

gesis

Leibniz Institute
for the Social Sciences



Survey Mastery

A Deep Dive into SQP 3.0 to Enhance
Questionnaire Development



MZES Social Science Data Lab

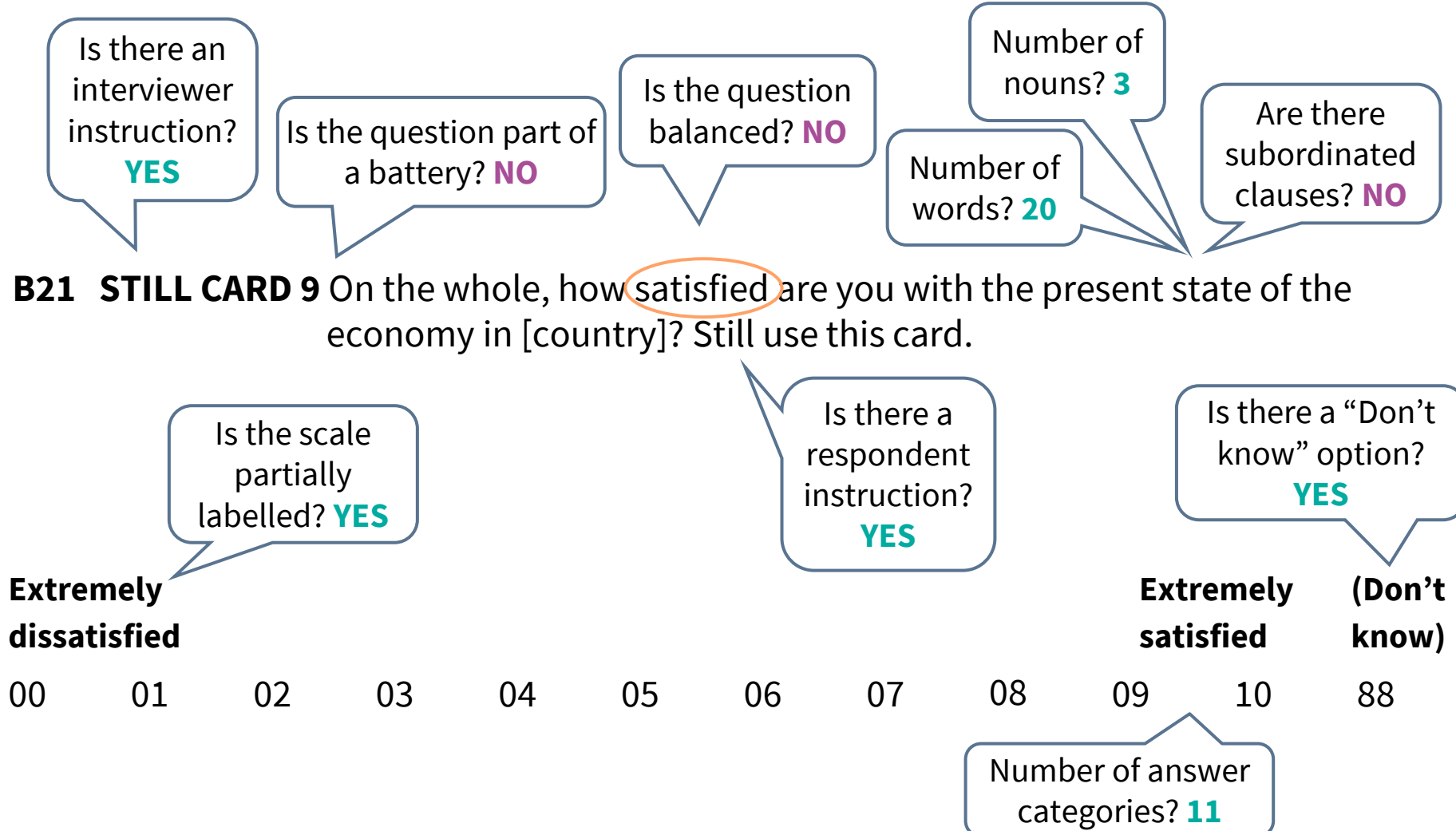
Dr. Lydia Repke

March 13, 2024

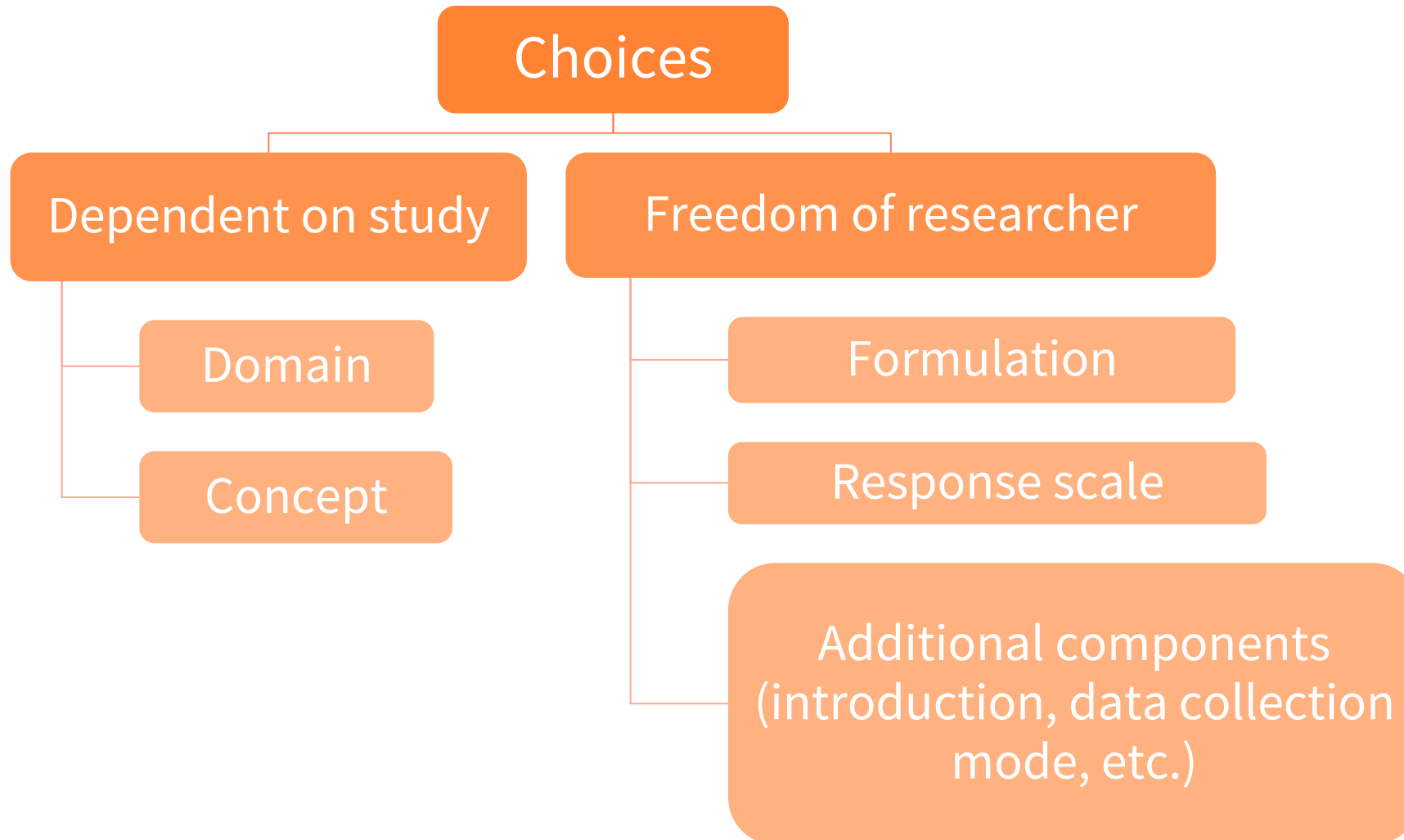
The image features a composite background. On the left, there is a textured, painterly surface with vibrant red, blue, and purple hues. On the right, a circular inset with a light blue background shows a laboratory scene with several glass test tubes. A glass dropper is positioned over one of the tubes, releasing a single drop of clear liquid. The overall design suggests a connection between art and science.

