VU Research Portal

Wording effects in survey research. Using meta-analysis to explain the forbid/allow asymmetry

Holleman, B.C.

published in

Journal of Quantitative Linguistics
1999

DOI (link to publisher)

[10.1076/jqul.6.1.29.4145](https://doi.org/10.1076/jqul.6.1.29.4145)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Holleman, B. C. (1999). Wording effects in survey research. Using meta-analysis to explain the forbid/allow asymmetry. *Journal of Quantitative Linguistics*, 6(1), 29-40. <https://doi.org/10.1076/jqul.6.1.29.4145>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

This article was downloaded by: [Vrije Universiteit Amsterdam]

On: 25 October 2012, At: 13:05

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Journal of Quantitative Linguistics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/njql20>

Wording Effects in Survey Research Using Meta-Analysis to Explain the Forbid/Allow Asymmetry

Bregje Holleman

Version of record first published: 09 Aug 2010.

To cite this article: Bregje Holleman (1999): Wording Effects in Survey Research Using Meta-Analysis to Explain the Forbid/Allow Asymmetry, Journal of Quantitative Linguistics, 6:1, 29-40

To link to this article: <http://dx.doi.org/10.1076/jqul.6.1.29.4145>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Wording Effects in Survey Research Using Meta-Analysis to Explain the Forbid/Allow Asymmetry*

Bregje Holleman
Utrecht Institute of Linguistics OTS

ABSTRACT

Answers to standardised attitude questions not only depend on the attitude respondents have on the issue, but are also influenced by the wording of the question. A well known example of this phenomenon is the forbid/allow asymmetry. Although the verbs 'forbid' and 'allow' are supposed to be each others' counterparts, the answers to questions worded with the verb 'forbid' turn out not to be opposite to answers to equivalent 'allow'-questions: respondents turn out to be more likely to respond 'no' rather than 'yes' to 'forbid' questions. It is commonly assumed that this is caused by the extreme connotations of both verbs. In this article a meta-analysis is reported over all forbid/allow research since 1940. First of all it is analysed whether the asymmetry can be generalised over questions and experiments. This turns out to be the case. The answer 'no, not forbid' is obtained more than 'yes, allow', and the mean effect size turns out to be quite large. The huge variance in asymmetry size however, offers room for a search after additional explanations. It turns out that the interaction between the complexity of the issue in the question, and the degree of abstractness of the question text, is systematically related to the size of the asymmetry. This stresses the importance of attitude strength as an explanatory variable, as well as the importance of looking at the communicative task as a whole.

INTRODUCTION

In designing questionnaires many choices have to be made, for example choices concerning question wording. Research has repeatedly demonstrated that small differences in the question wording causes large effects on the answers obtained (see Jobe & Mingay, 1991, for a review). These are called *wording effects*. For example, it turns out that answers to assertions are much less extreme than questions in question format (Petty, Rennier & Cacioppo, 1987).

Wording effects are a threat to the validity of survey questions, because the questionnaire designer can not be sure which particular question wording measures s/he intends to measure: the real opinion, the true attitude, can not be recovered. This communication problem can be solved by explaining the wording effect, using linguistic and psychological theories describing the differences between the two questions, and

conducting additional experiments in order to test these explanatory hypotheses. If the wording effect is explained, this can serve as a basis to solve the validity problem. At least equally important, however, is the more fundamental goal of gaining insight into the cognitive processes underlying question answering, and into the variables that affect responses (Cicourel, 1982). Wording effect research provides the opportunity to investigate question/answer processes and (differences in) meaning in a relatively natural task, thus providing ecological validity (Holleman, 1994). The theoretical as well as the practical research goals demand the same: more insight is needed into the relationship between the answers people give, and the opinion they have.

A wording effect that has received a lot of attention in more than half a century of research is the *forbid/allow asymmetry*, first identified by Rugg (1941). However, a satisfactory expla-

*Address correspondence to: Bregje Holleman, Utrecht Institute of Linguistics OTS, Trans 10, 3512 JK Utrecht, The Netherlands. Phone: + 31 30 2536068. Fax: + 31 30 2536000. E-mail: b.holleman@let.ruu.nl.

nation has not been found, so far. In this article a meta-analysis of all research into the forbid/allow asymmetry will be discussed. First of all it will be tested whether the asymmetry exists at all. Second, analyses will focus on explanations of the asymmetry using all previous research results that could be found in the literature.

FORBID/ALLOW ASYMMETRY

Although the verbs ‘forbid’ and ‘allow’ are considered to be each others’ counterparts, the answers to questions with the verb ‘forbid’ turn out not to be opposite to the answers to questions with the verb ‘allow’. The forbid/allow asymmetry means that a question worded with ‘forbid’ elicits more ‘no, not forbid’ answers than ‘yes allow’ in answer to the equivalent ‘allow’ question. Rugg (1941) found that respondents were more likely to support freedom of speech when the question was worded with the verb ‘forbid’, than when it was worded using the verb ‘allow’ – resulting in a difference of 21% between answers to two questions that are generally considered to be logically equivalent. As can be seen in the figure containing some examples of forbid/allow questions (Fig. 1), an experiment by Hippler and Schwarz (1986) elicited a difference of 14% between ‘not forbid’ and ‘yes allow’, and a forbid/allow question posed by Holleman a difference of 11.1%.

