DOUBLE TROUBLE: MULTIPLE COVERS IN SPECIFIC ISSUE READERSHIP MEASUREMENT

Michal Galin, Ph.D., Julian Baim, Ph.D., Kerry Zarnitz, GfK MRI Dr. Irena Petric, NOM Alke Bassler, Intomart GfK

Introduction

For the measurement of specific issue readership, respondents are shown the covers of magazines. Intomart GfK uses this method in the Dutch National Readership Study of NOM, while GfK MRI uses the same in their Issue Specific Study in the United States.

As opposed to the recent reading method, which can profit from learnings based on thirty years of symposia of the WRRS/P&DRF (among others), specific issue readership in its current form must still develop its own best practices for measurement. This paper, comparing the approaches used in the American and the Dutch studies to deal with multiple covers, is one of the starting points. (Papers delineating the methodology of either study in greater detail are listed in the bibliography.)

Both studies are faced with an increasing number of magazine issues which have more than one cover. Analyses have been made in both countries to determine the best way to deal with these situations: whether or not all covers should be shown, how respondents react when multiple covers are shown, the potential impact of including more than one cover, etc. This paper will discuss the findings and compare the different approaches in the Netherlands and the United States.

The challenge of measuring multiple magazine covers can be expected to be a constant issue in the measurement of specific issue readership, as multiple covers are a tactic which publishers are using increasingly as they seek to attract audiences in this age of fragmentation and competition with other media and platforms.

Measurement

Issue-specific measurement began in both countries in 2006, although the Dutch and U.S. studies differ in their ultimate objectives. The Dutch National Readership Study (NPM) measures specific issue readership in order to determine Average Issue Readership, while GfK MRI's Issue Specific Readership Study aims to capture issue-to-issue variability. In the United States, the GfK MRI National Study produces Average Issue Readership information.

For the Dutch National Readership Study, an annual sample of 22,000 respondents is interviewed either via an online questionnaire or random probability sample (see Petric & Appel 2007 for more details on the fieldwork). 58% of the sample is comprised of respondents from the Intomart GfK Online Access Panel, while 42% of the sample is recruited from a random probability sample. This 42% is then further subdivided, with 40% of the random probability sample completing the questionnaire with an interviewer present, while the other 60% fill in the questionnaire on their own. GfK MRI's study, though, relies solely on an Internet sample (derived from two sample panel companies) and accumulates 5,000 completed surveys on a weekly basis (over 260,000 completed surveys annually).

The GfK MRI survey is open from Wednesday afternoon until Thursday night each week with each week an independent sample of completed surveys. Once a respondent has participated in the survey, he/she is not eligible to be invited back for six months. Data for the Intomart GfK survey is collected continuously throughout the year.

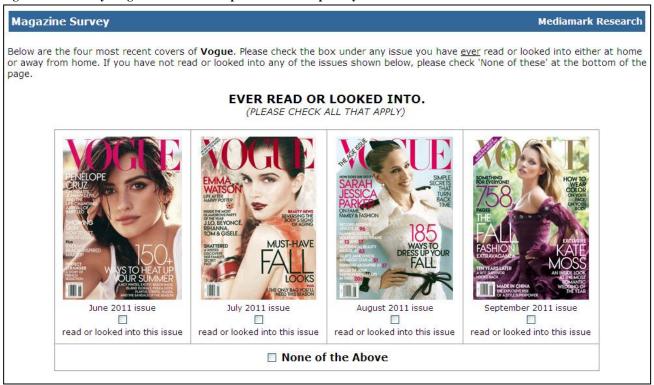
The structure of the two questionnaires is comparable to most readership surveys. Both begin with a screen-in: in Intomart GfK's, respondents are asked to indicate for each of 140 magazines and approximately 15 national and regional newspapers whether or not they have read or looked at the title in the past 12 months by clicking on 'yes' or 'no', GfK MRI's Issue Specific study begins with a six-month screen question of 246 magazine logos, excluding newspapers. Once GfK MRI's respondents have indicated which titles they have read or looked into in the last six months, the survey proceeds to the issue-by-issue section, while Intomart GfK determines source of copy next, before the specific issue readership is measured. In both studies, however, specific issue readership is determined by showing respondents covers of the most recent issues. In Intomart GfK 's study, the 6 most recent issues are displayed for weeklies and monthlies, and the 4 most recent issues are displayed for bimonthlies and quarterlies. In GfK MRI's study, on the other hand, for weekly and bi-weekly titles, the respondent is shown the

