Reliability and validity of the Alcohol Marketing Assessment Rating Tool (AMART)

Jonathan K. Noel
Johnson & Wales University, USA and University of Connecticut School of Medicine, USA

Thomas F. Babor
University of Connecticut School of Medicine, USA

Katherine Robaina
University of Connecticut School of Medicine, USA

Abstract

Objective: Current national responses to the monitoring and prevention of alcohol marketing to vulnerable populations, such as youth, typically rely on partial bans, co-regulation and industry self-regulation. Efforts to evaluate compliance with existing regulations are often cumbersome and resource intensive. We sought to develop a short questionnaire to assess compliance of alcohol advertising to existing alcohol industry self-regulated marketing codes. **Methods**: Questions for the Alcohol Marketing Assessment Rating Tool (AMART) were taken from a longer rating instrument that was originally developed to detect violations of a self-regulated alcohol marketing code. Secondary analysis of data from three studies of alcohol advertisements was performed to determine the reliability and validity of AMART at detecting code violations, using the longer form as the criterion. One sample was used as an exploratory set. Two samples were used to validate the questionnaire. **Results**: In the exploratory set, the reliability of AMART was considered near perfect (kappa = 0.92). Sensitivity was 97%, and specificity was 100%. Positive predictive value (PPV) was perfect and negative predictive value (NPV) was approximately 90%. In the validation sets, reliability was considered substantial to near perfect (kappa = 0.71–0.94). Specificity and PPV remained perfect, and NPV was 86%–90%. **Conclusion**: The AMART is a reliable tool to detect violations of a self-regulated marketing code in alcohol advertising. It significantly decreases the amount of resources needed to evaluate a finite number of advertisements.

Corresponding author:
Jonathan K. Noel, Department of Health Science, Johnson & Wales University, 8 Abbott Park Place, Providence, RI 02903, USA.
Email: jknoel@hotmail.com

Keywords

Alcohol marketing is widespread in most parts of the world. Multiple communication channels are used to promote alcohol brands, including youth-oriented radio, television, sports events, popular music concerts, websites, social media, mobile phone apps and product placements in movies and TV shows (Noel, Babor, & Robaina, 2017). Exposure to alcohol marketing is associated with the earlier initiation of drinking and binge drinking in adolescents (Jernigan, Noel, Landon, Thornton, & Lobstein, 2017), and for that reason, alcohol marketing regulations can be justified on the grounds of public health, public safety and human rights (Babor, Robaina, Noel, & Ritson, 2017; Chapman, 2017). Current national responses, most of which rely extensively on partial bans, co-regulation, and industry self-regulation, are insufficient in fulfilling the public health mission to prevent youth exposure to alcohol marketing (Noel, Lazzarini, Robaina, & Vendrame, 2017; Pantani et al., 2016).

Where alcohol advertising is not regulated by legal statute, there is often an informal agreement between a governing body and the alcohol industry to conform to minimal standards, which are created, implemented, and enforced by the alcohol or advertising industry. The primary self-regulators of alcohol marketing tend to be industry trade associations, such as spiritsEUROPE, or industry social aspects public relations organisations (SAPROs). In 2011, the largest industry SAPRO, the International Alliance for Responsible Drinking (IARD), published a model self-regulated alcohol advertising code that was endorsed by 10 of the largest transnational alcohol producers (IARD, 2011). Known as the Guiding Principles, this model code applies to all alcohol marketing communications published in all media in countries where self-regulation is the primary means of alcohol marketing regulation.

While other industry voluntary codes of practice have been developed (e.g., spiritsEUROPE, the Brewers of Europe), the most common provisions across all codes are included in the Guiding Principles.

Independent efforts to assess the alcohol industry's compliance with self-regulated alcohol marketing codes have been made since 1991 (Noel, Babor, & Robaina, 2017). In a review of 19 studies performed in 19 countries, code violation rates ranged from 0% to 100% depending on the medium studied and the advertisement sampling procedure. For example, in an evaluation of advertising appearing on British television, a majority of participants rated 86% of the alcohol advertisements as violating the United Kingdom's (UK) Code of Broadcast Advertising (Searle, Alston, & French, 2014). Similarly, 74% of websites for UK alcohol brands contained a violation of this code (Gordon, 2011). Of the 19 studies appearing in the review, 15 concluded that self-regulation was ineffective at preventing potentially harmful content from appearing in alcohol advertising, and no study recommended self-regulation as a method to restrict potentially harmful alcohol advertising content (Noel, Babor, & Robaina, 2017).