SQP_{3.0}

Decisions When Designing Survey Items



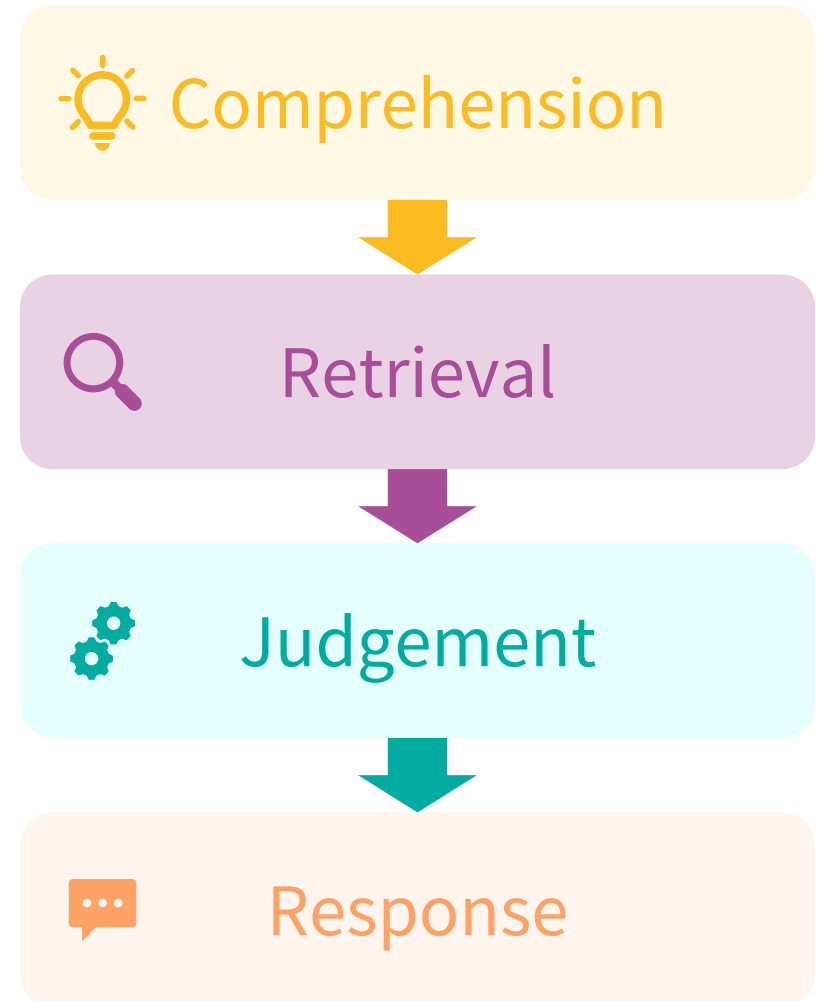
Designing a Survey Item



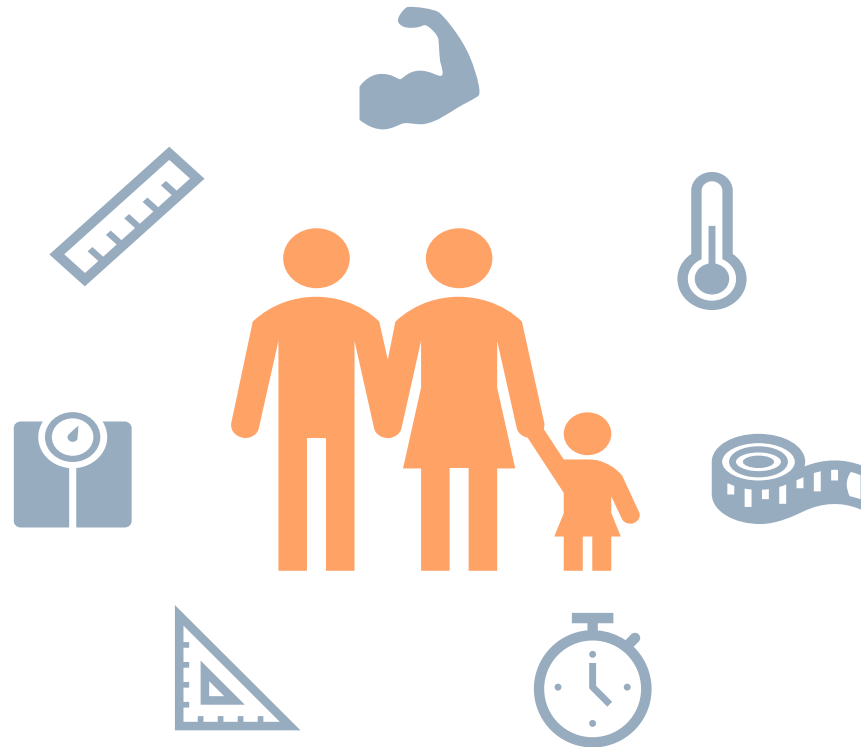
Why Do We Have to Bother?

Because all these decisions have an effect on the **response behavior** and as such also influence the **quality**

Measurement quality/measurement error



One Concept, but Many Ways to Measure



How do you decide *which* question or measurement method captures a concept best?



Immigration (ESS Round 3)

Now some questions about people from other countries coming to live in Britain.

Now, using this card, to what extent do you think Britain should allow people of the same race or ethnic group as most British people to come and live here?



- 1 Allow many to come and live here
- 2 Allow some
- 3 Allow a few
- 4 Allow none



Or Maybe Like This...?

Now some questions about people from other countries coming to live in Britain.

Using this card, please say how much you agree or disagree with each of these statements. Britain should allow more people of the same race or ethnic group as most British people to come and live here.

- 1 Agree strongly
- 2 Agree
- 3 Neither agree nor disagree
- 4 Disagree
- 5 Disagree strongly



Change in wording of
question + answer scale

*Does it make
any difference?*





What Difference Does It Make?