eight most recent covers, while for tri-weeklies, monthlies, and bi-monthlies, the respondent is shown the four most recent covers. In another similarity, both studies arrange a title's covers in order, with the newest issue's cover at the farthest right position. Examples of each study's cover recognition question appear in Figures 1 and 2, respectively.



Figure 1. A Monthly Magazine in the Dutch National Readership Study

Figure 2. A Monthly Magazine in the Issue Specific Readership Study



As with the screen-in question, the Dutch study requires respondents for each issue to indicate whether they have read or looked at the issue by choosing one of three possible answers:

- 1. Yes, read or looked at
- 2. No, not read or looked at
- 3. Don't know

If respondents click on the cover, an enlarged image is shown, so that headlines can be read for better recognition. The answer "yes, read or looked at" is counted as issue readership.

In GfK MRI's study, on the other hand, issue readership is determined simply by whether or not a respondent selects the positive response option underneath the issues' cover images.

After the cover recognition sections, both studies include questions regarding socio-demographics, with Intomart GfK also asking about other media and GfK MRI asking where-read and how-obtained questions for the screened-in titles. Furthermore, Intomart GfK allows respondents to give general feedback, as well as to rate the survey overall.

And Then There Were Two (or More)

When specific issue measurement was set up in the two countries, neither anticipated the growing obstacle of multiple covers for a single issue. In its very first week of employing cover recognition however, the Dutch Readership Survey was faced for the first time with an issue with two covers: one for male subscribers and one for female subscribers. GfK MRI learned of the existence of splits covers more gradually, though since then, multiple covers have appeared with increasing frequency in both countries. Publishers implement them for any number of reasons, such as in commemoration of a special event (e.g., the end of a popular television show) or as collector's items, among other reasons, and there are many different permutations of splits/cover treatments employed by magazine publishers. In the Netherlands, they have seen issues with up to 7 different covers, while an issue with as many as 26 cover variations has been encountered in the United States. Furthermore, there are complex implications when some titles produce multiple covers and other titles do not, as well as when, within a title, some issues have more than one cover while other issues do not. Additionally, there are difficulties when even a title that is consistent in publication of multiple covers has variance in the extent to which the cover versions differ from one another.

In the original surveys developed by both countries, there were no structural procedures for displaying multiple covers in place. Both had strict guidelines for representing the recent covers of an issue. Any deviation from that formula would render any data collected incomparable to the previous data collected, though neglecting the existence of multiple covers for an issue potentially limited the accurate measurement of the issue. Therefore, Intomart GfK and GfK MRI each had to devise approaches to the measuring multiple covers as well as single covers.

Approaches

Both studies found two possible problems with displaying a double cover:

- Respondents might think that the two covers represent two different issues. Do they understand that an issue should be marked "yes, read or looked at" if they have read the magazine with one of the covers?
- The size of each cover needs to be decreased to fit in the space in the questionnaire allotted to one issue, therefore each cover is more difficult to recognize.

In the Dutch Readership Survey, multiple covers are combined into one image, which is then displayed in the appropriate space in the cover recognition question. The combined image has the same size as is allotted for an issue with a single cover. Obviously, this means that each of the multiple covers is smaller than a single cover display. Up to 4 covers can be combined; if there are more, those covers with the highest circulation will be used. Respondents can enlarge the image by clicking on it, making the individual covers more recognizable. Even though two or more covers are displayed, the combination still stands for one issue with one set of yes/no/don't know as possible answers.

