


Please cite the Published Version

Ashraf, R  and Merunka, D (2017) The use and misuse of student samples: An empirical investigation of European marketing research. *Journal of Consumer Behaviour: an international research review*, 16 (4). pp. 295-308. ISSN 1472-0817

DOI: <https://doi.org/10.1002/cb.1590>

Publisher: Wiley

Version: Accepted Version

Downloaded from: <https://e-space.mmu.ac.uk/629350/>

Usage rights:  In Copyright

Additional Information: This is an Author Accepted Manuscript of an article published in *Journal of Consumer Behaviour*.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

The use and misuse of student samples: An empirical investigation of European marketing research

ROHAIL ASHRAF^{1*} and DWIGHT MERUNKA^{2,3}

¹Suleman Dawood School of Business, Lahore University of Management Sciences (LUMS), Lahore, Pakistan

²Aix-Marseille University, Cergam, Aix-Marseille Graduate School of Management, Aix en Provence, France

³Kedge Business School, Marseille, France

ABSTRACT

In spite of the five decade old debate on the merits of student samples, to date, no systematic review of the practice is undertaken. The need for such a review is warranted considering the impasse in the debate and inconsistencies among scholars in their approach to student sample usage. This paper thus presents a systematic review of student sample usage in European Marketing Research (EJM, IJRM and IMR during 2005–2014, inclusive) to highlight existing reporting practices, identify sub-domains of marketing where the usage is more prevalent and to report best practices. Results demonstrate that 99 (19.96 per cent) papers making generalization claims used student samples exclusively, had inconsistent reporting practices (e.g. demographic profile, limitations) and demonstrated trivial concern (e.g. bias estimation treatment, identification of moderators) for the implications of student sample usage on their study findings. In addition, 11 clusters representing sub-domains of marketing research and where the practice is prevalent are identified. These clusters provide novel direction to the debate on student sample usage by framing it away from the broader discipline of marketing and bringing it closer to the interests of scholars, i.e. linking it to sub-domains of marketing research. Finally, best practices related to student samples usage are reported to help academicians enhance the validity of their findings. Copyright © 2016 John Wiley & Sons, Ltd.

“We have developed an impressive corpus of scientific knowledge and, indeed, have learned a great deal from studying college sophomores in the laboratory. But it may be appropriate to be somewhat more tentative about the portrait of human nature we have developed from this database”

(Sears, 1986, p. 527)

Marketing research relies heavily on student samples for empirical testing, an issue about which social scientists have been debating merits and dangers for over five decades. Scholars against the use of student samples argue that students have a demographic and psychographic profile which is not representative or even close to that of other consumers or individuals. On the contrary, proponents contend that if the field of medicine can bring considerable good to humanity by using animals then why would not social scientists through the use of student samples? A more balanced view is that of Sears (1986, p.515) who emphasizes that every field has methodological idiosyncrasies; it is however important that “researchers trust that they have a reasonably good grasp of the biases introduced by their own particular methodological proclivities and that they can correct their conclusions for whatever biases are present.” This refers to the notion that the use of student samples per se does not automatically invalidate findings. However, researchers are expected to understand the biases introduced and to present adequate evidence about the solutions used to mitigate biases which might influence the results.

The debate on the use of student samples is better understood when contrasting two types of studies: “effect application

studies” and “theory application studies”. In effect application studies, the objective is to estimate effect sizes between constructs which are then generalized to situations other than those studied. Winer (1999) explains that quite often effect application studies may share characteristics similar to theory application studies. However, they primarily differ in that effect application studies aim to find results applicable in the real world. Empirical findings demonstrate that effect application studies should not be conducted with student samples (Peterson, 2001) or at least not with student samples alone. In “theory application studies”, the objective is to test scientific theories such as the influence of attitudes on behavioral intent (e.g. Bagozzi, 1992). Opinions concerning the use of student samples in such studies remain divergent. The proponents argue that because the objective is to develop theories dealing with fundamental human behavior, students as humans do represent a legitimate sample choice (e.g. Calder *et al.*, 1981; Lucas, 2003). In turn, the opponents argue that the major concern is not the legitimacy of student sample usage per se, but rather their unabated use without regard to the constraints which student demographic and psychological characteristics impose on theory testing procedures (e.g. Lynch, 1982, 1983; Bello *et al.*, 2009). Therefore, findings about consumer behavior derived solely from student samples may be suspect because of the peculiarities of their constituents.

Most of the debate on the merits of student sample usage in marketing research has come either from the editors of leadings journals (e.g. Ferber, 1977; Bello *et al.*, 2009), or from authors arguing the concept of external validity (e.g. Wells, 1993, 2001; Lynch, 1999; Winer, 1999), or from studies demonstrating empirical differences in results between student and non-student samples (e.g. James and Sonner, 2001; Peterson, 2001). However, in spite of these efforts, the issue remains unresolved with some scholars questioning

*Correspondence to: Rohail Ashraf, Suleman Dawood School of Business, Lahore University of Management Sciences (LUMS), Sector U, DHA, Lahore Cantt. 54792, Lahore, Pakistan.
E-mail: rohail.ashraf@lums.edu.pk

the practice while others embrace it. Most consider it as a minor problem to their work and if pressed by reviewers, they acknowledge the fact in the limitations section of their manuscript. Considering this impasse, we propose that a viable way forward is to move the debate's focus from the broader marketing discipline to more specific domains of study. This change of focus will help researchers and reviewers to specifically question the practice in their area of interest rather than posing general questions (e.g. heterogeneity and non-representativeness) which achieved little success both in terms of improving the practice and in terms of reducing usage frequency. Although such recommendations have been put forward by other researchers (Lynch, 1999; Sears, 2008), to date, no systematic review of student sample usage has been conducted to set a direction. We believe such a review is warranted as it will help develop a deeper understanding of how student samples are currently being used in marketing research and help scholars working in sub-areas of the marketing discipline to be more concerned about this practice.

Thus, the objective of this paper is threefold: (i) review the debate on student sample usage and put the issue in perspective; (ii) analyze existing practices of student sample usage and identify sub-areas of marketing research where student samples are often used; and (iii) highlight best practices in sample selection procedures. To provide an integrated view of the debate, the manuscript first examines (a) the demographic and psychological differences between students and consumers, as it forms the bases for (b) the generalization problem for studies using student samples. The discussion is further extended (c) by presenting a systematic review of studies using student samples in leading European marketing journals (i.e. *European Journal of Marketing*, *International Journal of Marketing Research* and *International Marketing Review*) in order to (d) identify major patterns of reporting and to (e) highlight sub-areas in the marketing discipline where student sample usage is more prevalent. Finally, the article concludes (f) with best practices in sample selection as a way to improve the modus operandi of student sample utilization while striking a balance between the convenience offered by student sample usage and the necessity to report generalizable research results.

STUDENT POPULATION CHARACTERISTICS

The opponents of student sample usage argue that students have demographic and psychological profiles which differ from the more heterogeneous adult population. Thus, they cannot be considered representative of consumers in general. We discuss these two major sets of characteristics (demographic and psychological) and report areas of significant differences between students and non-student consumers.

Demographic characteristics

Census data provides empirical evidence about the differences between student and non-student population characteristics. We illustrate these differences for the five most populated European countries (i.e. Germany, France, UK, Italy and Spain). For Germany, Italy and Spain, data was retrieved from the

United Nations Statistic Division (UNSD, 2011), for UK from the Office of National Statistics (ONS, 2011), and for France from the National Institute of Statistics and Economic Studies (INSEE, 2012).

A vast majority of university students are between 20 and 24 years of age. This age bracket represents between 5.1 per cent and 8.4 per cent of the population in major European countries, a population density similar to all subsequent 5-year age intervals. However, similarities end here. Table 1 shows that on other demographic indicators (employment and marital status), the 20–24 age bracket represent a particular transitory life stage where one is exiting from early life stages and is entering into a more stable adulthood phase.

Obviously, university students differ from adults in their professional situations. Most individuals, even in the 20–24 age group are economically active. In economically well-off European countries (Germany, France and UK), between 50 and 69 per cent of individuals in this age group are employed, whereas in Spain and Italy, percentage of employment is 30 and 38 respectively. The transition from studentship to economic activity continues in older age groups and stabilizes after 30 (Germany, France, UK \approx 80 per cent; Spain and Italy \approx 70 per cent). Similar transitory changes are also observed in family lifestyle. Most individuals in the 20–24 age group are bachelors (>93.6 per cent) and start to get married in the following age bracket groups. The percentage of married people increases substantially in the 25 to 29-year age group and continues to rise during adulthood to reach approximately 70 per cent in the 50–54 age group. Therefore, if employment or family life/marital status are an issue, university students cannot be a basis for inferences regarding other individuals within the same age group or in other life stages.

University students differ from other adult population on age, professional activity and marital status (family experiences). However, many other characteristics are seldom observed with university students such as retirement, raising children, owning a home, medium and high income levels, health issues, etc. (Andersen *et al.*, 2010). These characteristics should not only impact consumption patterns and situations in many product categories (e.g. purchase of luxury goods, club memberships, baby and children clothes, car size, vacations, housing, mortgages, healthcare, etc.) but, more importantly, they impact ones psychological development and consequently perspective about different life situations and decision scenarios.