Use of stimulus or statement in the request	No stimulus or statement	Stimulus or statement is present
› Absolute or comparative judgment	An absolute judgement	An absolute judgement
› Response scale: basic choice	More than 2 categories scales	More than 2 categories scales
› Number of categories	4	5
› Labels of categories	Fully labelled	Fully labelled
› Labels with short text or complete sentences	Short text	Short text
› Order of the labels	First label positive	First label positive
› Correspondence between labels and numbers of the scale	Low correspondence	Low correspondence
› Theoretical range of the concept bipolar/unipolar	Theoretically unipolar	Theoretically bipolar
› Range of the used scale bipolar/unipolar		Bipolar
› Symmetry of response scale		Symmetric
› Neutral category		Present
› Number of fixed reference points	1	1

According to the quality prediction of SQP...

Quality prediction	Prediction of Question #1	Prediction of Question #2
Reliability = 1 - random error (r^2)	0.834	0.798
Validity = 1 - method effect (v^2)	0.903	0.754
Quality = reliability (r^2) x validity (v^2)	0.753	0.602



What Difference Does It Make?



Quality prediction	Prediction of Question #1	Prediction of Question #2
Reliability = 1 - random error (r^2)	0.834	0.798
Validity = 1 - method effect (v^2)	0.903	0.754
Quality = reliability (r^2) x validity (v^2)	0.753	0.602

WHAT IS SQP?