A lot of research into the forbid/allow asymmetry has been done. In the literature 52 forbid/allow questions could be found, in 15 different

experiments, administered in different languages and countries (Bishop, Hippler, Schwarz, & Strack, 1988; Hippler & Schwarz, 1986; Holleman, MS; Krosnick & Schuman, 1988; Rugg, 1941; Schuman & Presser, 1981; Waterplas, Billet, & Loosveldt, 1988;). An important problem of research into the forbid/allow asymmetry (and of research into many other wording or response effects as well), is that most experiments consist of one manipulated question only. This causes difficulties generalising the wording effect. On first sight, looking at all the experimental results, it is not certain whether the asymmetry exists at all, as the forbid/allow asymmetry is not always found. When it is not, researchers tend to formulate post-hoc hypotheses to explain the absence of the effect. Those hypotheses are hardly ever tested in subsequent research. Neither is it obvious whether this would be worthwhile, because as most forbid/allow experiments consist of one manipulated question only, there are always several possible causes for the asymmetry not to occur: differences in respondent characteristics, as well as administration mode, question issue, etc. (Waterplas et al., 1988).

GENERALISING THE FORBID/ALLOW ASYMMETRY

In order to be able to generalise the effect, analysis has to done over all experiments at the same time, instead of per single question. This was the first thing that was done in the meta-analysis, by using a (sort of) *t* test over all 52 forbid/

1. Do you think the US should forbid (allow) public speeches against democracy?			
yes, forbid	54%	yes, allow	25%
no, not forbid	46%	no, not allow	75%
2. Meinen Sie, der Gesetzgeber sollte die Vorführung pornographischer Filme in öffentlichen Kinos generell verbieten (erlauben?) [Do you think the legislator should forbid/allow X-rated movies in cinemas?]			
ja, verbieten [yes, forbid]	26%	ja, erlauben [yes, allow]	60%
nein, nicht verbieten [no, not forbid]	74%	nein, nicht erlauben [no, not allow]	40%
3. Vindt u dat de overheid militaire oefeningen in of vlakbij natuurgebieden moet verbieden (toelaten)? [Do you think the government should forbid/allow military exercises in or nearby nature areas?]			
ja, verbieden [yes, forbid]	73,5%	ja, toelaten [yes, allow]	15,4%
nee, niet verbieden [no, not forbid]	26,5%	nee, niet toelaten [no, not allow]	84,6%

Fig. 1. Some examples of forbid/allow questions: 1. Rugg (1941); 2. Hippler & Schwarz (1986); 3. Holleman (MS).

allow questions found in literature at the same time. The answers were weighted based on the number of respondents per question, so that questions answered by 1500 respondents weigh heavier than questions answered by 40 respondents.¹ Results of this first analysis show that the wording effect does exist. The mean size of the asymmetry was 14%: the answer 'no, not forbid' is given 14% more than the answer 'yes, allow' ($p < .001$). This is a large wording effect ($d = 1.44$).² It indicates that the forbid/allow asymmetry exists: 'forbid' and 'allow' do not function as real counterparts, at least not when used in attitude questions.

At the same time, huge differences in the size of the wording effect are found: the standard deviation is 9.85, the variance is large (97). This means that for 95% of the questions the difference between 'not forbid' and 'yes allow' lies between -6% and +34%: in some cases the asymmetry does not occur, in some cases it is large.

1. Analysis was done using a method of Iterative Generalised Least Squares, distinguishing between two levels of variance: variance between questions within experiments, and variance between experiments. This was done because questions within the same experiment can be expected to be more similar to each other than questions administered in different experiments (because of different respondents, for example). Distinguishing two levels facilitated interpretation of the results. However, here results will be reported using the overall variance figures.
2. Effect size is "the degree to which the phenomenon is present in the population" or "the degree to which the null hypothesis is false" (Cohen, 1977, pp. 9–10). This can be calculated by division of the mean effect size by the standard deviation of the difference between 'forbid'- and 'allow'-answers. The obtained effect size can then be evaluated using Cohen's criteria, by which an effect of about 0.2 standard deviation is defined as a small effect, effects of about half a standard deviation as medium effects, and effects of about 0.8 standard deviation are judged as large effects (Wolf, 1986).

EXPLANATIONS FOR THE FORBID/ALLOW ASYMMETRY

The basic explanation for the forbid/allow asymmetry, point of departure in all forbid/allow literature, focuses on the *connotations* of 'forbid' and 'allow': "the former sounds harsher and may therefore be more difficult to endorse, whereas the latter in some context might seem to encourage a deviant behavior and therefore may invite opposition" (Schuman & Presser, 1981, p. 296). "Thus what we have called tone of wording, could be the sole source of the effect." (Schuman & Presser, 1981, p. 280).

Although it does seem plausible for the connotations of 'forbid' and 'allow' to cause the asymmetry, two problems are attached to this explanation.

