertise (counts/category total in parenthesis)

Percentage of 4 and 5 ratings on the factors	High Self-Assessed Expertise (Ratings of 4 and 5)	Moderate Self-Assessed Expertise (Ratings of 1,2 and 3)	Is the difference significant?
Knowledge about the statistical techniques used to generate the forecasts	68.42% (26/38)	40.63% (39/96)	Fisher's exact p = 0.004
Multiple techniques are used in generating forecasts	86.84% (33/38)	75.00% (72/96)	N.S.
My past experiences with technique(s) used to generate the forecasts	65.79% (25/38)	64.58% (62/96)	N.S.
Information about the person(s) making the forecasts	71.05% (27/38)	72.92% (70/96)	N.S.
Forecasts are elicited from more than one forecaster	73.68% (28/38)	80.21% (77/96)	N.S.
Consensus forecasts is given (by more than one forecaster)	76.32% (29/38)	79.17% (76/96)	N.S.
My past experiences with the forecaster(s) making the forecasts	78.95% (30/38)	72.92% (70/96)	N.S.
My knowledge about the expertise of the person(s) making the forecasts	86.84% (33/38)	79.17% (76/96)	N.S.
My knowledge about the data used to generate the forecasts	78.95% (30/38)	72.92% (70/96)	N.S.
Forecasts are prepared by consulting with me	34.21% (13/38)	28.13% (27/96)	N.S.
Explanations accompanying the forecasts are given	76.32% (29/38)	83.33% (80/96)	N.S.
Forecasts are presented using graphic tools/plots	60.53% (23/38)	59.38% (57/96)	N.S.
Forecast format (point forecasts, interval forecasts, etc.)	60.53% (23/38)	56.25% (54/96)	N.S.
Information that the forecasts will be revised in known intervals	76.32% (29/38)	69.79% (67/96)	N.S.
Presented forecasts are supported by scenarios	68.42% (26/38)	75.00% (72/96)	N.S.

Table 4. Factor loadings (pattern matrix) of features that affect trust in forecasts

FEATURES THAT AFFECT TRUST IN FORECASTS	ROTATED FACTOR LOADINGS			
	Factor 1	Factor 2	Factor 3	Factor 4
Presented forecasts are supported by scenarios	0.715	0.008	-0.140	-0.119
Forecasts are presented using graphic tools	0.641	0.143	-0.161	-0.006
Forecast format	0.559	-0.171	0.125	0.150
Info that forecasts will be revised in known intervals	0.518	-0.161	0.102	0.046
Explanations accompanying the forecasts are given	0.401	-0.216	-0.096	-0.071
Forecasts are prepared by consulting with me	0.394	0.099	-0.008	0.044
My knowledge about expertise of persons making the forecasts	-0.128	-0.810	-0.102	0.108
My past experiences with the forecasters making the forecasts	0.081	-0.773	-0.018	-0.171
Information about the person(s) making the forecasts	0.032	-0.671	-0.062	-0.018
My knowledge about the data used to generate the forecasts	0.060	-0.588	0.003	0.235
My past experiences with technique(s) used to generate the forecasts	0.010	-0.379	-0.045	0.314
Consensus forecasts are given by the forecasters	0.029	-0.076	-0.748	0.051
Forecasts are elicited from more than one forecaster	-0.024	-0.046	-0.695	-0.033
Multiple techniques are used in generating forecasts	0.325	-0.004	-0.382	0.188
Knowledge about the statistical techniques used to generate the forecasts	0.051	0.013	-0.043	0.821
Eigenvalues	4.553	2.033	1.336	1.075
% of total variance explained	30.36	13.56	8.91	7.16

Note. Factor loadings over 0.30 appear in bold.

Figure 1. Scree Plot

Peer Review Only

- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60

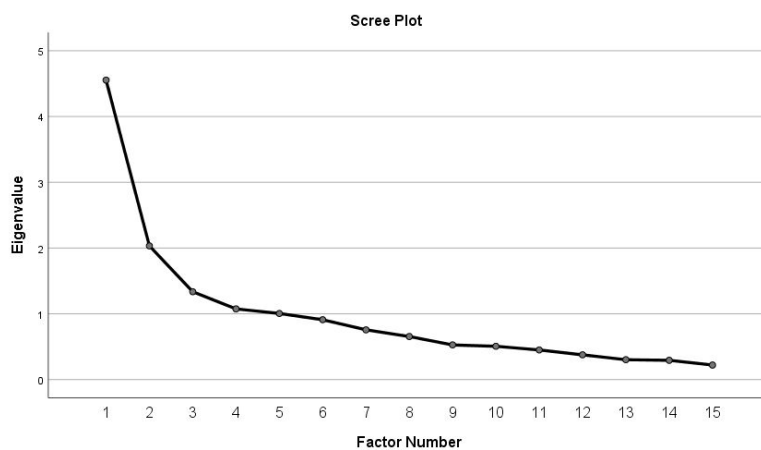


Figure 1. Scree Plot

Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60