Common violations of the self-regulatory codes identified in a systematic review included promotions of excessive alcohol consumption; associating alcohol with social, financial, physical, or sexual success; and targeting of minors under the legal minimum purchase age (Noel, Babor, & Robaina, 2017). In a study of advertisements collected from Germany, the Netherlands, and the UK (Winpenny et al., 2012), music, human characters, and technological effects were found to be particularly attractive to youth.

While previous research has consistently demonstrated that alcohol marketing self-regulation is ineffective, these findings have yet

to have a major impact on the regulation of alcohol marketing, perhaps because the research is concentrated in only a few countries, existing studies often evaluated a small number of unique advertisements, and recent evaluations have focused on large media events that may not be representative of alcohol marketing throughout the year (Noel, Babor, & Robaina, 2017).

A major reason for the limitations of existing alcohol marketing research is the lack of an efficient, standardised, and psychometrically sound procedure to identify violations of the self-regulated marketing codes, particularly the content guidelines. In an attempt to develop a standardised and valid instrument that could be used for research and public health surveillance, a procedure was developed using a questionnaire containing 29 items to assess compliance with the 1997 US Beer Institute code (Babor, Xuan, & Proctor, 2008). Because the procedure uses the Delphi method, which requires two rounds of ratings to increase inter-rater reliability, 58 responses were required to evaluate each beer advertisement. Searle et al. (2014) used 40 items to compare UK alcohol advertisements to the UK code. In two other studies, a 37-item questionnaire was used to evaluate alcohol advertising during the 2014 FIFA World Cup tournament in eight countries and Facebook advertising published in conjunction with the 2015 NFL Super Bowl (Noel, Babor, Robaina, Feulner, et al., 2017; Noel & Babor, 2017b). The FIFA and Facebook studies also incorporated a version of the Delphi method.

Although these studies have produced valuable findings, the time needed to evaluate each alcohol advertisement under review has limited the wider application of these rating techniques, including independent evaluations of the effectiveness of self-regulated alcohol advertising codes. Within this context, we sought to develop a short questionnaire that could be rapidly implemented to assess alcohol advertising with results comparable to the longer procedures previously described (Babor et al., 2008; Babor, Xuan, Damon, & Noel, 2013; Noel, Babor, Robaina, Feulner, et al., 2017).

A secondary analysis of data from three samples of alcohol advertisements was performed. One sample was used in the development of the short-form questionnaire. Two samples were used to validate the questionnaire.

Methods

Advertisement selection

Advertisements from three separate alcohol advertisement evaluation studies were included in this study. These advertisements originated from the Monitoring Alcohol Marketing in Africa (MAMPA) study (282 advertisements) (Robaina, Babor, & Noel, 2017), an evaluation of alcohol advertising shown during the 2014 FIFA World Cup Tournament (41 advertisements) (Noel, Babor, Robaina, Feulner, et al., 2017), and an evaluation of alcohol advertising posted on Facebook around the 2015 NFL Super Bowl (50 advertisements) (Noel & Babor, 2017b). For the current study, the FIFA advertisements were considered the exploratory set, selected because they were published on a single, traditional advertising medium. The alcohol advertisements evaluated from the Facebook study, a non-traditional advertising medium, and MAMPA, which consisted of multiple advertising media, were used as the validation sets. Combined, these samples represent alcohol advertising from eight different countries in North America and Africa and include five different media.