First, the explanation as worded by Schuman and Presser predicts that the asymmetry will occur quite constantly, and always have about the same size. Differences in the size of the asymmetry, then, should only reflect random error, and not be related systematically to other question characteristics, to the administration mode or to respondent characteristics. The amount of variance found in the asymmetry size raises serious doubts on this point, and offers room to look for explanations in addition to the connotations hypothesis.

Therefore, a second part of the meta-analysis (see the next section) was conducted, focusing on which characteristics of the question, the respondents, or the administration mode are systematically related to the size of the asymmetry. Insight into the causes of the variance may well provide insight into the mechanisms that cause the asymmetry, next to, or in interplay with, the connotations of 'forbid' and 'allow'. For in the forbid/allow literature, many explanations have been offered in addition to the connotations hypothesis. One of these focuses on *attitude strength*, and theorises that respondents holding a weak or very balanced opinion would be mainly responsible for the asymmetry. They find an affirmative answer too strong (because of the connotations of both verbs) and answer 'no' without realising that this implies an affirmative answer as well: 'no, not forbid' implies 'yes,

allow' (Hippler & Schwarz, 1986; cf. Fazio et al., 1982). This is about the only explanation that has been tested explicitly. Results are ambivalent however, probably partly because of differences in the operationalisation of the concept attitude strength.

In addition many other explanatory hypotheses for variation in the size of the asymmetry have been offered. Most of these have not been tested. Next to the hypothesised influence of *psychological factors*, such as attitude strength, three categories of hypotheses can be distinguished. First there are *linguistic hypotheses*, that focus on the effects of indicators of linguistic complexity. Questions that are linguistically complex would demand such an amount of working memory and attention in order to interpret the text that respondents during question answering do not realise that not forbidding actually implies allowing, thus causing a larger asymmetry for linguistically complex questions (Schuman & Presser, 1981). Second, there are hypotheses that focus on the influence of *question content* on the nature of the asymmetry between 'forbid' and 'allow'. For example, if a question is about an issue that is forbidden at the time the question is posed, the verb 'allow' may get a more 'active' and 'changing' meaning than when the issue was allowed at the time the question was posed. Third, characteristics of the *administration mode* of the question may influence the size of the asymmetry. For example: a question posed by phone or face-to-face gives the respondent less time to process the question and think of an answer than a written questionnaire. The respondent has less time to realise the implications of a negative answer - therefore, the asymmetry found in oral questionnaires may be larger.

A second problem connected to the connotations explanation is that it does not provide any real insight into the cognitive mechanisms underlying the asymmetry - causing it to remain a posthoc explanation rather than a hypothesis to be tested further. When explaining the forbid/allow asymmetry one wants to obtain insight into the relationship between the answers given to forbid- or allow-questions, and the attitudes the answers reflect. Generally, four stages are

distinguished within the question/answer process: interpretation of the question, location of the relevant attitude structure, retrieval of the attitude (or formation of the attitude), and fitting the judgment onto one of the precoded answering categories (cf. Sudman, Bradburn, & Schwarz, 1996; Tourangeau & Rasinski, 1988). The connotations hypothesis does not make explicit whether the difference between forbid/allow answers is caused in the first two stages of the question/answer process (during attitude localisation and retrieving, or formation); or whether similar attitudes are being measured and the asymmetry stems from the last stage, in which the opinion is mapped onto the answering options 'yes' or 'no'. Does the explanation, as worded by Schuman and Presser, mean that answers to 'forbid' questions reflect different attitudes than answers to equivalent 'allow' - questions? The connotations, or semantic fields of both verbs in general, may be that strong that not only the attitude towards a specific issue (abortion, for example) is measured, but also a general attitude towards forbidding or allowing. But it may also be the case that the asymmetry results from slight changes in perceptions of the meanings of attitude questions response options. Krosnick and Schuman (1988) theorise the asymmetry to be caused by differences in the way respondents map their answers to the answering options due to the use of 'forbid' and 'allow': "[...] 'not allowing' is perceived as a less extreme stance than is 'forbidding'."

This distinction between different attitudes being measured, versus similar attitudes being expressed differently onto the response options, is not only relevant for a better understanding of the question/answer process, but it is relevant for practice as well. If forbid/allow questions measure different attitudes, forbid/allow questions differ in validity. Either, or both, may not measure what the researcher intends to measure. But if forbid/allow questions measure a similar attitude, then they are equally valid, even though the answers to the questions differ. In that latter case, however, it is difficult, if not impossible, for the questionnaire designer to translate the yes/no answers back correctly to the true attitude.