Psychological characteristics

A student population is unique in terms of psychological development. Carlson (1971, p. 212) proposes that “students are ‘unfinished personalities’ in a relatively early adult life stage”. The demographic data presented above substantiates this as students are either undergoing or are on the verge of major transitions that will redefine their lives with new sets of responsibilities. They show many psychological differences from other adults (Sears, 1986):

1. *Weak self-definition*: Self-definition deals with a person's values, preferences, abilities and emotions which gel

Table 1. Population demographics of Germany (G), France (F), United Kingdom (U), Spain (S) and Italy (I)

Age group	Professional activity (%)																														
	Students					Searching job					Employed					Marital status (%)															
	G	F	U	S	I	G	F ^a	U	S	I	G	F ^a	U	S	I	G	F	U	S	I	G	F	U ^b	S	I						
15-19 ^c	5.0	6.0	6.3	4.7	4.8	58.3	78.0	48.4	46.7	79.7	3.6	6.4	13.8	15.2	7.0	32.3	12.4	34.2	4.2	5.4	99.8	99.7	99.4	94.1	99.1	0.2	0.2	0.2	0.5	2.0	0.9
20-24	6.0	6.0	8.4	5.4	5.1	18.4	26.1	17.3	25.4	32.6	5.4	17.7	10.8	34.1	16.2	68.8	50.6	63.2	30.0	37.6	95.2	95.1	95.0	97.6	93.6	4.6	4.6	4.8	5.3	6.2	6.2
25-29	6.1	6.1	8.4	6.6	5.5	5.9	2.9	4.0	4.2	9.3	5.1	15.2	6.9	30.4	13.6	77.0	74.8	77.8	57.7	62.2	75.2	78.1	75.2	94.1	76.6	22.9	20.7	23.6	19.9	22.1	22.1
30-34	5.9	6.2	8.0	8.3	6.4	1.4	0.6	1.9	1.0	1.8	4.6	11.7	5.4	25.2	10.1	81.3	79.8	79.9	67.5	72.3	50.2	56.5	49.2	78.5	51.9	45.4	40.1	47.1	44.7	44.9	44.9
35-39	5.9	6.6	8.2	8.7	7.7	0.4	0.1	1.2	0.5	0.6	4.1	9.9	5.0	23.9	8.1	83.5	82.1	80.0	68.8	74.7	32.3	42.6	32.8	51.6	35.5	59.7	50.9	59.5	60.4	58.4	58.4
40-44	7.9	6.9	9.0	8.2	8.1	0.1	0.0	0.8	0.4	0.3	3.6	8.9	4.7	23.4	7.2	85.8	83.1	81.2	67.4	74.7	23.8	33.4	23.5	33.0	24.6	64.2	56.1	64.3	67.4	66.1	66.1
45-49	8.7	6.9	9.1	7.7	8.0	0.1	0.0	0.5	0.4	0.3	3.7	7.9	4.2	21.9	6.3	85.3	82.9	81.9	66.4	73.8	17.3	26.4	16.8	22.7	17.8	67.0	59.6	67.3	71.0	70.5	70.5
50-54	7.7	6.7	8.0	6.9	7.0	0.0	0.0	0.0	0.3	0.2	3.8	7.1	4.0	20.7	5.3	81.5	80.0	79.2	62.7	70.6	12.3	19.4	11.6	17.1	13.5	70.4	63.3	70.6	74.1	73.6	73.6
55-59	6.8	6.4	7.0	5.8	6.3	0.0	0.0	0.2	0.4	0.2	4.0	7.0	3.7	19.1	3.8	74.2	64.2	71.1	53.5	56.5	8.7	13.9	8.1	12.5	10.8	73.4	66.7	72.9	75.3	75.9	75.9
60-64	5.9	6.3	7.4	5.3	6.3	0.0	0.0	0.2	0.4	0.1	2.7	1.9	1.8	12.6	1.7	46.0	19.4	46.9	30.6	23.2	5.9	9.9	5.8	10.0	9.1	75.0	69.1	74.8	76.2	76.4	76.4
65-69	5.2	4.6	5.9	4.7	5.3	0.0	.0	0.2	0.4	0.1	0.4	0.5	0.4	1.6	0.3	14.7	9.1	20.7	6.1	9.3	4.9	8.0	4.7	8.5	8.4	73.0	69.0	73.5	73.9	73.3	73.3
70-74	6.1	3.7	4.8	3.7	5.2	0.0	0.2	0.3	0.3	0.1	0.1	0.2	1.2	0.1	8.6	9.1	9.6	2.8	4.7	4.6	7.1	5.1	7.9	8.1	68.5	65.7	70.2	69.1	67.3	67.3	
75 ^c	9.3	9.2	9.5	9.0	10.4	0.0	0.2	0.2	0.1	0.0	0.1	0.1	0.1	1.3	0.0	4.0	3.5	1.7	1.5	5.1	7.4	8.4	8.4	7.4	8.7	45.2	45.3	54.3	47.9	43.8	43.8

^aProfessional activity data for France was available as an aggregate for the age group 65-69, 70-74 and 75+.

^bFor UK, employed population marital status data is used. We assume figures provide reasonable estimates for the aggregate level as well.

^cFor UK, the data was available for 16 to 19-year age group.

- together to form self-identity. Students tend to have weak self-definitions because of their limited life experience. These characteristics continue to develop as one matures.
2. *Un-crystallized attitudes*: As students are undergoing the process of developing their “self”, their attitudes towards many issues tend to be unstable. Sears (1986) identifies four different research streams (i.e. panel analysis, changes to political events, cohort analysis on political party identification and cohort analysis on social welfare issues) which substantiate claims of unstable attitude in younger participants. These unstable attitudes lead to:
 - a. *Proneness to external influence*: Entering a university requires years of conformation to academic norms and requirements of schools and colleges. This makes students somewhat conformist in their attitudes. They are easily influenced because they have been conforming to the authority of dozens of instructors and administrators during their academic pursuit. Therefore, when students are exposed to stimuli in an academic setting, they may try to comply with the instructions instead of exposing their “real” behavior.
 - b. *Weak self-perception*: People with poorly defined attitudes tend to indulge in creating self-perceptions based on external cues rather than on introspections. Therefore, students are generally assumed to be more receptive to external cues and prone to their influences. Such consideration becomes important in studies where consumption objects (brands, products, celebrities etc.) are studied as a means to influence self-image.
 - c. *Attitude-behavior inconsistency*: There is substantial evidence that attitude-behavior consistency increases when attitudinal preferences are stable and established (Kelley and Mirer, 1974; Norman, 1975). Because students have unstable attitudes, their attitude-behavior consistency should be low when compared to non-students.
 3. *Group norms and social support*: As people grow older, they tend to form attitudinally supportive groups. On the contrary, the late adolescence and early adulthood stages are marred by abrupt disruptions of social groups because of many factors including geographical mobility, changing environment, work demands, increasing responsibilities, etc. For these reasons, students are continuously redefining their group surroundings and adapting to the norms. Coupled with their tendency to be easily influenced, changes in group structure also contribute to un-crystallized attitudes.
 4. *Improved cognitive skills*: Students undergo a variety of academic evaluations in their academic career. Therefore when they are presented with a research study in a typical classroom setting as part of a course, they may apply the same academic principles of finding the “right answer”, critical thinking and close attention to the studies. Such a situational disposition of cognitive resources is less representative of a consumer decision making situation in the marketplace.
 5. *Egocentric bias*: Students are in a life stage characterized by an overwhelming focus on personal needs, desires and

emotions. This supplements the notion that students exhibit less responsibility towards others surrounding them.

All these student-specific characteristics may limit the possibility to generalize results obtained on student samples.

ISSUES OF GENERALIZATION

The debate on whether the difference between student and non-student sample actually biases study results centers around issues of generalizability, commonly referred to as “external validity”. The objective of external validity is to produce knowledge which is applicable to a range of situations irrespective of time and place (Lucas, 2003). However, researchers differ on how studies establish external validity. Some relate it to methodological considerations, i.e. similarity between study settings¹ and real marketplace (Ferber, 1977; Calder *et al.*, 1981; Winer, 1999), while others relate it to theoretical wholesomeness, i.e. incorporation of contextual factors that can act as moderators (Petty and Cacioppo, 1996; Lynch, 1999). The former belong to the methodological school and the latter to the theoretical school of external validity. The former set of researchers argue that the more a study’s design resembles real life marketplace, the more logical it is to extrapolate study findings to the marketplace. The latter set argues that external validity is about theoretical extrapolation, and thus robust and wholesome theories (and not methodological similarities in study design) are the way forward. A common way (e.g. Lynch, 1999; Lucas, 2003) to understand these arguments is by framing them using the generalization classification of Cook and Campbell (1979): (i) ‘across’ persons, settings and times, and (ii) ‘to’ particular target persons, settings and time. Each of these types engenders specific problems that need to be elaborated in a marketing context.