The FIFA project used a total survey approach for advertisement collection. Investigators were instructed to record each match, as well as the half-time show and at least 30 minutes of the pre- and post-game shows associated with each match (Noel, Babor, Robaina, Feulner, et al., 2017). All unique alcohol advertisements were then abstracted and stored in separate video files. The Facebook study collected a sample of 50 alcohol advertisements posted on Facebook by the official alcohol sponsors of the 2015 National Football League's Super Bowl one month before and after the event, which were randomly selected for

evaluation (Noel & Babor, 2017b). The Super Bowl was selected as an anchor point because it was the largest media event of 2015 and Facebook was selected because it is the most popular social networking platform among teens and young adults. The two-month advertisement collection window was selected to ensure all advertisements relevant to the event were included in the sample.

Alcohol advertisement collection varied by country within the MAMPA study due to variations in local resources (Robaina et al., 2017). For television and radio advertisements, at least one national television and radio station were monitored during weeknight and weekend hours. Unique advertisements were abstracted from these recordings and saved in individual files. Print advertisements were collected by monitoring daily newspapers and monthly magazines. Outdoor advertising was collected using systematic searches in urban and rural locations. While a large number of alcohol advertisements were collected during the MAMPA project, due to resource restrictions, the advertisment sample may not be representative of all alcohol advertising in the participating countries.

Violation ratings

In each study (i.e., FIFA, Facebook, MAMPA), a panel of expert raters determined whether each alcohol advertisement was compliant with the Guiding Principles. The FIFA study panel consisted of 14 experts; the Facebook study panel contained 11; and eight members were on the MAMPA panel. Raters were identified and defined as experts because they had experience conducting substance abuse, mental health, and/or public health services and research or had other expertise relevant to protecting vulnerable populations. Raters were invited to participate by email. Participation was restricted to those who affirmatively responded to the invitation.

Although each study used a unique panel of expert raters, the advertisements were rated using the same technique. In each case, the Delphi technique, which is an iterative process designed to build group consensus regarding policy-relevant issues, was employed (Hasson, Keeney, & McKenna, 2000; Powell, 2003). Here, a two round rating procedure was used. In round one, each expert rater evaluated the advertisements independently. In round two, each rater evaluated the advertisements independently again, but they were provided with the mean item-level ratings for each advertisement and any comments about the advertisements provided by the other expert raters during round one.

Alcohol Marketing Assessment Rating Tool (AMART) development

Questions for the AMART were taken from a longer, 37-item, questionnaire that was originally developed to detect code violations of the 1997 and 2006 US Beer Institute code (Babor, Xuan, Damon, & Noel, 2013; Babor et al., 2008). While the MAMPA study used a slightly longer ad rating questionnaire (i.e., 40 items), the additional items were relevant to national policies and not the Guiding Principles (Robaina et al., 2017). Three types of questions were used in the full questionnaire employed on the FIFA, Facebook, and MAMPA studies. First, raters were instructed to indicate the extent to which they agreed or disagreed with a list of 35 statements using a five-point Likert scale (e.g., "This advertisement portrays abstinence or moderate consumption in a negative way"). Second, one age perception question asked raters to identify the approximate age of the youngest actor/actress in the ad (i.e. "How old do you think the youngest person in this advertisement is?"). Third, one question asked raters to indicate the amount of drinking they perceived taking place in the ad (i.e. "How many drinks do you estimate this person is likely to consume in the situation shown in the advertisement?").

These questions have previously been found to be reliable (Babor et al., 2008) and were subsequently adapted to detect violations of the

Guiding Principles. The full rating questionnaire was used to evaluate alcohol advertisements in each of the previous studies (Supplemental Table 1). Although the questionnaire has been translated into Spanish and Portuguese (Noel, Babor, Robaina, Feulner, et al., 2017), only English-language items, and advertisements rated by the English-language items, were used in this analysis.

Potential AMART question sets were identified in three ways (Supplemental Table 2). First, three experts with experience conducting multiple evaluations of alcohol advertising selected items for the AMART. Items were selected based on their ability to detect content that may be harmful to populations considered to be vulnerable to alcohol marketing messages, such as youth. Although there are multiple sub-guidelines, the Guiding Principles contain five general content guidelines. Items were also chosen to ensure that each guideline was represented by at least one item. Second, items that were most often violated were identified and evaluated. Third, 50 random combinations of questions were identified and evaluated.