Whether the forbid/allow asymmetry is caused by the retrieval of (partly) different attitudes due to the use of both verbs, or whether it is caused by a difference in mapping of the answers to the answering options, was tested by setting up two experiments (one on attitudes towards environmental issues, one on attitudes towards ethnic groups) using a correlational design with a within- as well as a between-subjects design. Forbid- and allow questions can only measure the same attitude if they correlate with unity, i.e., if they have a correlation of one.³

Analysis revealed that 'forbid' and 'allow' questions correlate highly. This means that ranking of respondents based on their answer to 'forbid' questions results in a similar order of respondents as would a ranking on the basis of their answers to 'allow' questions. This may lead to the conclusion that questions worded with either verb do measure similar attitudes. However, in both experiments observed scores and error scores to 'forbid' questions differ from those of 'allow' questions: the percentages 'yes' to 'forbid' differ from the percentages 'no' to 'allow'. This means that similar attitudes are expressed differently on the answering scale due to the use of 'forbid' and 'allow'. The interpretation of Krosnick and Schuman (1988) of the connotations explanation is correct: the answering options have different meanings due to the question wording. The extreme connotations of 'forbid' and 'allow' may be an explanation for this. But first of all it is feasible to check whether Schuman & Presser's claim that this might be "the sole source of the wording effect" is correct and if not, which other explanations seem warranted. This will be discussed in the next section.

LOOKING FOR ADDITIONAL EXPLANATIONS

An affirmative 'yes' to 'forbid' does not mean the same as a 'no' to 'allow'. Schuman and

Presser theorise that the sole reason for this may be the rather extreme connotations of both verbs. This would suggest that the asymmetry occurs in every question, and is about the same size all the time. The huge variances found in the asymmetry size, however, suggest that there is room for additional explanations. Therefore, it was explored which question characteristics, respondent characteristics or other characteristics are related to differences in the asymmetry size. Insight into the 'causes' of the variance may provide insight into the mechanisms that explain the origin of the asymmetry.

As was described earlier, the many additional explanations that have been offered in addition to the connotations hypothesis can be divided into four groups: explanations concerning psychological factors, linguistic hypotheses, hypotheses concerning question content, and hypotheses concerning characteristics of the administration mode. Of course, it would have been possible to test all of these explanations for variation in the size of the asymmetry experimentally. However, most explanations were based on the results obtained with just one question. So it seemed more useful to check first of all whether those explanatory suggestions are altogether true when analysing over all forbid/allow questions at the same time.

This was done by coding as many characteristics (like the ones mentioned above) as possible. This resulted in 15 explanatory variables, which were coded with either 0 or 1, so that the asymmetry size for, for example, oral questions (1) could be compared to the asymmetry size for written questions (0). The dependent variable was the difference between the answering percentage 'not forbid' and 'yes allow', again weighted for the number of respondents.

For each of the coded characteristics three effect characteristics were taken into account: significance of the effect, the amount of variance being explained (within and between experiments), and the mean effect size. Each of these effect sizes is important. The percentage variance explained is an indication of the explanatory power of a variable, but can also be an indication of incidental variation in presence or absence of a particular variable in this set of

3. What is meant here is a correlation coefficient corrected for attenuation, cf. the concept of *congenericity* (Jöreskog, 1971). See Holleman (MS) for technicalities of the design used in both experiments.

questions. If only one question was administered orally in the question set, then it is not very likely for this variable to explain a lot of variance. The significance of the effect is also partly dependent on the occurrence of characteristics in this question set, and is a prerequisite to base conclusions on any effect. The effect size indicates whether the effect of an explanatory variable is large enough to be meaningful, for example for future research.

The effects⁴ of 15 explanatory variables were analysed. Here only a few will be discussed, as a demonstration of how this kind of analysis works, and what kind of interpretation problems can arise.

First of all, the *complexity of the issue* in the question was coded. The theory concerning attitude strength discussed earlier hypothesises that respondents with a rather weak opinion, or holding a moderate or ambivalent opinion, find agreeing to either question too extreme, because of the extreme connotations of forbid and allow. When the issue in the question is very complex, more respondents are likely to have an ambivalent or weak opinion. So then the asymmetry should be larger. Whereas complexity depends on individual respondents, 15 students were asked to judge the complexity of the issue (resulting in an interrater reliability of 0.8).

Also, there is some literature hypothesising that the asymmetry will be larger if the question is *linguistically complex*, because then it would take more working memory to process it during the interpretation stage, causing respondents not to realise in the answering stage that the answer 'no' to 'forbid' implies a 'yes' to 'allow'. In linguistic and psychological literature many indicators of linguistic complexity can be found, for example, the length of the sentences, or the length of the text. Longer questions are theorised to be more difficult, because they place a heavier load on working memory (Angleitner, John, & Löhr, 1986, p. 81). For this analysis, the

number of sentences in each question was coded: most questions consist of only one sentence, if more sentences were used, this variable was coded 1.⁵ Also the use of *nominalisations* is supposed to increase linguistic complexity, because the degree of abstraction of the text is increased by leaving out the actor of the nominalised verb (Onrust et al., 1993). In this question set, however, almost all questions contain nominalisations.⁶ For this analysis it was coded whether a question contained nominalisations followed by predicates indicating the actor or the object of the verb. Although the use of predicates causes the question content to be less abstract, it causes the question to be linguistically more complex as well. Lastly, the degree of *abstractness* is usually measured as an indicator of linguistic complexity. The use of abstract terms can cause difficulties for the respondent in comparing the abstract term to the more specific or concrete term in his/her own head. Specific information in the respondent's memory has to be integrated with inferences in order to compare it to the abstract term in the question, causing the respondent to pay less attention to the implications of answering 'no'.