In the Netherlands, most multiple covers are one time occasions. In such instances, the image of the multiple cover versions is displayed amidst the other single cover versions on the screen. However, there are some titles which have multiple covers on a regular basis. In January 2010, the bi-weekly automotive magazine, Autovisie, decided to use two different covers for every issue: one cover for subscribers and a different cover for selling in supermarkets or newsstands. Often, this cover has the same

or similar picture, and additionally, more headlines are used. Sometimes, two completely different covers are used, though the content of the issue is identical. Figure 3 shows three examples.

Figure 3. Autovisie: Examples of Two Covers for the Same Issue



For GfK MRI's Issue Specific Study a different approach to accurately represent split covers was tested. At present, GfK MRI can only include one cover per issue in the study, with one exception, which will be elaborated upon below. As seen by in the Netherlands, there are different types of splits; in the United States, newsstand/subscriber, geographical, and demographic splits all have been encountered. In the experience of GfK MRI, newsstand/subscriber splits tend to occur most frequently. If there is such a split for an issue(s) of a title, GfK MRI requests that the magazine provide the cover representing the greater proportion of its circulation (e.g., if the title has a larger proportion of subscribers, GfK MRI requests the subscriber cover in this case).

In the case of geographical splits, GfK MRI successfully tested an approach for including geographically split covers and found an effective method to measure such splits. This approach was first implemented in August 2009. In this approach, the survey unobtrusively reads the zip code of the computer on which the respondent is taking the survey, and then shows the respondent the correct cover for his/her geographical area. This method requires advance notice from publishers, as well as the zip code breakdowns by printing plant, provided by the publishers. Once the zip code breakdowns have been programmed, the designation of which covers will be sent to which areas is also required.

Also in the fall of 2009, GfK MRI entered into a development phase to review the current approach to split covers and developed a new standard of rules for such cases. The standards for releasing the data for issues with split covers, which were put into effect in September 2009, can be found in the appendix and also are posted on Issue Specific's deliverable.

However, the guidelines set forth by the Issue Specific Release Standards provide an imperfect solution for the quandary created by multiple cover versions. Firstly, the guidelines do not account for variability between cover differences within magazines that consistently publish multiple cover versions, thereby rendering the titles data less comparable from issue to issue. Furthermore, it seems reasonable to assume that when one cover version represents a miniscule portion of a title's total circulation its existence should not confuse respondents. Lastly, the standards rely heavily upon the subjectivity of GfK MRI, for individuals are tasked with evaluating the extent of the disparity between an issue's covers, which invites the possibility of human error into the process.

GfK MRI long dismissed the option of displaying more than one cover per issue in its study, out of concern that to do so might confuse respondents, or cause an unjustified spike in the number of reads for an issue. Moreover, such an approach is ineffectual in the instances of more than two covers per issue and it is most likely unsuitable when a title produces only occasional multiple cover versions.

Nonetheless, despite its misgivings about deviating from its standard procedure of presenting an issue's covers, the growing number of multiple cover issues compelled GfK MRI to investigate alternative procedures to its current methodology. Thus, in November 2010, GfK MRI fielded a small-scale test in an attempt to examine the ramifications of the various options of displaying a title's split covers. The overall objective of this test was to evaluate if including multiple covers for an issue served to increase the scores for the title simply because the respondent had more chances to select the issue. To that end, three parallel studies were incorporated into this examination: Version 1, in which respondents were shown the cover that had been determined as the "dominant" cover, and thus was the same cover shown in the main Issue Specific Study; Version 2, in which respondents were shown the predominant alternate cover (in some cases there were more than two cover versions); and Version 3, in which respondents were shown both the two most prevalent cover versions together.

The test was fielded for four weeks, using one sample panel and accruing 750 completes per week (n=3000 respondents across the 4-week test), and included nine titles. The titles were selected because each had produced multiple covers consistently for every one of its issues for the past year. GfK MRI also evaluated how significantly the title's split covers differed from one another.

Clearly, the test did not set out to answer how to handle an occasional issue with multiple covers. Instead, the objective was to establish how to best represent those covers that are affected by split covers for every issue.