Generalization “across”: issues of representativeness and boundary conditions

Generalization “across” refers to the notion that findings are generalized to a population which is different from the sample population (see Figure 1). Most often, the focus of marketing research is to study issues which explain consumption practices and buying behavior of consumers in general or of targeted groups of adult consumers (e.g. brand buyers, loyal consumers, seniors, etc.). Generalization “across” questions the fundamental assumption that students may represent all other consumers. As discussed previously, students differ from other consumers in many ways ranging from basic demographics to complex psychological development. Students tend to be relatively less experienced consumers (Kropp *et al.*, 2005) and one cannot assume their shopping behavior to be representative of other consumer segments (Yoon, 2013). In addition most of the empirical evidence also suggest that findings on student samples are not similar to the ones

obtained with non-student samples (cf: Peterson, 2001; Schepers and Wetzels, 2007; Wang and Yang, 2008; James and Sonner, 2001).

The methodological school argues that studies may ignore the fact that the findings are established on a homogeneous student sample which is characteristically different from the wider heterogeneous consumer population. In fact, in our empirical study described below, we find that nearly 95 per cent of studies using student samples and published in European marketing journals during the period 2005–2014 make generalization claims with an underlying assumption that similar findings would be observed on other consumers. This methodological school posits that if results are to be considered generalizable, it remains the responsibility of the researcher to establish sample representativeness and explain why the findings of studies with student samples should be considered applicable to other consumers. Sample representativeness can be established by comparing the distributions of the sample and of the population on characteristics which are relevant to the study (Ferber, 1977). For example in studies about hedonism, researchers can present relevant statistics, such as the number of hedonic products bought per month or the percentage of monthly budget allocated to hedonic consumption to demonstrate that students are similar to the more heterogeneous non-student population in terms of hedonic product consumption. If such supporting evidence is not presented, findings should be considered relevant to a subset of the consumer population (i.e. students). Thus, the methodological school puts the onus on the researchers to either demonstrate that the phenomenon studied functions similarly for students and other consumers or concede that the demographic and psychological differences impose boundary conditions on the findings.

On the other hand, the theoretical school argues that representativeness is difficult to establish and, although desirable, it is not a necessary condition for theory testing (Basil, 1996). A single study even with real world settings would lack generalizability as ‘external validity’ cannot be increased with higher levels of similarity between study settings and marketplace, but only by developing more wholesome theories which incorporate relevant boundary conditions (Petty and Cacioppo, 1996; Lynch, 1999). If study findings differ between a student sample and a non-student sample, it indicates boundary condition that must explicitly be treated as moderators for theory building (Petty and Cacioppo, 1996; Lynch, 1999). Once theories have adequately accounted for such moderators representing differences in study settings, it is only then that valid generalizations can be made from it. Researchers from this school of thought critic the banal reference to student sample as a limitation to study. Instead, they recommend researchers to actively treat sample specific moderators in their conceptualization and design stages or, at the minimum, explicitly discuss such moderators and their potential impact on study findings in the limitations section.

Generalization “to”: issues of relevance

Generalization “to” deals with the idea that findings of a study can be generalized to the population from which a sample is selected. When students are selected as study

¹Study settings involve considerations such as stimuli, study design, sample choice and measurements. For the scope of this article, we limit discussion to the choice of study sample only, i.e. student vs. non-student samples.

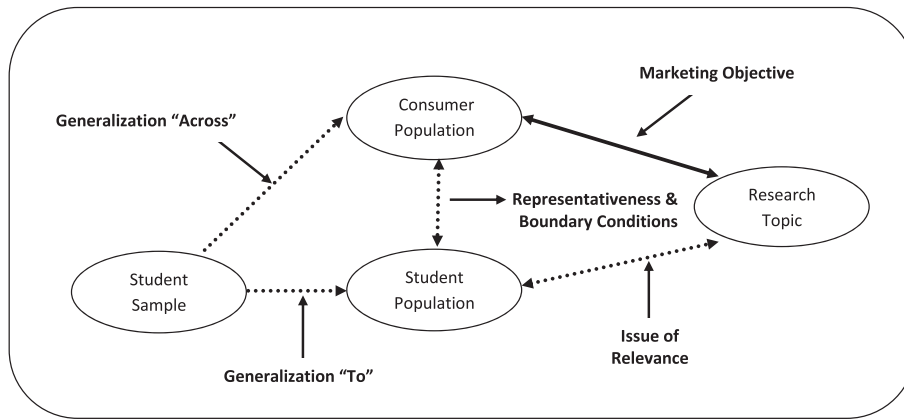


Figure 1. Issues of generalization in marketing research.

participants, then generalizations of results are made to the student population (see Figure 1). An established characteristic of student populations is their homogenous profile (demographics and life experiences) and inter-sample homogeneity is often used as an argument in studies where inter-group comparisons are to be made (e.g. Aaker and Sengupta, 2000; Martinez *et al.*, 2008). Because student samples have homogenous profiles, inter-group differences in results cannot be attributed to sample characteristics, but rather to the manipulation itself. Therefore, generalization “to” per se appears to be an argument supporting the use of student samples in marketing research. However, this raises a serious concern about sample/population relevance to marketing research. Does the research focus on an issue which is relevant to a student population? In order to have meaningful results, the fit of the issue under study with student participants should be analyzed. Emphasizing the importance of relevance in consumer studies, Ferber (1977, p.58), in his editorial, articulates that if researchers cannot provide rationale on the relevance of student samples to the study, “there is no reason why the GIGO principle should not be applied immediately, namely, garbage in, garbage out”.

The notion that students are consumers and that any consumption theory can be studied with them underestimates the importance of sample-issue fit. In some consumption situations students may lack relevant experience and may not be able to provide valid responses. For example, Liu *et al.* (2012) study consumption phenomenon (i.e. self-congruity, brand attitude and brand loyalty) for luxury brands using students among whom around 60 per cent spend less than AU \$200 dollars per month on non-necessities. Similarly, Yim *et al.* (2014) examine drivers of luxury brand attitudes using student samples. However, students as consumers have limited resources and financial independence and probably lack consumption experience for such high end products. In another set of studies, Laufer *et al.* (2010) and Skarmeas and Shabbir (2011) study charity giving and fund raising using student, but it is unclear how relevant this phenomenon is for students who usually are more concerned about how to make ends meet. With the use of students, Rampl and Kenning (2014) study issues related to employer brand trust in spite of the fact that students have very limited experience, if at all, as employees. Of course researchers rightly indicate

that students are ‘potential future’ consumers of luxury product, strong donors and employees. However, implicit in this acknowledgement is the fact that study participants lack experience as ‘actual’ consumers, donors and employees. Therefore, their responses could potentially differ from the population actually relevant to these phenomena.

Another type of relevance deals with sample-stimuli fit. Researchers most often choose stimuli which are relevant to students (e.g. while studying hedonism, choices could be mobile phones or watches instead of high end luxury brands). For example, Michaelis *et al.* (2008) in a study on consumer trust use mobile phones as a stimuli and indicate that nearly all the respondents had signed a mobile phone contract, and in a study on consumer reactions to food contamination, Carvalho *et al.* (2008) use food contamination in student cafeteria as a stimuli. Making stimuli relevant to the sample increases the validity of the responses collected from the subjects. However, in some cases, this fit between the sample and the stimuli remains questionable. In a study on new product evaluations using student samples, Odou (2005) uses automobiles as the product category and the Mercedes and Opel brands as stimuli. Such a methodological choice presents sample-stimuli fit issues (the product category—new cars—and the brands—Mercedes). Lack of sample-stimuli fit raises validity issues and James (2006) for example, deletes child car seats and hydraulic excavators from a prospective list of product categories to be used in a study, because they lack relevance to student samples. A good research practice is to pre-check whether the issue studied and the stimuli used are both and jointly relevant to the sample.

Does it matter? Empirical evidence

Much of the debate on student sample usage and its implications on generalization is conceptual in nature. Sears (1986) acknowledges that his student sample profiling is inferential in nature and that its actual impact on empirical relationships needs to be verified. He infers that using student samples may result in the strength of relationships being incorrectly reported with the sign and shape of the relationship being misrepresented. In line with his assertions, we review three meta-analysis that explicitly consider respondent type (students vs. non-student) differences in study results.

Peterson (2001) sensing the importance of the issue and realizing the need to support theoretical propositions with empirical data, conducted a second order meta-analysis to test differences in effect size between student samples and other samples ($N > 350,000$). He demonstrates that studies using student samples generally report higher effect sizes. Out of the total 64 relationships analyzed, students have a larger effect size in 35 instances (55 per cent) and a smaller effect size in 27 instances (42 per cent) whereas in two instances the effect sizes are the same. Especially intriguing is the finding that in 19 per cent of cases the directionality of the effect size is different, representing a proportion of 1 in 5 relationships being misrepresented. Moreover, in 29 per cent of relationships the size of the larger effect differs by a factor of more than two from the smaller one. When combined together, it appears that in nearly half of the cases (49 per cent), relationships differ either by size or directionality. Similar results are reported by Wang and Yang (2008) in their meta-analysis in international marketing. They analyze 196 effects in international business marketing and 878 effects in international consumer marketing. Studies with students in international business had significantly larger effect sizes than studies using managers. In international consumer marketing, studies with student samples have significantly lower effect sizes than those using non-student consumers. Schepers and Wetzels (2007) report similar results in their meta-analysis of the Technology Acceptance Model (TAM) based on 51 articles containing 61 studies. They find clear patterns of significant larger effect sizes for studies using students ($n = 5224$) than non-students ($n = 8896$) for both social and technology related variables. These meta-analyses involving thousands of study participants and hundreds of effects sizes provide strong empirical support to the notion that using student samples in research may bias results.