Also the effects of several content characteristics were looked into. Among others, the degree of morality of the issue at stake, and the status quo of the issue – i.e., whether or not the issue was forbidden at the moment the question was posed. Hypotheses on the effects of content characteristics were not very strong beforehand. For *morality*, offered by Waterplas et al. (1988) as a possible relevant content characteristic, the hypothesis was that issues with a high degree of morality are usually more abstract and without indication of the actor that should 'do' the forbidding, whereas less moral, and more legal, is-

4. It should be noted here that this article speaks about 'effects' where correlations are measured. Because no experimental manipulation was used, causal conclusions can not be drawn. However, as a certain causality is expected theoretically, I will refer to 'effects' when discussing the analyses.

5. Question length could have been coded differently, for example by counting the number of words, or clauses, instead of the number of sentences. However, this did not differentiate sufficiently between questions in this question set.

6. It is not very surprising that many forbid/allow questions consist of nominalisations: forbid and allow are often followed by nominalised verbs indicating which action is the object of forbidding/allowing - in English and German as well as in Dutch.

sues are possibly more concrete. If this is indeed the case, one would expect a smaller asymmetry for the more concrete act of forbidding/allowing, so for the 'less moral' issues. The effect of *status quo* was hypothesised to be that the active and changing connotation of 'forbid' weakens when the issue at stake is forbidden at the moment the question is posed. This would cause the asymmetry to diminish in size, respondents finding it easier to say 'yes' to a less extreme 'forbid'. At the same time, the connotation of 'allow' may get a more active connotation in this case, thus undoing this effect, respondents now finding it too extreme to answer affirmative to 'allow'. Intuitively one would say, however, that the connotations of 'forbid' are more extreme than those of 'allow', causing the asymmetry to diminish when the status quo of the issue at stake is forbidden.

FINDINGS

As reported earlier, when explanatory variables are not taken into account, the mean difference between 'not forbid' versus 'allow' answers is 14%. The variance turned out to be large, as the asymmetry size varied between -6% (6% more 'yes allow' than 'not forbid') and 34% (34%

more 'not forbid' than 'yes allow'). As will be discussed in this section, 60% of this variance in the asymmetry size could be explained.

The effect of issue complexity is as was predicted: 'not forbid' was answered 19% more than 'yes allow' for questions on complex issues, whereas the asymmetry size was only 11% for questions about non-complex issues. The total amount of variance explained by this variable is quite large: 44% (see Table 1).

Two indicators of linguistic complexity have the effect that was expected: the more linguistically complex, the larger the asymmetry. The effect of abstractness is very small: although it does explain a lot of variance, a difference of only 1% is negligible for future research. If a question consists of more than one sentence, the difference between the percentage 'not forbid' and 'yes allow' is larger. This hardly explains any variance, indicating that there is not much variation in this variable in the set of questions. Note that the effect of nominalisations is non-significant, so it is not necessary to have a further look into the amount of variance explained by it.

All in all, the effect of these indicators of linguistic complexity is not large or substantial. This is not very surprising, however. The indicators for linguistic complexity used here were

Table 1. Effects of Several Variables, when Estimated Separately.

Characteristic		Effects on asymmetry		p	Explained variance
		0	1		
attitude strength	issue complexity	11%	+8%	<.001	44%
linguistic complexity	abstractness	14%	+1%	<.001	36%
	number of sentences	13%	+6%	<.001	0.1%
	nominalisations	14%	+0.2%	n.s.	17%
question content	status quo	14%	+9%	<.001	0.8%
	morality	8%	+8%	<.001	40%

Note. The size of the asymmetry in percentages is given for cases in which the variable does occur, and does not occur (e.g. 0 is 'non-complex issues', 1 'complex issues'; 0 is 'non-abstract'; 0 is a question consisting of one sentence, 1 is one consisting of more sentences; 0 is no nominalisations + predicates; 0 is allowed at the moment the question was posed; 0 indicates questions on legal issues, 1 questions on moral issues), column p indicates the significance of the effect, and the last column signifies the total amount of explained variance within as well as between experiments.

the best possible, though not the most elegant. First of all it is hardly possible to code the degree of linguistic complexity without taking the context and content of the question into account: maybe then it would become clear that no simpler linguistic form could have been used given this specific context. Second, not every linguistic characteristic of a question necessarily has to play a role in the final representation (Verhagen, 1997). One nominalisation in one question may have no effect, whereas a questionnaire loaded with nominalisations can be experienced to be very linguistically complex. These are things that were not, and could not, be taken into account in the coding.

The effect of morality is as was predicted: questions on moral issues show a larger asymmetry (16%) than questions on legal issues (8%). In Table 1 can be seen that the asymmetry is larger if the issue is forbidden at the moment the question is posed. This is contrary to expectations. Several possible explanations for this finding can be thought of. First of all, on a semantic level, it is possible that the connotations do not change, or change differently from the way that was expected. Second, on a pragmatic level, it may be the case that the connotation of change of forbidding on a semantic level is quite strong, causing a clash in a context in which the act of forbidding implies a preservation of the status quo. Similarly, allow may have a stronger connotation of inactiveness than of activeness, thus causing a similar clash on the pragmatic level for the allow version. But, lastly, it may also very well be the case that this finding is difficult to interpret, because all issues that were forbidden at the moment the question was posed have other similarities apart from just this status quo variable. For example, in this question set, there is a distinct overlap between questions about issues that were forbidden at the moment the question was asked, and questions about ethnic politics, and questions with a high degree of morality.