Findings

When the first few double covers were used in the Dutch Readership Survey, a preliminary check was done by examining comments given at the end of the questionnaire and by asking interviewers for feedback. There was no indication that respondents were experiencing difficulties with the questionnaire.

Another indicator of confusion could be the percentage of "don't know"-answers in the cover recognition question. An increase in respondents who cannot indicate with certainty whether or not they have read an issue would be grounds to assume that the double cover images pose problems for respondents.

On average for all magazines, 10-12% of the respondents who screened in give the answer "don't know" to the cover recognition question. As can be expected, the percentage is higher for the oldest issues (average of 11.6%) than for the most recent issues (9.8%). Automotive magazines are the runner up for the highest percentages "don't know": from 12.8% for the most recent issue to 13.8% for the oldest issue (only home and garden magazines rank higher). Autovisie, in particular, has always had one of the highest percentages "don't know".

An analysis was done of the percentages of "don't know" answers given before and after double covers were introduced. After the introduction of double covers, the percentage of "don't know" increased significantly. The difference was especially notable for the oldest issue. Figure 4 shows the average percentages of "don't know"-answers in 2009 (1 cover per issue) vs. the first half of 2010 (2 covers per issue).

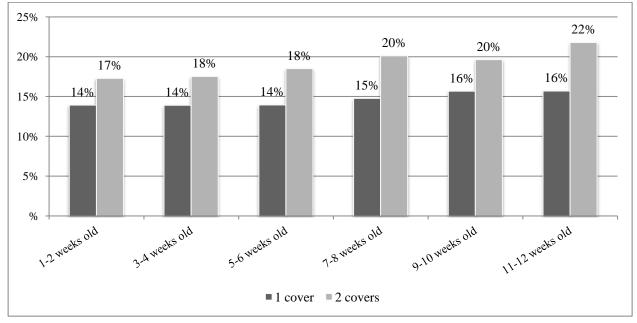


Figure 4. Percentages of "Don't Know" Responses for 1 vs. 2 Covers

Because of the increase, it was decided to add an explanation in the cover recognition question for Autovisie:

"Please note: each issue of Autovisie is published with two different covers. You only need to have read or looked at one of the two."

Figure 5 shows the percentages of "don't know"-answers after the explanation was added to the question. The levels of "don't know" are slightly lower, but not significantly so.

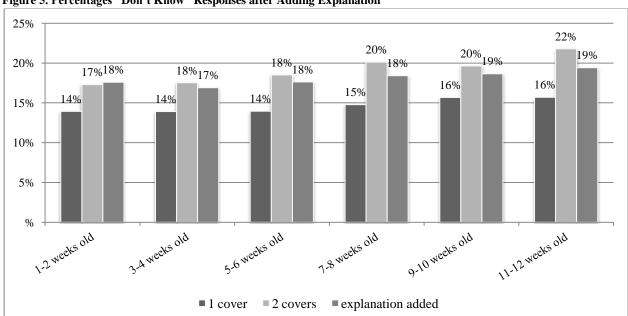


Figure 5. Percentages "Don't Know" Responses after Adding Explanation

Figure 6 shows that there is a difference between subscribers and non-subscribers. The display of double covers does not affect the subscribers. Due to their habitual reading, they can identify read issues with much greater certainty, as just 5% of the subscribers "don't know" whether or not they have read an issue. This level does not change after double covers are displayed. Respondents who buy their issues at a newsstand/in the supermarket, etc., are less sure to begin with: their percentage of "don't know"-answers is twice as high when one cover is used and more than triple that of subscribers when the double cover is displayed.

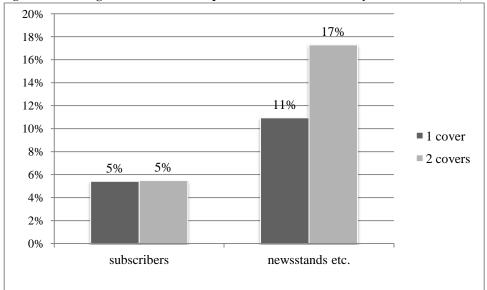


Figure 6. Percentages "Don't Know" Responses for Subscribers vs. Buyers at Newsstands, etc.