Apart from these meta-analysis, some individual studies examining differences in study results between student and non-student samples report mixed results. For example, James and Sonner (2001) confirm differences for advertising effectiveness on measurements such as liking and purchase intent; Völckner and Sattler (2007) report no differences in studies on brand extensions; and Maguire *et al.* (2003) demonstrate no differences in willingness to pay to a charity both in terms of likelihood and dollar value. Interestingly, single study citation remains the most commonly used argument by researchers using student sample to justify their choice of the sample. Such a practice necessitates caution as single study findings imply specific contextual factors (e.g. sample characteristics, stimuli choices, measurement errors etc.). Assertions based on a single study remain weak and should be substantiated either through multiple studies or through a meta-analysis supporting the claim.

Another important issue identified by Sears (1986) is the limited range of values on dimensions common to students (e.g. limited age bracket, similar education levels and environmental characteristics, etc.). The limited range of values makes student samples relatively homogenous and less representative of the heterogeneity that is present in a larger population. Scientifically, these value ranges have important

implications for theoretical relationships. In a study exploring the link between risk aversion and demographic characteristics, Andersen *et al.* (2010) find that with laboratory based student dominated samples, age and education do not impact risk aversion tendencies. On the contrary, the impact of age and education on risk aversion becomes significant when the relationship is tested in the field using a population representative sample. This occurs with students because the measured values have limited variance. It therefore becomes difficult to find support to the relationships (in the field experiments the respondent's age ranged from 19 to 75 years, while in the laboratory study, 98 per cent of participants were less than 30 years old). In another set of studies, Green *et al.* (1996) demonstrate that younger adults are more impulsive and make decisions that favor immediate rewards instead of delayed and bigger rewards. They posit that between 20 and 30 years, adults are going through transition and that their impulsivity, risk taking behavior and personality traits tend to become stable only after that. Therefore, the range of values normally observed in student samples may also influence the significance of explored relationships.

STUDENT SAMPLES IN EUROPEAN MARKETING RESEARCH

Referring to the unabated use of student samples, Ferber (1977) in his JCR editorial cautioned "In economics the phenomenon of a poor currency driving out a good currency is widely known and goes under the name of Gresham's law ... In consumer research ... this phenomenon is the tendency for researchers to turn increasingly to convenience samples and forego probability samples" (p.57). Ferber's concern regarding Gresham's Law remains relevant to marketing research and the use of student samples is still prevalent with no signs of decline.

In spite of the historic nature of the debate and student sample prevalence in marketing research, no systematic study has been done to identify patterns of student sample usage. It remains unclear what kind of studies use student samples, how they report their findings and in which areas of the marketing discipline the issue is more prevalent. A systematic review contributes to the debate in two ways. First, for proponents of student samples it helps assess the practice of student sample usage and identify directions for better reporting. Second, it enhances the relevance of the debate by framing the issue from the broader marketing discipline to the more specific domain or the relationships under study. Certain sub-domains may be more susceptible to student sample usage than others. Analyzing domain specific relevance of student sample will bring the debate closer to the interests of researchers. For example, a researcher studying brand extensions will be less convinced with the empirical evidence demonstrating student sample differences in advertising research, but, will be more concerned if the results are in the area of branding and brand extension. Thus, identification of sub-areas using student sample provides direction to researchers to assess the severity of the practice in their areas of interest and work towards findings a consensus among peers.

To explore this, we conducted a systematic review of papers using student samples to assess the magnitude of the issue, the way in which such a practice is discussed and reported, and to identify sub-areas in which the practice is more prevalent than others. We analyzed three leading marketing journals from Europe as ranked by ABS (2015), i.e. the European Journal of Marketing (EJM), the International Journal of Marketing Research (IJRM) and the International Marketing Review (IMR), for publications during the last decade (2005–2014, inclusive). In total, 817 papers published in EJM, 344 in IJRM, and 306 in IMR were analyzed. We selected empirical theoretical papers because the student sample usage debate remains inconclusive regarding the merits of student samples in theory application studies (Peterson, 2001), and filtered these based on the criteria that hypotheses are outlined and then tested using statistical tools. The filtering task was simultaneously conducted by two different judges who independently filtered out the papers in these outlets. Inconsistencies were reviewed by both judges and reconciled.

Results indicate that 768 theory testing empirical papers (EJM=388; IJRM=220; IMR=160) were published in these journals during the 2005–2014 period. Papers using panel data (e.g. Chen and Huang, 2013; Nicolau, 2013), and the ones defining either the target population (e.g. immigrants: Segev *et al.*, 2014; managers: Al-Khatib *et al.*, 2011) or the domain of application (retailers: Mohr and Batsakis, 2014; non-profit organizations: Lefroy and Tsarenko, 2013) in their title were ignored. Out of the remaining 496 papers (EJM=301; IJRM=93; IMR=102) with general titles alluding to market wide generalization, 99 used student samples exclusively (EJM=47; IJRM=27; IMR=25) to test their theoretical models i.e. 19.96 per cent of total theoretical papers (EJM=15.61 per cent; IJRM=29.03 per cent; IMR=24.51 per cent). Another 45 papers (EJM=28; IJRM=8; IMR=10) use student samples for some part of their study (e.g. pre-tests, initial theory testing) but then supplement tests of their theoretical models with heterogeneous non-student samples as well. Figure 2 provides yearly evolution of the percentage of theoretical papers using student samples (i) in total; (ii) exclusively; and (iii) in combination with other heterogeneous samples for the three European journals.

Reporting of results

The 99 papers using student samples are further analyzed to detect patterns regarding student sample characteristics, methodologies used, nature of conclusions drawn and limitations of external validity mentioned in the paper. Only four papers provide sufficient details (age range, average age and education level) about the student samples enabling readers to make detailed judgments about participants’ profile. Education level (34 papers) remains the most common reported characteristic and that 17 papers do not provide any details about student sample characteristics in their study. The absence of sufficient details from a majority of the papers (95 out of 99) alludes to the fact that researchers do not deliberate on the importance of the demographic or psychological characteristics of their samples. A detailed

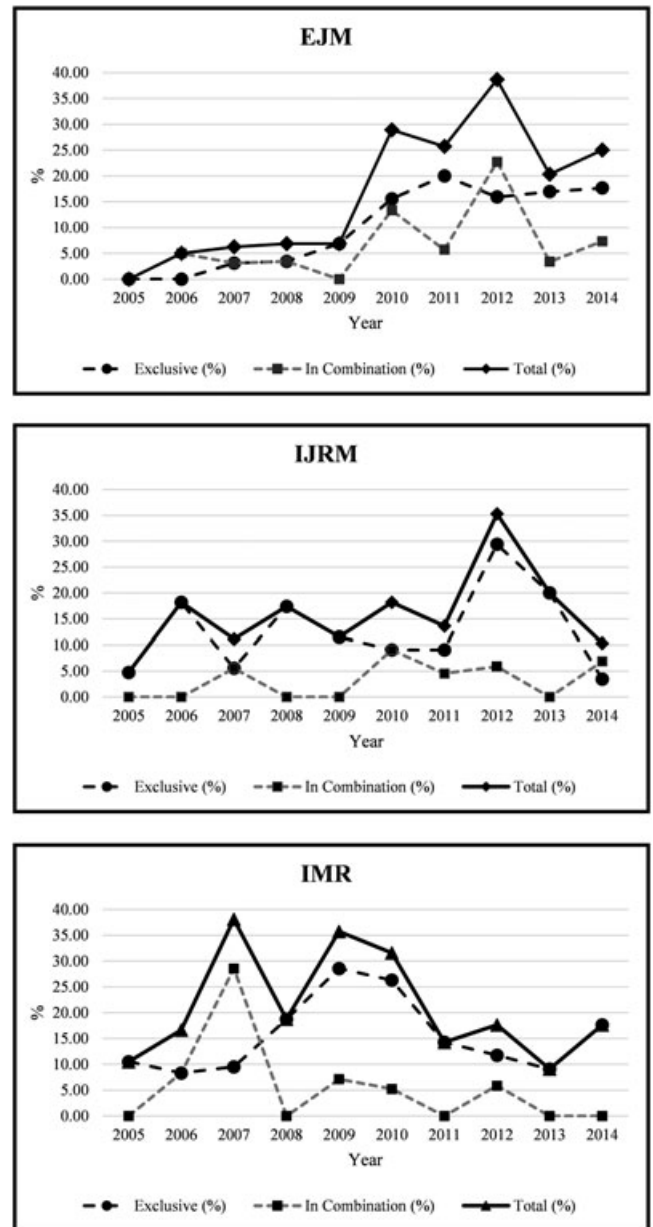


Figure 2. Studies using student samples as a percentage of total theoretical studies published during the year.

reporting would help assess the influence of sample characteristics (low mean age and limited life experience, restricted variance) on findings. Concerning methodologies used, 78 papers (78.8 per cent) used experimental designs to test hypotheses, while 20 used surveys and 1 used model estimation. Such a finding is not surprising considering the cost and effort required to conduct experiments with adults in a controlled laboratory setting.