SQP 3.0 is a survey quality prediction system for continuous questions used in survey research.

- ☒ CONSULT, COMPARE, AND EVALUATE CONTINUOUS QUESTIONS
- ☒ DESIGN NEW QUESTIONNAIRES
- ☒ CORRECT FOR MEASUREMENT ERRORS

The Survey Quality Predictor

But what Exactly Is SQP?



The **Survey Quality Predictor** (SQP) is...

- an open-access system for **predicting the quality** of *continuous* survey questions
- and
- a **database** of survey questions and information about their quality.

<https://sqp.gesis.org/>

Saris & colleagues started
conducting and analyzing a
series of MTMM experiments

AAPOR's Warren J. Mitofsky
Innovators Award (2014)

SQP 1.0 for Windows
Saris, Oberski, & Kuipers, 2004

SQP 3.0
~ 600 MTMM experiments;
33 countries; 28 languages
Felderer et al.

1980's

1990's

2000's

2010's

2020's

SQP 2.1 (2015)

Idea of SQP

Hands-on tool for predicting
the quality of *new* questions
based on the questions'
characteristics

SQP for DOS
*87 MTMM experiments in
English, German, and Dutch
Saris, 2001*

SQP 2.0
+ >250 MTMM experiments from the
ESS; 22 countries; 21 languages
Saris et al., 2011

Item Characteristics Influence the Response Behavior

Language & Country of Survey

Trait

- Domain
- Concept

Associated to Trait

- Social Desirability
- Centrality
- Reference period

Instructions

- Interviewer instruction
- Respondent instruction

Additional Information

- Knowledge provided
- Extra information or definition

Introduction

- Request in intro

Linguistic Complexity

- # sentences
- # subordinated sentences
- # nouns
- # abstract nouns
- # syllables

Formulation of Request

- Direct, indirect, no request
- Stimulus (battery)
- WH word used in the request
- Type of 'WH' word
- Interrogative, imperative, declarative
- Gradation
- Balance
- Encouragement to answer
- Emphasis on subjective opinion
- Information about the opinion of other people
- Absolute or comparative judgment

Response Scale

- Categories, yes/no scale, frequencies, magnitude estimation, line production...
- # categories
- Full/partial labels
- Long/short labels
- Order of labels
- Label-number correspondence
- Theoretical range (bipolar/unipolar)
- Range used
- Symmetry of response scale
- Neutral category
- Number of fixed reference points
- Don't know option

Visual Aid Characteristics

- Showcard or other visual aids used
- Horizontal or vertical scale
- Overlap of scale labels and categories
- Start of the response sentence
- Request on the visual aid
- Numbers or letters before the answer categories
- Scale with only numbers or numbers in boxes
- Picture provided?

Questionnaire Context

- Computer-assisted
- Interviewer
- Visual or oral presentation
- Position of item



SQP from the Program's Perspective



Collection of all existing knowledge about the **quality** of survey questions
~ 600 MTMM experiments; 6,074 questions; 28 languages; 33 countries



Analysis of experiments and coding of formal characteristics
e.g., domain, balance of request, response scale