A meta-analysis being a non-experimental design, this kind of overlap between possibly relevant question characteristics could not be controlled for. This is a data-technical reason to also estimate the effects of all question charac-

teristics in combination with each other - to find out whether, and where, there is overlap between the occurrence of variables, and to estimate interaction effects. Needless to say, there are not only data-technical reasons for doing this. Also from a linguistic-theoretical perspective this would be elegant: it may very well be expected that text characteristics, respondent characteristics (etc.) interact, or have certain threshold values before beginning to have any effect at all (Verhagen, 1997).

So, what could be done is to compute all possible interaction effects between any two or more variables, and interpret the interaction effects found either as a 'traditional' interaction effect, or as a sign of overlap between two or more variables. However, there are quite a few problems when doing this. First of all, of the 360 cells (the possible interaction effects with 6 variables) only 17 are filled in this data set. So not every interaction effect can be estimated, because of the fact that this dataset is based on a non-experimental design. Second, each interaction effect found would be difficult to interpret, because for each effect one has to find out whether the interaction should be interpreted as an overlap between two (or more) variables in this dataset (as is the case if one of the cells is empty), or as an interaction.

For this reason the analysis was restricted to one interaction effect that could be interpreted, and further the model was kept as simple as possible, mainly consisting of main effects. The only interaction effect estimated was the interaction between issue complexity and abstractness.⁷

As can be seen in Table 2, results of this analysis show that the size of the asymmetry is 17% for questions on complex issues – it does not matter whether these questions contain abstract terms or not. So, complexity leads to a 'ceiling

7. There is a certain overlap between these two variables in this question set, but all cells are filled, so the interaction can be estimated and can be interpreted as a 'traditional' interaction effect and not as the effect of overlap between variables. The overlap is as follows: many questions on non-complex issues are non-abstract (92% of the questions on non-complex issues), and most of the complex issues are abstract (90% of the complex issues).

Table 2. Interaction Effect of 'Complexity' and 'Abstractness'.

Characteristic	Asymmetry size
non-complex & non-abstract	11%
non-complex & abstract	8%
complex & non-abstract	17%
complex & abstract	17%

Note. The total amount of variance in the asymmetry size explained by this interaction effect is 44%.

effect': if the issue is complex, then other characteristics do not seem to play a role for the asymmetry size. Also, an interesting change is visible: estimated separately, questions on abstract issues lead to a larger asymmetry than questions on non-abstract issues. It was hypothesised that because of the load on working memory caused by abstractness, respondents would not realise that not forbidding implies allowing. Now it turns out that if complexity is taken into account, the degree of abstractness is only important for non-complex questions. Furthermore, now abstractness leads to a smaller asymmetry, which contradicts expectations as well as previous results.

This finding can only be explained by post-hoc hypothesising, to be tested in an experimental design. A possible explanation could be the following. Abstraction was coded as an indicator of linguistic complexity, which would increase the processing effort needed in the interpretation and attitude localisation stage. The large memory load would lead to a larger asymmetry in the answering stage, because respondents do not have enough attention for the implications of a negative answer. Maybe this is indeed the case for questions about complex issues. But issue complexity in itself already causes a maximum asymmetry size. At the same time, for questions on non-complex issues containing abstract terms, respondents may use the interpretation room the abstract term offers to choose an interpretation for the question that creates optimal possibilities to form or locate an extreme opinion, and be able to choose between the answering options offered, thus decreasing

Table 3. Effects of All Characteristics Estimated Together, to Detect Overlap.

Characteristic	Asymmetry size	<i>p</i>	Effect size
issue complexity	+6%	<.001	medium
abstractness	-2%	<.001	small
number of sentences	+5%	<.001	medium
nominalisations	0%	n.s.	—
status quo	+2%	<.001	small
morality	+4%	<.001	medium
issue complexity x abstractn.	+5%	<.001	medium

Note. The total amount of variance explained with this set of variables is 60%.

the asymmetry size. So when answering questions about non-complex issues containing abstract terms, respondents choose an interpretation for the abstract term that enables them to be extreme, and fulfill this specific communicative task: answering yes or no to a forbid/allow question.

Now that the interaction effect has been analysed, it is important to estimate the effects all together, instead of just individually, to detect possible overlap between characteristics in this dataset (see Table 3). Doing this, no important change can be found for most variables. All effects are a bit smaller, due to overlap between variables. Noticeably smaller are the effects of status quo and morality, indicating some overlap between these two variables, as was suggested earlier.⁸ So hypotheses on the effects of status quo or morality can hardly be tested on this data set. Note that the effects of abstractness and issue complexity in this estimation did not change substantially from the estimations in Table 2. Also, one has to be aware that, although the main effects of abstractness and complexity are in this Table 3, it is difficult to interpret them now that the interaction effect is significant.