Out of the 99 papers, only six papers focus on the theoretical dimensions of their findings and show a deliberate concern for not generalizing their findings. The remaining 93 papers (94.9 per cent) draw general conclusions from their findings by using words such as “consumers” (e.g. Fetscherin and Toncar, 2010; Buil *et al.*, 2013; Yan and Duclos, 2013), “public” (e.g. Laufer *et al.*, 2010), “people” (e.g. Carvalho *et al.*, 2008), “professionals” (Rampl and Kenning, 2014) etc., in their conclusions and managerial implication sections.

Student samples seem to be de facto considered as representative of consumers in general with an assumption that results obtained with students would be similar to other samples of consumers. However, in spite of these general conclusions, 57 out of these 93 studies (61 per cent) acknowledge the use of student samples as a potential threat to external validity in the limits section of the paper. This can succinctly be described in the words of Wells (1993, p. 493) "It is surprising to outsiders that the authors then go on to state broad, general conclusions as though the limitations had vanished on being confessed." The remaining 36 papers (39 per cent) making generalized claims about their findings did not refer to student sample as a potential limitation to study results.

Commonly occurring areas of research

In order to better understand which areas in marketing research tend to use student samples more frequently, we clustered the 99 articles² into distinct groups. This identification is aimed at bringing more focus to the debate on student sample usage, i.e. from the broader discipline of marketing to the more specific domain of study and, ideally, to the relationship(s) under study. We contend that the merits of using student samples or not will be better served if researchers in these domain collectively work to accumulate evidence to understand the implications of student sample usage in their domains.

After reviewing the article titles, abstracts and key words, we identified 11 clusters (see Table 2) that represent domains in which student sample usage is frequent. As these domains define broad categories, we also report frequently occurring themes within each cluster to account for intra-cluster variations. This exercise enabled identify two important issues. First, researchers in a domain are generally not well aware of the implications of student sample usage in their specific domain and tend to frame arguments at the broader marketing discipline level. For example, out of the 8 studies on 'Brand Extensions and Fit' (Cluster: Branding), only four acknowledge external validity limits to their study findings and that also because of the homogenous and non-representative profiles of study participants. Only one out of these four reports a relevant study in the area of brand extensions as an argument to support their use of student samples. This demonstrates that most researchers treat 'student' sample usage as a marketing discipline problem and that they do not perceive the possibility of it having less/more impact for their area of concern. We contend, that this presents opportunities for researchers to test for these differences more explicitly in their areas of interest to develop evidence whether study findings differ across student vs. non-student samples or not. For example, Völckner and Sattler (2007) in a study on empirical generalizability of brand extension research demonstrate that study findings are generally similar when obtained with student and non-student samples. As single study results cannot provide conclusive evidence, the efforts by Völckner and Sattler (2007) are a first step to settle the debate for brand extension research. Subsequent studies testing for such

differences will provide more clarity to researchers in the domain and may serve as an input to a meta-analysis.

The clustering also helps researchers to discern upon the implications for their study when strong evidence exists in their domains. For example, researcher working in the area of product/technology adoption will find useful the results of the meta-analysis conducted by Schepers and Wetzels (2007). These authors report significant differences between the results obtained using students vs. non-student samples for both social and technological variables related to adoption. With such differences, researchers should be prudent when making generalizing claims for their study (e.g. in the title, abstract, and conclusion sections) and aim to incorporate relevant boundary conditions for more robust testing of their theoretical frameworks.

BEST PRACTICES IN SAMPLE SELECTION

This section provides an overview of best practices identified in the literature that may help improve the heterogeneity of samples used in marketing research. Heterogeneous adult samples remain the most desirable choice for marketing researchers and theories tested on such samples provide higher validity of reported results. Therefore, we first discuss ways that may help researchers collect data more easily from a heterogeneous non-student sample (Figure 3). It is followed by an argument that when heterogeneous non-student samples are not feasible, researchers should prefer post-graduates (MBA/EMBA) rather than undergraduates, as they possess characteristics which make them more similar to heterogeneous non-student samples. Finally, we highlight circumstances in which student samples are suitable and situations where estimating student sample biases increases confidence in reported results.

Heterogeneous adult samples

Social psychologists often consider student samples not to be a threat as external validity is not a serious concern in their discipline. They classify political science and sociology as disciplines more concerned with the relevance to the real world (Fiske, 2008; Henry, 2008). Similarly, Wells (1993) and Winer (1999) contend that marketing is naturally different from psychology and social psychology as it intrinsically has an applied focus. Marketing research should therefore be relevant to the consumption marketplace which is better represented by "non-students" than by "students". Winer (1999) supports this case by highlighting that business schools are increasingly disseminating their research in the form of simple to understand newsletters to the business world and that scientific marketing journals (e.g. *Journal of Consumer Research*) are disseminating their published work as press notes. Such a practice is directed at establishing relevance of marketing research to practitioners.

Collecting data from heterogeneous adult samples is time consuming and costly. However, in a field experiment it is often possible to conveniently and cheaply access an adult population the members of which are waiting at different locations such as train stations, airports, bus stations, parks and

²Some articles could potentially appear in two clusters e.g. Wang and Sun's (2010) work on Advertising in Online environment. For such articles, we classified them in the most relevant cluster.

Table 2. Clusters of studies using student samples

Cluster	% of total (n = 99)	Major themes
Branding	21	Brand Extensions, Brand Alliances, Luxury Brands, Global Brands, Brand Cue Processing, Measurement issues in Branding
Consumer Behavior	20	Interpersonal Influence, Consumer Emotions, Consumer Risk Perception, Consumer Trust, Consumer Choice
Advertising	19	Ad Appeals (Sexual, Violence, Emotional, Comparative), Message Framing, Event Sponsorships
Country of Origin	9	Consumer Animosity, COO impact on Brand/Product evaluations, COO and Trust
Online Marketing	7	eWOM, Online Purchasing
Service Marketing	6	Service Recovery
Social Responsibility	5	Consumer Social Responsibility (donations, charity), Corporate Social Responsibility
Relational Marketing	3	Loyalty Programs
Consumer Co-creation	3	
Consumer Promotions	3	
Product Technology Adoption	3	

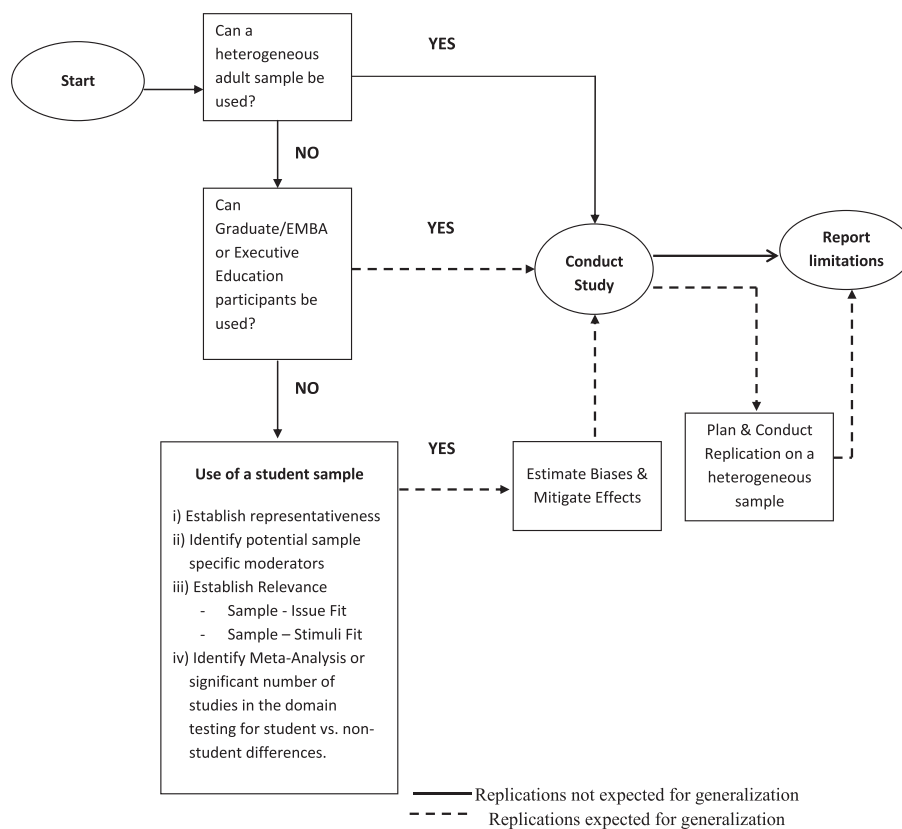


Figure 3. Sample selection process.

waiting rooms. For example, Dunn and Schweitzer (2005) in a study on trust and emotions conduct three different field experiments using adults waiting at a train station. Similarly, Mazodier *et al.* (2012) use trained assistants to conduct intercept interviews at public places such as parks and market-places. As Henry (2008) points out, energetic research assistants with clearly visible university affiliation are good candidates who can help to easily approach and collect data from adults at these waiting spots. These samples remain convenience samples but offer far more variability than convenience student samples.

Advances in the internet technology also open up more efficient avenues of accessing heterogeneous adult samples.