Statistical analysis

In each study, the true code violation rate, as determined by the full questionnaire, was calculated using a scoring criterion (Babor, Xuan, Damon, & Noel, 2013; Babor et al., 2008; Noel & Babor, 2017b) that specifies that if a rater indicated any item-specific violations among the items associated with the same sub-guideline, a sub-guideline violation was indicated. If any sub-guidelines associated with the same guideline were violated, a guideline violation was indicated. When 50% or more of the expert raters identified the same guideline violation, the advertisement was coded as containing a violation. A modified version of this criterion was used for the AMART. If a rater indicated any item-specific violation, the advertisement was coded as containing a rater-specific violation. When 50% or more of the

raters agreed the advertisement contained at least one item-specific violation, the advertisement was coded as containing a violation. Item-specific violations were defined as scores of ≥ 4 for Likert scale items, <21 for the age of the youngest actor or actress, and ≥ 5 for the perceived number of alcoholic drinks consumed.

The reliability of the AMART to detect code violations compared to the long form questionnaire was tested using Cohen's kappa. During the exploratory phase, a question set with a kappa ≥ 0.8, which is considered substantial or better (Landis & Koch, 1977), was specified. Validity of the potential AMART question sets was assessed by calculating the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) to detect code violations compared to the full questionnaire. The question set with the greatest reliability and validity was selected for evaluation in the validation sets.

The reliability of each item selected for potential inclusion in the AMART was assessed using (2, k) intra-class correlations (ICC). This calculation was performed within each separate sample of advertisements. An item was retained in the AMART if it had an ICC ≥ 0.6, which is considered moderate or better reliability (Landis & Koch, 1977), across all three advertisement sets. Statistical analysis was conducted using SPSS for Windows Version 24 (Armonk, NY: IBM Corp.).

Results

Nine items were initially selected for inclusion in the AMART (Table 1). The first seven questions were Likert-scale items where raters indicated their level of agreement or disagreement with a statement. The eighth question asked how old the youngest person in the advertisement was and required a numerical response in years. The ninth question asked raters to estimate how much alcohol the youngest person was likely to drink based on the situations depicted in the advertisement. This question required a numerical response that was

Table 1. Inter-rater reliability of Alcohol Marketing Assessment Rating Tool items across studies.

Question	Guideline covered	ICC (FIFA)	ICC (Facebook)	ICC (MAMPA)
This advertisement shows situations where people are drinking an alcoholic beverage excessively, or otherwise encourages immoderate consumption.	2	0.94	0.94	0.99
This advertisement uses symbols, language, music, gestures, or cartoon characters that are associated with or are intended to appeal primarily to persons below legal purchase age.	4	0.91	0.91	0.99
This advertisement suggests that drinking leads to an exciting adventurous life.	5	0.97	0.95	0.99
This advertisement associates social, professional, mental, educational, athletic or financial success with drinking the alcohol product.	5	0.95	0.92	0.99
This advertisement shows or suggests the use of an alcohol product before or during activities requiring sobriety or a high degree of alertness or coordination, such as driving an automobile, operating machinery, boats, working in a hazardous situation, playing sports, etc.	3	0.91	0.93	0.98
This advertisement shows illegal activity.	1	0.93	0.85	0.99
The advertisement depicts or appears to be addressed to at-risk groups, such as pregnant women, women of childbearing age, people under legal purchase age, college students, ethnic minorities, alcoholics, or other vulnerable groups.	3	0.90	0.89	0.99
How old do you think the youngest person in this advertisement=is?	4	0.98	0.99	0.88
How many drinks do you estimate this person is likely to consume in the situation shown in the advertisement?	2	0.96	0.95	0.88

ICC = intra-class correlations; FIFA = Fédération Internationale de Football Association; MAMPA = Monitoring Alcohol Marketing Practices In Africa.

measured in standard drinks of alcohol. These nine items were selected because of their relevance to protecting vulnerable populations from potentially harmful advertising content, as well as their public health implications (e.g., preventing excessive alcohol consumption) and coverage of all guidelines in the Guiding Principles.