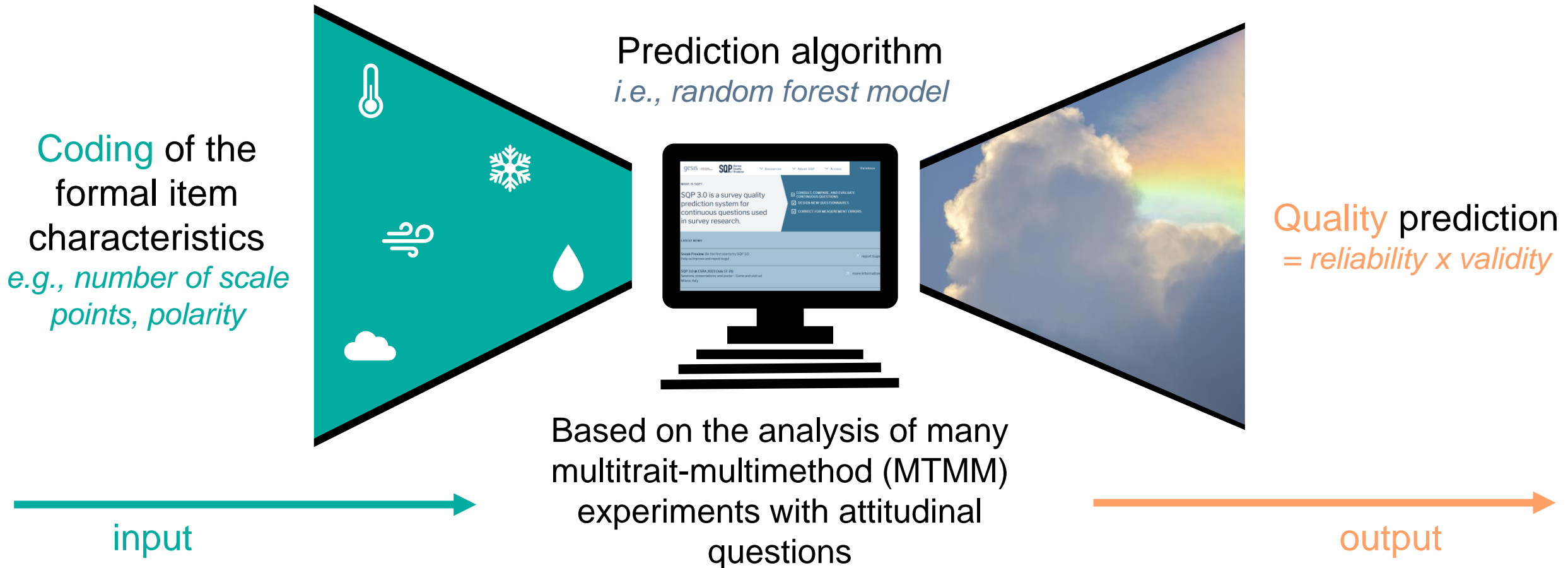
“Meta” analysis of experiments



Implementation of **prediction algorithm**
in software



SQP from the User's Perspective



cognitive pretesting

web probing

expert review

*“Many roads
lead to Rome.”*

Survey Quality Predictor 3.0



SQP as...



A dictionary
(for questionnaire development)



An evaluation tool
(comparing qualities, correcting for measurement error)



A translation check tool
(comparing the codings)

SQP as a “Dictionary” for Survey Questions



Database

Filter by ⓘ

Study	▼	Question owner	▼	Language	▼	Country	▼	Quality prediction	▼
-------	---	----------------	---	----------	---	---------	---	--------------------	---

Advanced filters



By text ⓘ

By coding characteristic:

Select characteristic	▼
-----------------------	---

Applied filters

Studies: Question owner: Language: Country: Quality prediction : Text: Characteristics:

SQP is a searchable database.

clear filters

apply filters

Database

Filter by ⓘ

Study	▼	Question owner	▼	Language	▼	Country	▼	Quality prediction	▼
-------	---	----------------	---	----------	---	---------	---	--------------------	---

Advanced filters

By text ⓘ

satisfied

By coding characteristic:

Select characteristic ▼

× RANGE OF THE USED SCALE BIPOLAR/UNIPOLAR

All

None

Delete



Bipolar



Unipolar

Applied filters

Studies:

Question owner:

Language:

Country:

Database

Filter by ⓘ

Study	▼	Question owner	▼	Language	▼	Country	▼	Quality prediction	▼
-------	---	----------------	---	----------	---	---------	---	--------------------	---

Advanced filters



Applied filters

Studies:	Question owner:	Language: <u>English</u>	Country:	Quality prediction :	Text: <u>satisfied</u>	Characteristics: <u>Range of the used scale bipolar/unipolar</u>
-----------------	------------------------	------------------------------------	-----------------	-----------------------------	----------------------------------	--

clear filters

apply filters



Question	Study	Language	Country	Quality	Actions
↑ HS9 / testc9 / <i>Political satisfaction: functioning of democracy</i>	ESS Round 4	English	United Kingdom	A	Replicate

Request for answer text:

And on the whole, how satisfied are you with the way democracy works in (Britain/the UK)? Use the same card.