8. Almost 97% of all legal issues were allowed at the moment the question was posed, whereas 20% of the moral issues was forbidden at the moment the question was posed. Only about 1% of questions on issues that were forbidden at the moment the question was posed were about legal issues.

CONCLUDING REMARKS

Concluding, several things can be learned on the basis of the analyses presented here. First of all, the main finding of this meta-analysis is that the forbid/allow asymmetry does exist. Overall, 'not forbid' is answered 14% more than 'yes allow': a small difference in question wording turns out to make a big difference for the answers. The connotations hypothesis only offers a partial explanation: it does explain why people prefer to say 'no' instead of 'yes'; as the asymmetry could as well have been in the opposite direction, as the preference for agreement-principle (cf. Sacks, 1987) would predict. But apart from the variation in the use of 'forbid' or 'allow', many other question characteristics turn out to influence the size of the wording effect, which indicates that research on the influence of (question) characteristics on the answering process remains essential and should not be limited to forbid/allow alone. The effect of the use of either verb seems to be diminished (abstraction), facilitated (complexity) or shaped (morality) by several question or respondent characteristics.

The large amount of variance in the asymmetry size offers room for additional explanations, apart from the connotations of 'forbid' and 'allow'. Based on the explanations for the asymmetry offered in the literature, the effects of fifteen explanatory variables were explored. In this article, the effects of six of these variables were discussed: the psychological variable 'issue complexity', the linguistic variables 'number of sentences', 'nominalisations', and 'abstractness', and two content variables 'status quo' and 'morality'.

As an individual effect, the most important characteristic turns out to be the degree of issue complexity, for it explains a lot of variance, and causes a large difference in asymmetry size. This supports theories on attitude strength, assuming more respondents holding a weak or ambivalent opinion will be found the more complex the issue in the question. The asymmetry is larger for complex issues, supporting the attitude strength hypothesis which states that mainly respondents holding a weak or ambivalent opinion are responsible for the asymmetry. They an-

swer 'no' to the forbid- as well as to the allow-question, finding agreeing to either question too extreme.

It is difficult to interpret this main effect of complexity however, because the effect is not constant for all levels of abstractness. This interaction effect is the most important explanatory variable found in this meta-analysis. It indicates that the source of the asymmetry has to be found in characteristics of the communicative task as a whole. When answering attitude questions in yes/no-format respondents have to interpret the question, locate or form the requested attitude, and map their answer, their opinion, onto one of the precoded answering options. Previous experiments already showed (Holleman, MS) that forbid/allow questions measure similar attitudes, but that the similar attitudes are expressed differently on the answering scales: forbid/allow questions are equally valid, but the conclusions based on the answers are not.

Based on the interaction effect found in this meta-analysis, one may hypothesise that especially the no-answer seems to represent a melting pot of several different opinions, such as "no, I do not think it should be forbidden/allowed", "no, I do not have an opinion on this, so I do not want to answer affirmative", or "no, I only moderately agree with forbidding/allowing, so I can not answer yes". Respondents holding a weak or ambivalent attitude answer 'no' because they find answering affirmative too extreme. Forbid and allow may well be each others' counterparts, but no to forbid is not a counterpart of yes to allow – not always, anyway. Respondents do try, however, to fulfill the communicative task: if the question leaves room for it, they prefer to interpret the (abstract or vague words in the) question in such a way that it is possible for them to answer affirmative, to answer extreme, instead of stretching the meaning of the answering option 'no'. Further, experimental, research should continue from this hypothesis.

The set of six explanatory variables tested in this meta-analysis explains 60% of the variance in the size of the asymmetry. Of this 60%, most is explained by the interaction effect between complexity and abstractness, again stressing the importance of these variables. The effects of the

content variables 'morality' and 'status quo' are difficult to interpret. Because of their huge overlap within this question set, no hypotheses can be based on these effects generalising beyond this data set. The effect of the number of sentences in each question is quite stable: questions consisting of more than one sentence lead to a larger asymmetry. However, the amount of variance explained by it is quite small (due to the fact that most questions in this set only consist of one sentence: variation on this point is small), and furthermore, the theory that might explain the effect of the number of sentences is rather thin. Longer questions supposedly are linguistically more complex, just because of their length and the load on working memory that places, whereas it is just as plausible to assume that questions consisting of more sentences are more concrete, and consist of more explanation of the object or backgrounds of the question. So for future research, one may assign a low priority to this explanatory variable.

In general this meta-analysis demonstrated that in natural language processing the effects of text or context characteristics overlap or interact. Within the large and complex communicative task of answering an attitude question several kinds of question characteristics that may be theorised to be of possible importance in themselves may turn out not to play a substantial role among other characteristics. Also it was shown that the context of survey research offers possibilities of doing research on language interpretation and question/answer processes in a (relatively) natural task, a task in which it is possible to relate text- or respondent characteristics to quantitative reactions (i.e., the percentages yes/no).

Finally, it has also been made clear that meta-analytic techniques can be very helpful in cumulating years of research. Once the analysis aims at explaining a phenomenon, such as the forbid/allow asymmetry, one has to be careful not to interpret findings too quickly in terms of generalisations beyond the data set. Analysis has to be done carefully step by step, using each result to get more insight into the structure of the dataset and the interdependencies between variables. Then an exploratory meta-analysis

performed on a dataset like this one can generate useful hypotheses that provide insight into the mechanisms behind the wording effect and that can be tested in experimental research.