Does the fact that the percentage of "don't know"-answers is higher for certain groups of readers (i.e., non-subscribers) influence the level of Average Issue Readership for Autovisie? After the introduction of double covers in the measurement of readership, the levels of AIR stayed relatively stable. In the same period, levels of screen-ins increased from 6.4% to 7.7%. These new readers did not generate average issue readership. Rather, the levels of "don't know" as well as "no, not read or looked at" increased.

Two explanations are possible: either the new readers are infrequent readers, which means the measurement of Autovisie is correctly representing reality, or AIR is underestimated, because readers are confused by the double covers.

There were some changes in the composition of the circulation: while there were no changes in newsstands sales, there was a slight decrease in subscriptions as well as an increase of circulation by reading circles; however, the overall level of circulation remained relatively stable. Circulation data thus do not help us to confirm or reject either of the possible explanations above.

Additional analysis is needed in order to determine the effect of displaying double covers in the measurement of specific issue readership. Are readers less sure about having read a certain issue of a magazine because they are infrequent/less engaged readers, or is the question itself confusing respondents? If confusion is the main factor, however, there should be some effect on the data received from subscribers as well. Nevertheless, the smaller image size of the displayed cover is a concern.

In GfK MRI's multi-covers test, the focus was placed on two factors presumed to significantly affect the measurement of issues with split covers: the circulation split between different versions for an issue and the degree of difference between the cover images for an issue. It was hypothesized that survey versions showing two covers would consistently generate higher issue-specific audiences than versions showing only one cover image, regardless of the circulation splits or difference in cover imagery. Regarding the latter, the disparity between versions may range from minor differences in cover lines (unlikely to affect measurement) to considerable changes in the number and presentation of cover lines as well as in the actual images used. As for the former, some of the titles have circulation splits that are predominantly skewed (90% or more of the circulation) towards a subscriber base (in the U.S., most titles have larger subscription bases than newsstand sales). In such cases, if such a title has covers split along circulation lines, GfK MRI feels it can reasonably assume that the dominant cover is representative of the issue; if each version accounts for more than 10% of the circulation, there is less assurance that a single cover stimulus will suffice.

The nine titles represented in GfK MRI's test all had consistently published split covers and all had split issues with differences. (One title presented unique complications to measurement beyond its split covers, and thus was eliminated from the analysis.) Among the remaining eight titles, however, the breadth of the difference between the cover versions fluctuated. For some titles, every issue had covers that were markedly different from one another, while some other titles published covers that were visually indistinguishable at first glance. Moreover, the total number of cover versions for the issues included in the study varied, with some issues having as many as ten different configurations.

Our findings were not always consistent with our initial hypothesis and, overall, because this was a small-scale test, the conclusions are suggestive. That said, from the data rendered by the test study, GfK MRI found that titles with substantial circulation versions and consistently different cover executions will produce larger issue-specific audiences from the presentation of two of its cover versions, as opposed to just one. The theory is that issues with **two drastically different covers** that were seen by **substantial portions of the circulation** (roughly more than 10%) are not accurately measured by presenting respondents with a single cover, while those issues with fairly similar versions and/or one version seen by a predominant proportion of the total circulation were adequately represented by the single dominant cover and therefore were not significantly benefited by the presence of two covers. Only two titles of the eight, Elle and Harper's Bazaar, satisfied both criteria, as seen below in Figure 7.