At least one item covered each guideline of the Guiding Principles, with guidelines 2 (prohibiting depictions of irresponsible consumption), 3 (suggestions that alcohol has health benefits), 4 (targeting of minors), and 5 (social, physical, and sexual consequences of alcohol use) being covered by two items each. The inter-rater reliability of the questions was substantial to near perfect across studies (ICCs ≥

0.85) (Table 1). The reliability of the AMART was considered substantial (kappa = 0.92) and was above the desired cut-off point (Table 2). Sensitivity was 97%, and specificity was 100%. PPV was perfect and NPV was 88%. The expert selected items outperformed all other potential questions sets (Supplemental Table 3), and based on these results, the nine expert selected items were applied to both validation sets of advertisements.

When applying the AMART to the Facebook advertisements, reliability was considered substantial, and sensitivity, specificity, PPV, and NPV were similar to for the FIFA advertisements (Table 2). Reliability of the AMART among the MAMPA advertisements was lower

Table 2. Reliability, validity, and code violation rates of the Alcohol Marketing Assessment Rating Tool (AMART).

Variables	Full questionnaire	AMART	*Kappa*	Sensitivity	Specificity	PPV	NPV
FIFA			0.92	0.97	1.00	1.00	0.88
Total number of advertisements	41	41					
Numbers of advertisements with a code violation	34	33					
Code violation rate (%)	82.9	80.5					
Facebook			0.94	0.98	1.00	1.00	0.90
Total number of advertisements	50	50					
Numbers of advertisements with a code violation	41	40					
Code violation rate (%)	82.0	80.0					
MAMPA			0.71	0.63	1.00	1.00	0.86
Total number of advertisements	282	282					
Numbers of advertisements with a code violation	78	49					
Code violation rate (%)	27.7	17.4					

PPV = positive predictive value; NPV = negative predictive value; FIFA = Fédération Internationale de Football Association.

but still considered moderate to substantial. Although sensitivity was decreased, specificity and PPV remained perfect and NPV was still considered excellent.

Discussion

The adoption of the Guiding Principles as an international model code for the self-regulation of alcohol advertising has provided researchers and the public health community with the ability to evaluate alcohol advertising internationally and in diverse media in a truly comparable way. Efforts to complete such evaluations by independent third parties have been limited largely due to the resource-intensive nature of evaluating alcohol advertising for regulatory compliance. The development of the AMART may reduce this burden. By reducing the number of rating questions by approximately 75%, significantly fewer resources will be needed to rate a sample of advertisements, or alternatively more advertisements can be rated using the same amount of resources. Moreover, because self-regulation is often seen as a minimum regulatory standard for alcohol advertising, advertisements in countries that have statutory restrictions on the content of alcohol advertising can also be assessed. That is, if an alcohol advertisement fails to comply with the Guiding Principles based on the AMART, the advertisement may also be unlikely to comply with existing national laws.

Because the AMART has perfect PPV and a high NPV, it may be an ideal screening tool. The AMART may also be ideal because it is a conservative measure of code violations. Indeed, there were no false positives in either the exploratory set or the validation sets of advertisements when using the AMART. Furthermore, the AMART demonstrates that a relatively small number of questions can be used to detect violations of the Guiding Principles even though the code contains numerous sub-guidelines. This may occur because the majority of violations occur in a small number of

guidelines and sub-guidelines, whereas other guidelines are rarely violated because of their esoteric nature (e.g., depictions of Santa Claus). Conversely, because the AMART does not cover all the sub-guidelines in the Guiding Principles, the implementation of this questionnaire may result in an unintended consequence. Regulatory compliance among the sub-guidelines that the AMART does not expressly cover may decrease even as regulatory compliance among the sub-guidelines that the AMART does cover may increase. Because of this, a random sub-sample of advertisements should continue to be evaluated using the full rating questionnaire to ensure full regulatory compliance.

The speed at which the AMART can be completed should aid in its adoption and implementation. We estimate that it will take 1–2 minutes to complete the AMART for any given alcohol advertisement. If the Delphi method is used, then 2–4 minutes will be needed, although this additional step may not be necessary if the AMART is used solely for screening purposes.