Knowing the results of this meta-analysis then, which question should be preferred, the forbid question or the allow question? The large variance in asymmetry size means that the asymmetry is sometimes large, and sometimes does not occur. In other words, sometimes the conclusions based on the answers to a forbid/allow question are equally valid, whereas sometimes the conclusions based on one of either questions is more valid than a conclusion based on the other.

This meta-analysis indicates that the asymmetry size is dependent of question- and respondent characteristics that are not related to the attitude the question intends to measure. However, as the dependent variable was the difference between the 'not forbid' and 'yes allow' answers, and because in this meta-analysis it was not possible to compare equivalent questions with and without a certain characteristic (with and without nominalisations, for example), this analysis did not reveal whether the asymmetry size changed because of a change in the forbid-answers, because of a change in the allow-answers, or because of a change in both. Further research could try to find out whether it is the forbid-answers, or rather the allow-answers, that are sensitive to change caused by background variables like attitude strength. As a criterion for the best question wording one could decide that answers to questions that fluctuate least caused by background variables unrelated to the attitude intended to measure are preferable.

REFERENCES

- Angleitner, A., John, O.P., & Löhr, F.J. (1986). It's what you ask and how you ask it: An itemmetric analysis of personality questionnaires. In A. Angleitner & S. Wiggins S. (Eds.), *Personality assessment via questionnaires: Current issues in theory and measurement* (pp. 61–107). Berlin: Springer Verlag.
- Bishop, G.F., Hippler, H.-J., Schwarz, N., & Strack, F. (1988). A comparison of response effects in self-administered and telephone surveys. In R.M. Groves et al. (Eds.), *Telephone survey methodology* (pp. 321–340). New York: Wiley.

- Cicourel, A.V. (1982). Interviews, surveys and the problem of ecological validity. *The American Sociologist*, 17, 11-20.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (revised ed.). New York: Academic Press.
- Fazio, R.H., Sherman, S.J., & Herr, P.M. (1982). The feature-positive effect in the self-perception process: Does not doing matter as much as doing? *Journal of Personality and Social Psychology*, 42, 401-411.
- Hippler, H.-J., & Schwarz, N. (1986). Not forbidding isn't allowing: The cognitive basis of the forbid-allow asymmetry. *Public Opinion Quarterly*, 50, 87-96.
- Holleman, B.C. (1994). Analyse van vragenlijstonderzoek vanuit een taalgebruiksperspectief [Analysis of surveys from a language use perspective]. In A. Maes, P. van Hauwermeiren & L. van Waes (Eds.), *Perspectieven in taalbeheersingsonderzoek* (pp. 240-249). Dordrecht: ICG Publications.
- Holleman, B.C. (MS, in review with Sociological Methods and Research), *The nature of the forbid/allow asymmetry: two correlational studies*. 36 pp.
- Jobe, J., & Mingay, D.J. (1991). Cognition and survey measurement: History and overview. *Applied Cognitive Psychology*, 5, 175-192.
- Krosnick, J.A., & Schuman, H. (1988). Attitude intensity, importance, and certainty and susceptibility to response effects. *Journal of Personality and Social Psychology*, 54, 940-952.
- Petric, I. (1992). *Predicting listening performance for news texts*. Dissertation Utrecht University.
- Petty, R.E., Rennie, G.A., & Cacioppo, J.T. (1987). Assertion versus interrogation format in opinion surveys. Questions enhance thoughtful responding. *Public Opinion Quarterly*, 51, 481-494.
- Rugg, D. (1941). Experiments in wording questions: II. *Public Opinion Quarterly*, 5, 91-92.
- Sacks, H. (1987). On the preferences for agreement and contiguity in sequences in conversation. In G. Button & J.R.E. Lee (Eds.), *Talk and social organisation* (pp. 54-69). Clevedon: Multilingual Matters.
- Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys. Experiments on question form, wording and context*. New York: Academic Press.
- Sudman, S., Bradburn, N.M., & Schwarz, N. (1996). Answering a survey question: Cognitive and communicative processes. In *Thinking about answers* (pp. 55-79). San Francisco: Jossey-Bass.
- Tourangeau, R., & Rasinski, K.A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, 103, 299-314.
- Verhagen, A. (1997). Context, meaning and interpretation, in a practical approach to linguistics. In L. Lentz & H. Pander Maat (Eds.), *Discourse analysis and evaluation: Functional approaches* (pp. 7-39). Amsterdam: Rodopi.
- Waterplas, L., Billiet, J., & Loosveldt, G. (1988). De verbieden versus niet toelaten asymmetrie. Een stabiel formuleringseffect in survey-onderzoek? [The forbid/allow asymmetry. A stable wording effect in survey research?] *Mens en Maatschappij*, 63, 399-417.
- Wolf, F.M. (1986). Meta-analysis, quantitative methods for research synthesis. *Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-059*. Beverly Hills: Sage Publications.