Figure 7. Satisfaction of Criteria by the 8 Studied Titles

Title	Circulation: % Subscription	Circulation: % Newsstand	Both Versions Have Circulation > 10%	Major Differences Every Issue	Cover for
Elle	73.7%	23.6%	$\sqrt{}$	√	
Entrepreneur	94.1%	5.9%		√	
Family Circle	84.3%	15/7%	$\sqrt{}$		
Harper's Bazaar	78.6%	21.4%	\checkmark	√	
Ladies' Home Journal	95.6%	4.4%			
Men's Health	78.4%	21.6%	$\sqrt{}$		
Midwest Living	94.5%	5.5%			
Road & Track	91.6%	8.4%			

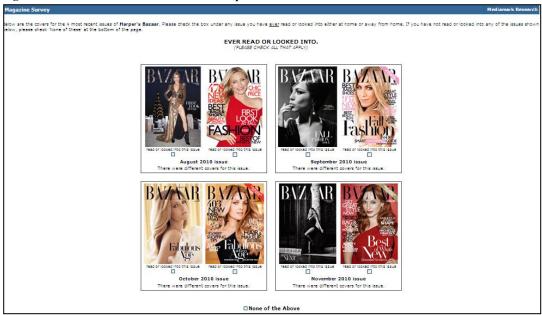
Note: circulation figures in Table 7 come from the title's ABC Statement.

For those two highlighted titles, Elle and Harper's Bazaar, the data for the test version which presented respondents with both covers for each issue was noticeably higher than the data for either of the other two versions. The test versions showing both covers for Elle and Harper's Bazaar are shown in Figures 8 and 9, respectively.

Figure 8. Two Cover Versions Shown for Elle



Figure 9. Two Cover Versions Shown for Harper's Bazaar



Two other magazines, Men's Health and Family Circle, had comparable circulation subscriber/newsstand splits to Elle and Harper's Bazaar, though their cover versions were not as dissimilar (see Figures 10 and 11, below, which reproduce the test version showing both covers). And, Entrepreneur presented the scenario in which its two covers looked completely different one another (see Figure 12) but the circulation was skewed predominantly toward subscriptions (94.1% of total circulation is subscriptions). In each of these three examples the data is not in the expected direction, and GfK MRI surmises that this is due to the fact that the two conditions, circulation split and cover image disparity, are not met.

Figure 10. Two Cover Versions Shown for Men's Health



Figure 11. Two Cover Versions Shown for Family Circle



Figure 12. Two Cover Versions Shown for Entrepreneur

Megazine Survey

Mediamark Resear

Below are the covers for the 4 most recent issues of Entrepreneur, Please check the box under any issue you have <u>over</u> read or looked into either at home or away from home. If you have not read or looked into any of the issues shown below, please check

None of these if the bestimen of the page.

EVER READ OR LOOKED INTO.



None of the Above

The issue-specific audiences for Men's Health and Family Circle, however, were not significantly higher when both covers were shown which, GfK MRI hypothesizes, is due to the fundamental similarities between the issue covers. On the other hand, in the case of Entrepreneur, the issue-specific audiences were not significantly higher when both covers were shown, but the power of this small-scale test was probably not strong enough to detect differences for a title with one portion of its circulation, newsstand sales, comprising less than 10% of the total. Additional testing would be required to evaluate such circumstances with more precision.

From the findings yielded by this initial test, though, some conclusions can be drawn. Firstly, GfK MRI believes that for titles that have consistently split covers and that satisfy both more balanced circulation and divergent cover criteria, the presentation of two cover versions is beneficial to measurement. Therefore, GfK MRI is planning to implement that presentation in such instances moving forward.

Conclusion

The appropriate measurement of multiple cover versions for a single issue in a specific issue readership survey seems to be a challenge that is arising with growing frequency and in a variety of countries. NOM/Intomart GfK and GfK MRI have similar readership surveys, though with different objectives for those studies. Nevertheless, our two issue specific studies find themselves confronting the same obstacle—how to accurately assess the readership of a magazine issue that publishes multiple covers, within the bounds of their existing surveys.