As implied above, the AMART could be an integral part of an independent surveillance system to monitor alcohol marketing practices. The reasons for such a system are two-fold. First, while most of the extant research has demonstrated that self-regulation is ineffective (Noel, Babor, & Robaina, 2017), substantially greater quantities of data are needed from more countries in order to provide a basis for national and international policy. Second, the current compliance systems implemented by the alcohol industry have not effectively prevented harmful content from appearing in alcohol advertisements and are fraught with conflicts of interest (Noel & Babor, 2017a). Complaints against an alcohol advertisement that is considered to contain a code violation by experts are often dismissed, and when alcohol advertisements are removed from the marketplace, this typically occurs after a marketing campaign has concluded. Furthermore, the raters employed by the alcohol industry do not have experience protecting vulnerable populations, and often use the

reasonable adult standard that is not applicable to youth (Noel, Lazzarini, et al., 2017).

Because the parent questionnaire has been developed and implemented for expert raters, we recommend that the AMART be deployed within similar populations pending additional research using laypersons. Experts could include research, public health practitioners, and marketing executives who work for the alcohol industry. Moreover, we recommend that the AMART be used to evaluate alcohol advertisements from all media, whether traditional formats (e.g., television, print, radio) or new formats (e.g., websites, apps, social media).

Strengths and limitations

A strength of this study is the diverse nature of the alcohol advertising that was originally assessed for compliance with the Guiding Principles. Combined, the advertisements were published on television, on radio, in print, on social media, and in multiple outdoor, public settings. The advertisements also originated from multiple countries in North America (i.e., Canada, the USA) and Africa (i.e., Ghana, Kenya, Malawi, Namibia, Nigeria, and Uganda).

There are several limitations to this study. First, the AMART was not applied to non-English-language advertising. Further research will be needed to validate the full rating questionnaire and the AMART in other languages. Second, the study results may be biased because of the reliance on experts, rather than representatives of vulnerable groups, to complete the advertisement evaluations. However, prior research suggests such biases are non-existent. Comparisons of alcohol advertisement evaluations indicate that expert ratings are equivalent to, or more conservative than, ratings from community non-experts (Babor, Xuan, & Damon, 2013; Vendrame et al., 2015).

The AMART may only perform well in samples of advertisements where violations of the Guiding Principles are prevalent. However, the AMART may still produce acceptable, if not ideal, reliability and validity within

advertisement sets with a low prevalence of code violations, as seen in the MAMPA advertisement set. Further testing of the AMART is needed in other low-prevalence code violation advertisement sets. However, even when the prevalence of code violations is low, the AMART may be useful to rule in code violations because specificity for code violations will likely remain high. That is, each advertisement that contains a code violation according to the AMART will likely contain a violation according to the full questionnaire regardless of the advertisement's origin. Third, because the AMART was not tested separately, answers to the items within the AMART may have been influenced by the surrounding questions. Further research is needed to determine whether raters' answers to the AMART items are context dependent. Finally, the AMART is only partially able to provide precise information on the specific nature of a code violation, which may be needed for research purposes or to file a complaint.

Conclusion

The AMART is a nine-item questionnaire that can reliably determine regulatory compliance of alcohol advertising to an international marketing code that has been adopted by many segments of the alcohol industry. The small number of questions compared to the standard version will significantly decrease the amount of resources needed to evaluate a finite number of advertisements or significantly increase the number of advertisements that could be evaluated with finite resources. The AMART can be readily implemented by both alcohol marketers and independent researchers, although a sub-sample of advertisements should continue to be evaluated by the full rating questionnaire.

Data access

Data used in this study can be accessed through correspondence with the lead author.

Supplementary material

The supplementary material is available online with the article.

References

Babor, T. F., Robaina, K., Noel, J. K., & Ritson, E. B. (2017). Vulnerability to alcohol-related problems: A policy brief with implications for the regulation of alcohol marketing. *Addiction*, *112*(S1), 94–101.

Babor, T. F., Xuan, Z., & Damon, D. (2013). A new method for evaluating compliance with industry self-regulation codes governing the content of alcohol advertising. *Alcoholism: Clinical and Experimental Research*, *37*, 1787–1793.