Answer options:

- . 0 Dissatisfied
- . 1
- . 2
- . 3
- . 4
- . 5
- . 6
- . 7
- . 8
- . 9
- . 10 Satisfied

Information	Quality	Options
A Authorized prediction	0.712	  
M My quality prediction		

Filtering is based on all codings.

i.e., my codings, authorized codings, and other users' codings

The background of the slide is a composite image. On the left, there is a close-up, low-angle shot of a light bulb, showing its internal filament and the base. The lighting is dramatic, with warm orange and yellow tones from the bulb and cooler blue and teal tones from the surrounding environment. On the right, there is a dark blue, semi-transparent overlay that contains the title text. A pen nib is visible, appearing to write the text on this surface.

SQP as an Evaluation Tool

Consider This Version...

Question	Study	Language
↑ A9 / PplFair / <i>social trust, take advantage</i>	ESS Round 4	English

Request for answer text:

Using this card, do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?

Answer options:

- 00 Most people would try to take advantage of me
- 01
- 02
- 03
- 04
- 05
- 06
- 07
- 08
- 09
- 10 Most people would try to be fair

...

Or This Version

Question	Study	Language
↑ A9 / PplFair / <i>social trust, take advantage</i>	ESS Round 4	English

Request for answer text:

Using this card, do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?

Answer options:

- **00** Most people would try to take advantage of me
- **01**
- **02**
- **03**
- **04**
- **05**
- **06**
- **07**
- **08**
- **09**
- **10** Most people would try to be fair

↑ HS5 / testc5 / <i>Social trust: take advantage</i>	ESS Round 4	English
---	-------------	---------

Request for answer text:

Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair? Choose your answer from this card.

Answer options:

- **1** Most people would try to take advantage of me
- **2** Most people would try to be fair

And Their Quality Predictions

Quality prediction	Prediction of Question #1	Prediction of Question #2
Reliability = 1 - random error (r^2)	0.716	0.657
Validity = 1 - method effect (v^2)	0.879	0.838
Quality = reliability (r^2) x validity (v^2)	0.629	0.551
Reliability Coefficient (r)	0.846	0.811
Validity Coefficient (v)	0.937	0.916
Quality Coefficient (q)	0.793	0.742
Method Effect Coefficient (μ)	0.348	0.402
Reliability Coefficient Interquartile Range	(0.810, 0.910)	(0.770, 0.890)
Validity Coefficient Interquartile Range	(0.917, 0.990)	(0.870, 0.990)

SQP as a Translation Check Tool



Multilingual and Cross-National Surveys

- **Aim:** keeping comparability across languages and countries as high as possible
- SQP: comparison of formal characteristics between translated item and source item (or different countries)
 - Detection of formal differences due to translation (or country adaptation)
 - Allows to suggest corrections
 - But: differences might be due to the properties of the language (or societal context) itself

Compare code

ESS Round 5 / I15 / testd15 / Trust in police, effectiveness quickly arrive / United Kingdom / English **VS.** ESS Round 5 / I15 / testd15 / Trust in police, effectiveness quickly arrive / Spain / Catalan

download as CSV file

Legend: Difference due to the coding Difference due to the coding of the previous characteristic