GfK MRI's test was done on a relatively small scale; therefore, additional testing is required. From the findings yielded by this initial test, though, some conclusions can be drawn. Firstly, GfK MRI believes that for titles that have consistently split covers and that satisfy both more balanced circulation and divergent cover criteria, the presentation of two cover versions is beneficial to measurement. Therefore, GfK MRI is planning to implement that presentation in such instances moving forward. Furthermore, the results from NOM/Intomart GfK's test are reassuring. They suggest that displaying two covers for an issue does not cause inflation of the readership numbers. There might be some confusion about the display of two covers, however this is not the case for subscribers/frequent readers. It might be the case for infrequent readers, thus negatively influencing audience figures.

For the Dutch survey it is interesting to learn that in case of similar covers, the choice of which cover is displayed has little effect on the readership figures. The same is true when both covers are displayed. In case of magazines which use split covers on single occasions, the displaying double covers might be abandoned if covers are similar, thus avoiding any possible confusion for respondents (and extra work necessary to make the double cover display for the survey). This will have to be discussed in the Netherlands.

Findings from the individual work done by GfK MRI and NOM/Intomart GfK suggest there are possible solutions to at least some of the issues presented by multiple covers. Though some magazine examples may pose issues that make measurement more difficult, both companies have made progress towards more appropriately measuring a greater number of issues with split covers. These steps made toward better measurement, however, will require continued study as the methods are further evaluated.

Appendix: Issue Specific Split Cover Release Standards

This note summarizes criteria used in determining the release of issues with split covers. These criteria are listed in order from relatively easy decisions to the most difficult decisions. All split covers will be evaluated before release.

- 1. If a magazine has consistently published different covers for subscribers and newsstand buyers and GfK MRI consistently uses the dominant cover in the IS study, GfK MRI will release indices for all issues.
- 2. If a magazine has a split cover between newsstand and subscriber and the only difference is the number of cover stories listed on the cover (or something minimal like this example), we will release that issue in IS. Release is predicated on the notion that modest differences in the covers will have an immaterial impact on IS audiences.
- 3. If a magazine that generally does not have a split cover between newsstand and subscriber publishes a split cover for a particular issue, GfK MRI will assess release in the following manner:
 - a. If newsstand is less than 5% of total circulation, GfK MRI is likely to release in IS unless there is some overtly compelling reason to suppress the IS index. We have not encountered any situations along these lines. Release is predicated on the notion that such a small percentage of newsstand circulation has a minimal impact on the overall index.
 - b. If the newsstand circulation is greater than 5%, we will review the cover to determine the magnitude of the differences between the covers. In most cases, if the covers are not radically different, we will release the IS index. Examples of this issue includes having the same personality on the covers, but in different poses, having the same cover stories but somewhat different cover executions, or having dramatically different covers. Once again, the similarity of the cover execution supports a release decision.

In either event, GfK MRI must use the subscriber cover as the stimulus.

- 4. If a magazine has a random split cover (or covers), GfK MRI will assess each case, individually.
 - a. If the index exceeds 100 for that issue, we will release in IS. The rationale is straightforward: We must presume that showing only one of a number of covers inherently depresses the IS index. In the event the index is over 100, we believe the index, while likely understating issue performance, clearly shows that the specific issue outperforms the average issue.
 - b. If the index is equal to or below 100 for that issue, we will not release in IS. Instead, we will replace the index with an average index of 100 and use the national study profile of the magazine to "composition target" that issue. Since we presume that indices are inherently depressed, we are unable to estimate the "true" index of any issue with a calculated index below 100. In order not to assume the specific issue's "true" index was below 100, we use the average index as the best estimate of issue performance.

References

Baim, Julian, Martin Frankel, Michal Galin, Joseph Agresti, and Kerry Zarnitz. "Measuring Issue Specific Audiences." Paper presented at the Worldwide Readership Research Symposium, Vienna, Austria: 2007.

Petric, Irena, and Marion Appel. "The Readership Currency: Dutch Design." Paper presented at the Worldwide Readership Research Symposium, Vienna, Austria: 2007.

Petric, Irena, and Marion Appel. "New Learnings from Specific Issue Readership: Accumulation of Readership and Calculation of Average Issue Readership." Paper presented at the Worldwide Readership Research Symposium, Valencia, Spain: 2009.