The traditional way is to use on-line panels of consumers. Although efficient, this remains expensive. In recent years, online survey software companies developed their own panels which are accessible at a relatively low price (e.g. SurveyMonkey). New systems are also developing such as the Mechanical Turk (MTurk) service. It offers a scalable workforce which people can hire to perform their jobs for compensation, including but not limited to surveys. Paolacci *et al.* (2010) and Buhrmester *et al.* (2011) demonstrate that MTurk population is more representative of the general non-student US population than other traditional sources of convenience sampling and that data quality meets or exceeds psychometric properties expected in academic research.

Studies using MTurk appear in marketing journals such as *Journal of Marketing Research* (Amar *et al.*, 2011), *Journal of Consumer Research* (Smith *et al.*, 2013) and *Journal of Consumer Psychology* (Paolacci *et al.*, 2011). Advances in technology provide researchers more incentives to update their modus operandi for collecting data on heterogeneous adult samples. However, collecting data on the internet raises other issues such as attention, motivation and absence of identification. Therefore, internet based data collection should ideally be supplemented with other data collection methodologies.

At times researchers find it difficult to collect adequate data from non-student sources and tend to merge the student sample with the non-student one. In situations where two sample groups are pooled together (e.g. student + non-students), it is advisable that the responses be tested for similarities before pooling the datasets. For example, Malär *et al.* (2011) in a study on emotional brand attachment and brand personality pooled students (69 per cent of all respondents in study 1 and 60 per cent of all respondents in study 2) with non-students. However, authors tested the extent to which students' answers were comparable to other respondents and conducted a mean difference tests for all focal constructs. If the sample is sufficiently large, conducting formal tests between students and non-students sub-samples to establish measurement invariance between the two groups before the data is merged together is warranted (Steenkamp and Baumgartner, 1998).

Graduate/MBA/EMBA students

The age cohort of non-traditional student samples (i.e. MBAs, Executive MBAs, Evening Programs, Executive education) demonstrates higher levels of heterogeneity than traditional student samples (e.g. undergraduates). In a relevant study, James and Sonner (2001) compare a sample of non-traditional students who are older adults pursuing education during evening classes alongside their work and family responsibilities with the traditional undergraduate and heterogeneous adult samples. Across different constructs such as emotions elicited by advertisements, ad liking and purchase intentions, undergraduate students' exhibit marked differences in preferences with adult samples. By contrast, non-traditional students show similar preferences to other heterogeneous adult samples. Marketing researchers are therefore encouraged to look in this direction as a viable middle ground between non-student samples and "controversial" student samples. For example, Kabadayi and Lerman (2011) used graduate students for their study (average age of 35 years, 75 per cent with full time work positions, 63 per cent being married and 59 per cent being parents). If researchers cannot collect enough respondents from the graduate student pool then an effective way is to combine both the undergraduate and graduate students. For example, in a study on celebrity endorsements for products, Liu and Brock (2011) collected data from a sample of undergraduate and graduate students. The sample comprised 32 per cent respondents in the 18 to 23-year age bracket, 28 per cent were 24–29, 17 per cent were 30–35 and 23 per cent were above 35 years of age. Thus, by using non-traditional graduate students, researchers may test theoretical models on a more

heterogeneous samples with little effort. However, before pooling datasets, researchers should conduct formal tests of invariance between the datasets.

Undergraduate students

In spite of the controversies surrounding the use of student samples in scientific research, undergraduate students are valuable resources and their discretionary use helps the progress of knowledge. They are easily accessible and their use is cost effective. Researchers such as Sears (1986), Henry (2008) and Peterson (2001) who warn the scientific community against the overwhelming use of students sample also support the notion that careful usage of student samples can be useful. They concur to the notion that although the use of student sample alone may limit the understanding of consumer behavior, a strategy using both students and non-students is a productive option.

Theoretical studies and bias estimation

Calder *et al.* (1981), from the theoretical school, were the first to advocate the use of student sample in marketing studies. They argue that the use of student sample for theory testing is justified because it is the theory which is supposed to be generalized and not the results per se. However, these advocates of student sample utilization do not abscond researchers from added responsibility. As, Lynch (1982, p.234) notes, "If findings supporting one's theory lack external validity, the theory lacks construct validity. The theory is at a minimum incomplete, and it is quite possibly just plain wrong."

Lynch (1982, 1999) and Petty and Cacioppo (1996) debated that theories tested on students may have to be modified when tested with adults or more heterogeneous samples because of interaction effects with background factors. This indicates that although theoretical phenomena may be studied with students, the peculiarities of their life experiences may bias findings. For example, sex appeal advertising (Liu *et al.*, 2009) and nudity (Dianoux and Linhart, 2010) are issues relevant to younger consumers. However, the use of student samples may bias results as younger participants tend to have a more tolerant view of the phenomenon than older ones for provocative appeals. Similarly, a study on authority based advertising (Jung *et al.*, 2009) can be studied with student samples but results may be biased as students are more exposed to authoritative environments (university and dependent family status) and have yet to earn their independence (e.g. financial). Therefore, studies using student samples should estimate whether the sample introduces biases in study findings. When such an exercise is conducted at a pre-study phase, it helps in two ways. First, it encourages researchers to deliberate about the established differences between students and heterogeneous adults, bringing the issue to a conscious and pro-active realm. Second, it lends credence to the findings as interaction effects and boundary conditions are actively treated and addressed.

Winer (1999) and Wells (2001) share the same argument and propose that theoretical application studies should not be given a free pass for the use of student samples and that it is the responsibility of the researcher to establish external validity. Winer (1999) recommends that such studies must

either extend their findings by using alternative designs (such as panel data or other samples) in the same research or at the very least propose direction for studies that should augment the generalizability of findings. While saying so, Winer (1999) is not referring to the current practice of customarily stating that the studies are to be replicated with a heterogeneous adult population but to the explicit identification and discussion of the biases and boundary conditions that may limit generalizability of current findings.

Exploratory/preliminary studies

Ferber (1977) advocates that student samples may be helpful in exploratory studies where the objective is to understand multiple explanations of a phenomenon. Students who form convenient and cost-effective samples may be used for preliminary diagnosis of a theory (pre-tests, initial experiments) without committing extensive resources, energy and time. Such a practice offers two advantages. First, as student samples constitute a homogenous group, they provide high levels of internal validity, a viable first step for studies which subsequently demonstrate external validity of their theoretical framework on a more heterogeneous sample. This stepwise procedure helps the development of sound theoretical frameworks that satisfy both internal and external validity concerns. Second, it helps identify the presence or absence of boundary conditions in theoretical models once they are tested both on students and non-student samples. For example, if an effect works for students and not for the general population then it alludes to possible moderations of background variables that must be incorporated to make theories more reliable (Calder *et al.*, 1981; Petty and Cacioppo, 1996; Lynch, 1999).

Our review of articles published in European marketing journals identified 45 studies (EJM: 28; IJRM=7; IMR: 10; see Figure 1 for evolution), representing 9.1 per cent of the total theoretical empirical papers published in these outlets, where student samples were used discretionally for preliminary studies before re-validating the theoretical model using heterogeneous non-student samples. For example, Andrews and Kim (2007) in a study on multi-national brands use 210 university students in their pilot study before conducting the main study using 213 consumers with an age range of 19–70 years old (mean age=39.8 years). López and Sicilia (2013) study the effect of word of mouth on product adoption, and test their hypotheses with 217 undergraduates in Study 1, and then establish external validity of their findings using 170 respondents of an online panel. Similarly, Van Doorn and Verhoef (2011) in a study on willingness to pay for organic products, first test their conceptual model with a group of 172 students in a large Dutch University, and then re-examine the model with a large scale survey among 737 Dutch consumers responsible for their household food purchases. In all such cases, researchers gain insights using student dataset before they venture on a more burdensome data collection procedure involving heterogeneous adults.

It may appear demanding to think in terms of multiple studies in order to earn credibility. However, publications in top tier journals test their findings across multiple samples (e.g. the Journal of Consumer Research articles contain on average 3.5 samples in 2009). The paradigmatic shift from

a single study to multiple studies involving heterogeneous samples gives more credibility to findings and enhances confidence on the scientific knowledge produced. The benefit of using a combination of student and non-student samples across different studies is the balance between convenience and producing reliable research results.

CONCLUSION

In spite of the debate on the merits of student sample in academia for the past five decades, no substantial change in practice is observed. This review provides a systematic analysis of the studies using student samples to identify trends in usage practices and research areas with the aim to take the debate forward. It provides insight to researchers to more prudently use and report on student samples issues. Researchers should be conscious of the basic demographic and psychological profile of student populations and how it may impact study results. It proposes that the merits of student samples be closely weighted by looking into the areas of interest rather than the broader discipline of marketing or social sciences in general. In this pursuit, identification of areas more frequently using student samples provides researchers with opportunities to embrace the debate and explicitly test for differences in order to establish a pool of findings that can be relied upon more conclusively.