Questions	Question #1	Question #2
Study	ESS Round 5	ESS Round 5
Question name	I15 / testd15 / Trust in police, effectiveness quickly arrive	I15 / testd15 / Trust in police, effectiveness quickly arrive
Country/Country prediction	United Kingdom	Spain
Language	English	Catalan
Request for an answer text	If a violent crime were to occur near to where you live and the police were called, how slowly or quickly do you think they would arrive at the scene? Choose your answer from this card, where 0 means very slowly and 4 means very quickly.	Si succeís un delicte amb violència a prop d'on viu vostè i truquessin a la policia, com de lent o ràpid creu que arribaria al lloc del delicte? Triï la resposta d'aquesta targeta.
Answer options	Very slowly Rather slowly Neither slowly nor quickly Rather quickly Very quickly	Molt lent Bastant lent Ni lent ni ràpid Bastant ràpid Molt ràpid
User	authorized coding	authorized coding

Codings	Coding of Question #1	Coding of Question #2
› Domain	National politics	National politics
› Domain: national politics	Local institutions	Local institutions
› Concept	All other simple concepts	All other simple concepts
› Concept: other simple concepts	Evaluation	Evaluation
› Social Desirability	Not present	A bit
› Centrality	Not at all central/salient	A bit central
› Reference period	Present	Future
› Formulation of the request for an answer: basic choice	Direct request	Indirect request
› WH word used in the request	WH word used	WH word used
› 'WH' word	How (extremity)	How (intensity)
› Request for an answer type	Interrogative	Imperative
› Use of gradation	Gradation used	Gradation used
› Balance of the request	Balanced or not applicable	Balanced or not applicable
› Presence of encouragement to answer	No particular encouragement present	Encouragement present
› Emphasis on subjective opinion in request	Emphasis on opinion present	Emphasis on opinion present

Also Resulting in Different Quality Estimates

Quality prediction	Prediction of Question #1	Prediction of Question #2
Reliability = 1 - random error (r^2)	0.725	0.723
Validity = 1 - method effect (v^2)	0.916	0.848
Quality = reliability (r^2) x validity (v^2)	0.664	0.613
Reliability Coefficient (r)	0.851	0.850
Validity Coefficient (v)	0.957	0.921
Quality Coefficient (q)	0.815	0.783
Method Effect Coefficient (μ)	0.290	0.390
Reliability Coefficient Interquartile Range	(0.820, 0.890)	(0.820, 0.900)
Validity Coefficient Interquartile Range	(0.950, 0.990)	(0.870, 0.990)

Two open hands, palms facing forward, are positioned on either side of the central text. The hands are light-skinned and have fingers spread. A semi-transparent blue horizontal bar is overlaid across the middle of the image, containing the text.

Hands-on Demonstration

A photograph of a stage with heavy red curtains. A spotlight illuminates a circular area on the curtains. A microphone on a stand is positioned in the center of the spotlight. The floor is made of wood.

Some Final Words...

Limitations

- Only for survey questions measuring **continuous** latent variables
≠ factual, observable questions such as demographics, facts, behavioral questions
- Only for **formal** characteristics
cannot substitute cognitive pretesting, expert review, or web probing techniques
- Empirical program aimed at **prediction**
≠ explanation
- Prediction limited to type of **questions in the MTMM experiments**

Must-Know Features of SQP 3.0



Search the database



Add your questions, code them, and get a quality prediction



Compare codes of questions



Download your codings

Summing Up

The **Survey Quality Predictor** (SQP) is...

- an open-access system for predicting the quality of *continuous* survey items and
- a database of survey questions and information about their quality.

Application areas of SQP are...

- 1) Questionnaire development
- 2) Evaluation of question quality
- 3) Translation checks

But SQP Is More Than Just a Software

- Big ongoing research project
- A lot of possibilities to collaborate
- Tons of data to analyse for papers

sqp@gesis.org
<https://sqp.gesis.org>



✉ lydia.repke@gesis.org
👕 <https://www.linkedin.com/in/lydiarepke/>
🦋 @lydiarepke.bsky.social
🖋 @LydRep

The Team Behind and Many More...