Babor, T. F., Xuan, Z., Damon, D., & Noel, J. (2013). An empirical evaluation of the US Beer Institute's self-regulation code governing the content of beer advertising. *American Journal of Public Health*, *103*, e45–e51.

Babor, T. F., Xuan, Z., & Proctor, D. (2008). Reliability of a rating procedure to monitor industry self-regulation codes governing alcohol advertising content. *Journal of Studies on Alcohol and Drugs*, *69*, 235–242.

Chapman, A. R. (2017). Can human rights standards help protect children and youth from the detrimental impact of alcohol beverage marketing and promotional activities? *Addiction*, *112*(S1), 117–121.

Gordon, R. (2011). An audit of alcohol brand websites. *Drug and Alcohol Review*, *30*(6), 638–644.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

International Alliance for Responsible Drinking (IARD). (2011). *Guiding principles: Self-regulation of marketing communications for beverage alcohol*. Retrieved from http://www.iard.org/wp-content/uploads/2016/01/Guiding-Principles.pdf

Jernigan, D., Noel, J., Landon, J., Thornton, N., & Lobstein, T. (2017). Alcohol marketing and youth alcohol consumption: A systematic review of longitudinal studies published since 2008. *Addiction*, *112*(Suppl 1), 7–20.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Noel, J. K., & Babor, T. F. (2017a). Does industry self-regulation protect young persons from exposure to alcohol marketing? A review of compliance and complaint studies. *Addiction*, *112*(Suppl 1), 51–56.

Noel, J. K., & Babor, T. F. (2017b). Predicting regulatory compliance in beer advertising on Facebook. *Alcohol and Alcoholism*, *52*(6), 730–736.

Noel, J. K., Babor, T. F., & Robaina, K. (2017). Industry self-regulation of alcohol marketing: A systematic review of content and exposure research. *Addiction*, *112*(Suppl 1), 28–50.

Noel, J. K., Babor, T. F., Robaina, K., Feulner, M., Vendrame, A., & Monteiro, M. (2017) Alcohol marketing in the Americas and Spain during the 2014 FIFA World Cup tournament. *Addiction*, *112*(Suppl 1), 64–73.

Noel, J., Lazzarini, Z. R., Robaina, K., & Vendrame, A. (2017). Alcohol industry self-regulation: Who is it really protecting? *Addiction*, *112*(Suppl 1), 57–63.

Pantani, D., Peltzer, R., Cremonte, M., Robaina, K., Babor, T., & Pinsky, I. (2017). The marketing potential of corporate social responsibility activities: The case of the alcohol industry in Latin America and the Caribbean. *Addiction*, *112*(S1), 74–80.

Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, *41*(4), 376–382.

Robaina, K., Babor, T., & Noel, J. (2017). *Evaluating compliance with alcohol industry self-regulation in seven countries in Africa: An external evaluation of the MAMPA (Monitoring Alcohol Marketing Practices In Africa) project*. FORUT – Campaign for Development and Solidarity, Gjøvik, Norway. Retrieved from http://www.add-resources.org/getfile.php/3720760.994.cabbuuassq/MAMPA_Final+Report+2016+email+version.pdf

Searle, R., Alston, D., & French, D. P. (2014). Do UK television alcohol advertisements abide by the Code of Broadcast Advertising rules regarding the portrayal of alcohol? *Alcohol and Alcoholism*, *49*(4), 472–478.

Vendrame, A., Silva, R., Xuan, Z., Sparks, R., Noel, J., & Pinsky, I. (2015). Self-regulation of beer advertising: A comparative analysis of perceived violations by adolescents and experts. *Alcohol and Alcoholism*, *50*, 602–607.

Winpenny, E., Patil, S., Elliott, M., Villalba van Dijk, L., Hinrichs, S., Marteau, T., & Nolte, E. (2012). *Assessment of young people's exposure to alcohol marketing in audiovisual and online media*. RAND Europe, Cambridge, UK. Retrieved from: http://ec.europa.eu/health/alcohol/docs/alcohol_rand_youth_exposure_marketing_en.pdf