A gradual shift towards solving this issue also requires active participation of editors and reviewers. Wells (2001) and Henry (2008) suggest that editors and reviewers should assign additional weights to studies using non-student samples and that researchers be encouraged to reverse the ‘file drawer effect’ when inconsistent results are found with student vs. non-student samples. They also propose that a special section on replications be opened in which theoretical models extended from student samples to a wider population are more openly accepted as a scientific contribution. An important step forward in this direction is the ‘Replication corner’ introduced in the International Journal of Research in Marketing (IJRM) (Goldenberg and Muller, 2012), where the editors consciously intend to address the dearth of generalizable conclusions in marketing research. Evidently, such replications are designed as extensions related to boundary conditions of the existing theoretical models. The IJRM editors define it as conceptual replications with a pre-defined deviation. This is a rejoinder to the statement of Andrew Ehrenberg’s position on generalization: ‘‘A result can be regarded as routinely predictable when it has recurred consistently under a known range of different conditions. This depends on the previous analysis of many sets of data, drawn from different populations. There is no such basis of extensive experience when a prediction is derived from the analysis of only a single set of data. Yet that is what is mainly discussed in our statistical texts’’ (Ehrenberg and Bound, 1993, p. 207).

Although this study is restricted to three major European marketing research outlets, it is encouraging to observe that prevalence of student samples is not the dominating rule in European research (EJM=15.61 per cent; IJRM=29.03 per cent; IMR=24.51 per cent). Similar empirical investigations of international journals in marketing (Peterson, 2001) and

social psychology (Henry, 2008) find higher rates (around 80 per cent) of student sample usage in published research. There is no doubt that students represent a user-friendly convenience population and outright rejection of such an important research resource would be detrimental to the overall development of science. Student samples may be used in consumer research but not without caution. A viable way forward is to use student samples for initial theory testing and then follow on to validate study findings with a more heterogeneous adult population. Such a practice permits academic scholars to balance the convenience of using student samples with the responsibility of developing robust theories. Thus, as Wells (1993, p. 492) succinctly explains: "This is not to say that findings based on students are always wrong. It is only to say that findings based on students should be carefully reported by authors and interpreted by readers. Our findings would be substantially more credible if students were not so often the first and only choice."

ACKNOWLEDGEMENTS

The project was part of first author's doctoral thesis at Aix-Marseille Graduate School of Management, Aix-Marseille University, France and for which funding was provided by the La Poste Chair at Kedge Business School, France. The project was completed using LUMS Faculty Start-up Grant (OSP-323-02) # 460. We also thank Faheem Ahmed for data collection.

BIOGRAPHICAL NOTES

Rohail Ashraf is an Assistant Professor at the Suleman Dawood School of Business, Lahore University of Management Sciences (LUMS), Pakistan. He holds a PhD from Aix-Marseille Graduate School of Business (Aix-Marseille University), France. His research interest includes corporate brands, corporate advertising, consumer emotions, and consumer identities.

Dwight Merunka is Professor of Marketing at Aix-Marseille Graduate School of Management (Aix-Marseille University) and at KEDGE Business School in Marseille, France. He previously was at Paris-Dauphine University and at Toulouse University. He taught in various international universities such as the University of Texas at Austin, Renmin University of China, Shanghai Jiao Tong University, or the Universities of Dakar, Abidjan and Cotonou in Africa. He holds an MBA from Ecole des Hautes Etudes Commerciales (HEC Paris) and a PhD from Aix-Marseille University. His work is published in scientific journals such as International Journal of Research in Marketing, Journal of Business Ethics, Journal of the Academy of Marketing Science, Journal of Business Research, Journal of Forecasting, Journal of Personal Selling and Sales Management, or International Marketing Review. His research interests cover branding, brand management, consumer behavior, consumer well-being, and cross-cultural consumer research.

REFERENCES

- Aaker JL, Sengupta J. 2000. Additivity versus attenuation: the role of culture in the resolution of information incongruity. *Journal of Consumer Psychology* 9(2): 67–82. DOI:10.1207/S15327663JCP0902_2.
- Al-Khatib JA, Malshe A, Sailors JJ, Iii IC. 2011. The impact of deceitful tendencies, relativism and opportunism on negotiation tactics: a comparative study of US and Belgian managers. *European Journal of Marketing* 45(1/2): 133–152. DOI:10.1108/03090561111095621.
- Amar M, Ariely D, Ayal S, Cryder CE, Rick SI. 2011. Winning the battle but losing the war: the psychology of debt management. *Journal of Marketing Research* 48(SPL): S38–S50. DOI:10.1509/jmkr.48.SPL.S38.
- Andersen S, Harrison GW, Lau MI, Rutström EE. 2010. Preference heterogeneity in experiments: comparing the field and laboratory. *Journal of Economic Behavior and Organization* 70(2): 458–469. DOI:10.1016/j.jebo.2009.09.006.
- Andrews M, Kim D. 2007. Revitalizing suffering multinational brands: an empirical study. *International Marketing Review* 24(3): 350–372. DOI:10.1108/02651330710755339.
- Bagozzi RP. 1992. The self-regulation of attitudes, intentions and behavior. *Social Psychology Quarterly* 55(2): 178–204. DOI:10.2307/2786945.
- Basil MD. 1996. Standpoint: The use of student samples in communication research. *Journal of Broadcasting and Electronic Media* 40(3): 431–440. DOI:10.1080/08838159609364364.
- Bello D, Leung K, Radebaugh L, Tung RL, Van Witteloostuijn A. 2009. From the Editors: Student samples in international business research. *Journal of International Business Studies* 40(3): 361–364. DOI:10.1057/jibs.2008.101.
- Buhrmester M, Kwang T, Gosling SD. 2011. Amazon's mechanical turk: a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science* 6(1): 3–5. DOI:10.1177/1745691610393980.
- Buil I, De Chernatony L, Montaner T. 2013. Factors influencing consumer evaluations of gift promotions. *European Journal of Marketing* 47(3): 574–595. DOI:10.1108/03090561311297463.
- Calder BJ, Phillips LW, Tybout AM. 1981. Designing research for application. *Journal of Consumer Research* 8(2): 197–207.
- Carlson R. 1971. Where is the person in personality research? *Psychological Bulletin* 75(March): 203–219. DOI:10.1037/h0030469.
- Carvalho SW, Block LG, Sivaramakrishnan S, Manchanda RV, Mitakakis C. 2008. Risk perception and risk avoidance: the role of cultural identity and personal relevance. *International Journal of Research in Marketing* 25(4): 319–326. DOI:10.1016/j.ijresmar.2008.06.005.
- Chen HN, Huang CY. 2013. An investigation into online reviewers' behavior. *European Journal of Marketing* 47(10): 1758–1773. DOI:10.1108/EJM-11-2011-0625.
- Cook TD, Campbell DT. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin, Boston: MA.
- Dianoux C, Linhart Z. 2010. The effectiveness of female nudity in advertising in three European countries. *International Marketing Review* 27(5): 562–578. DOI:10.1108/02651331011076590.
- Dunn JR, Schweitzer ME. 2005. Feeling and believing: the influence of emotion on trust. *Journal of Personality and Social Psychology* 88(5): 736–748. DOI:10.1037/0022-3514.88.5.736.
- Ehrenberg A, Bound J. 1993. Predictability and prediction. *Journal of the Royal Statistical Society, Series A* 156(2): 167–206. DOI:10.2307/2982727.
- Ferber R. 1977. Research by convenience. *Journal of Consumer Research* 4(June): 57–58. DOI:10.1086/208679.
- Fetscherin M, Toncar M. 2010. The effects of the country of brand and the country of manufacturing of automobiles: an experimental study of consumers' brand personality perceptions. *International Marketing Review* 27(2): 164–178. DOI:10.1108/02651331021037494.
- Fiske ST. 2008. Perseverance furthers: aims of and sympathy for both student and nonstudent sampling. *Psychological Inquiry* 19(2): 106–107. DOI:10.1080/10478400802049829.
- Goldenberg J., Muller E. (2012). IJRM replication corner—structure and process. Retrieved from <http://www.journals.elsevier.com/international-journal-of-research-in-marketing/news/ijrm-replication-corner-2012/> (accessed 25 November, 2012)

- Green L, Myerson J, Lichtman D, Rosen S, Fry A. 1996. Temporal discounting in choice between delayed rewards: the role of age and income. *Psychology and Aging* 11(1): 79–84. DOI:10.1037/0882-7974.11.1.79.
- Henry PJ. 2008. Student sampling as a theoretical problem. *Psychological Inquiry* 19(2): 114–126. DOI:10.1080/10478400802049951.
- INSEE. (2012). Les résultats des recensements de la population [Results of the population census]. Retrieved from <http://www.insee.fr/fr/bases-de-donnees/default.asp?page=recensements.htm> (accessed 15 July 2015)
- James DO. 2006. Extension to alliance: Aaker and Keller's model revisited. *Journal of Product and Brand Management* 15(1): 15–22. DOI:10.1108/10610420610650846.
- James WL, Sonner BS. 2001. Just say no to traditional student samples. *Journal of Advertising Research* 41(5): 63–71.
- Jung JM, Polyorat K, Kellaris JJ. 2009. A cultural paradox in authority-based advertising. *International Marketing Review* 26(6): 601–632. DOI:10.1108/02651330911001314.
- Kabadayi S, Lerman D. 2011. Made in China but sold at FAO Schwarz: country-of-origin effect and trusting beliefs. *International Marketing Review* 28(1): 102–126. DOI:10.1108/02651331111107125.
- Kelley S Jr, Mirer TW. 1974. The simple act of voting. *American Political Science Review* 68(2): 572–591. DOI:10.1017/S000305540011740X.
- Kropp F, Lavack AM, Silvera DH. 2005. Values and collective self-esteem as predictors of consumer susceptibility to interpersonal influence among university students. *International Marketing Review* 22(1): 7–33. DOI:10.1108/02651330510581154.
- Laufer D, Silvera DH, McBride JB, Schertzer SMB. 2010. Communicating charity successes across cultures: highlighting individual or collective achievement? *European Journal of Marketing* 44(9/10): 1322–1333. DOI:10.1108/03090561011062862.
- Lefroy K, Tsarenko Y. 2013. From receiving to achieving: the role of relationship and dependence for nonprofit organizations in corporate partnerships. *European Journal of Marketing* 47(10): 1641–1666. DOI:10.1108/EJM-06-2011-0338.
- Liu F, Cheng H, Li J. 2009. Consumer responses to sex appeal advertising: a cross-cultural study. *International Marketing Review* 26(4/5): 501–520. DOI:10.1108/02651330910972002.
- Liu F, Li J, Mizerski D, Soh H. 2012. Self-congruity, brand attitude, and brand loyalty: a study on luxury brands. *European Journal of Marketing* 46(7/8): 922–937. DOI:10.1108/03090561211230098.
- Liu MT, Brock JL. 2011. Selecting a female athlete endorser in China: the effect of attractiveness, match-up, and consumer gender difference. *European Journal of Marketing* 45(7/8): 1214–1235. DOI:10.1108/03090561111137688.
- López M, Sicilia M. 2013. How WOM marketing contributes to new product adoption: testing competitive communication strategies. *European Journal of Marketing* 47(7): 1089–1114. DOI:10.1108/03090561311324228.
- Lucas JW. 2003. Theory-testing, generalization and the problem of external validity. *Sociological Theory* 21(3): 236–253. DOI:10.1111/1467-9558.00187.
- Lynch JG. 1982. On the external validity of experiments in consumer research? *Journal of Consumer Research* 9(3): 225–239. DOI:10.1086/208919.
- Lynch JG. 1983. The role of external validity in theoretical research. *Journal of Consumer Research* 10(1): 109–111. DOI:10.1086/208949.
- Lynch JG. 1999. Theory and external validity. *Journal of the Academy of Marketing Science* 27(3): 367–376. DOI:10.1177/0092070399273007.
- Maguire KB, Taylor LO, Gurnu S. 2003. Do students behave like adults? Evidence from valuation experiments. *Applied Economics Letters* 10(12): 753–756. DOI:10.1080/1350485032000138395.
- Malär L, Krohmer H, Hoyer WD, Nyffenegg B. 2011. Emotional brand attachment and brand personality: the relative importance of the actual and the ideal self. *Journal of Marketing* 75(4): 35–52. DOI:10.1509/jmkg.75.4.35.
- Martinez E, Polo Y, Chernatony LD. 2008. Effect of brand extension strategies on brand image: a comparative study of the UK and Spanish markets. *International Marketing Review* 25(1): 107–137. DOI:10.1108/02651330810851908.
- Mazodier M, Quester P, Chandon JL. 2012. Unmasking the ambushers: conceptual framework and empirical evidence. *European Journal of Marketing* 46(1/2): 192–214. DOI:10.1108/03090561211189284.
- Michaelis M, Woisetschläger DM, Backhaus C, Ahlert D. 2008. The effects of country of origin and corporate reputation on initial trust: an experimental evaluation of the perception of Polish consumers. *International Marketing Review* 25(4): 404–422. DOI:10.1108/02651330810887468.
- Mohr A, Batsakis G. 2014. Intangible assets, international experience and the internationalization speed of retailers. *International Marketing Review* 31(6): 601–620. DOI:10.1108/IMR-09-2013-0186.
- Nicolau JL. 2013. Direct versus indirect channels: differentiated loss aversion in high-involvement, non-frequently purchased hedonic product. *European Journal of Marketing* 47(1/2): 260–278. DOI:10.1108/03090561311285547.
- Norman R. 1975. Affective-cognitive consistency, attitudes, conformity and behavior. *Journal of Personality and Social Psychological* 32(1): 83–91. DOI:10.1037/h0076865.
- Oudou P. 2005. L'heuristique d'ancrage et d'ajustement comme mode d'évaluation d'un produit nouveau [The heuristic anchoring and adjustment as mode of evaluation of a new product]. *Recherche et Applications en Marketing* 20(3): 21–38. DOI:10.1177/076737010502000303.
- ONS. (2011). Dataset and reference tables. Retrieved from <http://www.ons.gov.uk/ons/datasets-and-tables/index.html> (accessed 10 July, 2015)
- Paolacci G, Chandler J, Ipeirotis P. 2010. Running experiments on Amazon mechanical turk. *Judgment and Decision Making* 5(5): 411–419.
- Paolacci G, Burson KA, Rick SI. 2011. The intermediate alternative effect: considering a small tradeoff increases subsequent willingness to make large tradeoffs. *Journal of Consumer Psychology* 21(4): 384–392. DOI:10.1016/j.jcps.2011.04.005.
- Peterson RA. 2001. On the use of college students in social science research: insights from a second-order meta-analysis. *Journal of Consumer Research* 28(3): 450–461. DOI:10.1086/323732.
- Petty RE, Cacioppo JT. 1996. Addressing disturbing and disturbed consumer behavior: is it necessary to change the way we conduct behavioral science. *Journal of Marketing Research* 1(1): 1–8. DOI:10.2307/3152008.
- Rampl LV, Kenning P. 2014. Employer brand trust and affect: linking brand personality to employer brand attractiveness. *European Journal of Marketing* 48(1): 218–236. DOI:10.1108/EJM-02-2012-0113.
- Schepers J, Wetzels M. 2007. A meta-analysis of the technology acceptance model: investigating subjective norm and moderation effects. *Information and Management* 44(1): 90–103. DOI:10.1016/j.im.2006.10.007.
- Sears DO. 1986. College sophomores in the laboratory: influences of a narrow database on social psychology's view of human nature. *Journal of Personality and Social Psychology* 51(3): 515–530. DOI:10.1037/0022-3514.51.3.515.
- Sears DO. 2008. College student-itis redux. *Psychological Inquiry* 19: 72–77. DOI:10.1080/10478400802050181.
- Segev S, Ruvio A, Shoham A, Velan D. 2014. Acculturation and consumer loyalty among immigrants: a cross-national study. *European Journal of Marketing* 48(9/10): 1579–1599. DOI:10.1108/EJM-06-2012-0343.
- Skarmeas D, Shabbir HA. 2011. Relationship quality and giving behaviour in the UK fundraising sector: exploring the antecedent roles of religiosity and self-construal. *European Journal of Marketing* 5(45): 720–738. DOI:10.1108/03090561111120000.

- UNSD. (2011). United Nations Statistic Division. Retrieved from <http://data.un.org/Explorer.aspx?d=POPandf=tableCode%3a29> (accessed 1st July, 2015)
- Van Doorn J, Verhoef PC. 2011. Willingness to pay for organic products: differences between virtue and vice foods. *International Journal of Research in Marketing* 28(3): 167–180. DOI:10.1016/j.ijresmar.2011.02.005.
- Völckner F, Sattler H. 2007. Empirical generalizability of consumer evaluations of brand extensions. *International Journal of Research in Marketing* 24(2): 149–162. DOI:10.1016/j.ijresmar.2006.11.003.
- Smith RW, Faro D, Burson KA. 2013. More for the many: the influence of entitativity on charitable giving. *Journal of Consumer Research* 39(5): 961–976. DOI:10.1086/666470.
- Steenkamp JBEM, Baumgartner H. 1998. Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research* 25(1): 78–90. DOI:10.1086/209528.
- Wang X, Yang Z. 2008. A meta-analysis of effect sizes in international marketing experiments. *International Marketing Review* 25(3): 276–291. DOI:10.1108/02651330810877216.
- Wang W, Sun S. 2010. Examining the role of beliefs and attitudes in online advertising. *International Marketing Review* 27(1): 87–107. DOI:10.1108/02651331011020410.
- Wells W. 1993. Discovery oriented consumer research. *Journal of Consumer Research* 19(4): 489–504. DOI:10.1086/209318.
- Wells W. 2001. The perils of N = 1. *Journal of Consumer Research* 28(3): 494–498. DOI:10.1086/323737.
- Winer RS. 1999. Experimentation in the 21st century: the importance of external validity. *Journal of the Academy of Marketing Science* 27(3): 349–358. DOI:10.1177/0092070399273005.
- Yan D, Duclos R. 2013. Making sense of numbers: effects of alphanumeric brands on consumer inference. *International Journal of Research in Marketing* 30(2): 179–184. DOI:10.1016/j.ijresmar.2012.09.007.
- Yim MYC, Sauer PL, Williams J, Lee SJ, Macrury I. 2014. Drivers of attitudes toward luxury brands: a cross-national investigation into the roles of interpersonal influence and brand consciousness. *International Marketing Review* 31(4): 363–389. DOI:10.1108/IMR-04-2011-0121.
- Yoon SJ. 2013. Antecedents and consequences of in-store experiences based on an experiential typology. *European Journal of Marketing* 47(5): 693–714. DOI:10.1108/03090561311306660.