The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Invited paper presented at the 6th African Conference of Agricultural Economists, September 23-26, 2019, Abuja, Nigeria

Copyright 2019 by [authors]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

What are the essential qualities of domestic biogas plants? Selecting attributes for a discrete choice experiment

*Stephen K. Mailu
Kenya Agricultural and Livestock Research Organization
Dairy Research Institute
kmailu@gmail.com

Collins U. Nwogwugwu
Nnamdi Azikiwe University, Awka,
Department of Economics
uchenwogwugwu@gmail.com

Kevin Kinusu
Kenya Biogas Programme
kkinusu@kbp.co.ke

Peterson Njeru
Kenya Agricultural and Livestock Research Organization
Food crops Research Institute, Muguga
Peterson.Njeru@kalro.org

Thomas Rewe
Pwani University, Kilifi
Department of Animal Sciences
t.rewe@pu.ac.ke

*Corresponding author: JEL codes [Q51, Q13, Q57, Q12]

Abstract
The paper discusses steps through which the selection of attributes of domestic biogas plants for use in a discrete choice experiment (DCE) can be implemented. From an initial list of more than 20 biogas plant attributes, the paper discusses the prioritization of nine which were subjected to two conjoint exercises featuring university students and researchers. These conjoint exercises conducted online ranked the most to the least important of nine attributes which could feature in a discrete choice experiment. The correlation between rankings by these groups of subjects is considered high enough to confirm consistency in the results. Gas production, installation cost, durability, maintenance cost, reliability and ability to translocate plants are the first six most important attributes in that order. These results are considered robust as input in a DCE designed for an audience of potential biogas plant users.
Keywords: conjoint analysis, choice experiment, attributes, ranking, focus group discussion

Introduction

The decision over which important attributes, usually from a longer list and which to use in a Choice Experiment (CE) is crucial. However, many times, the process for doing so is typically given casual treatment despite its essence in providing robust results from studies utilizing Stated Preference (SP) methods such as CE. As one of the valuation methods available, CEs are based on the notion that, any good (or service) can be described in terms of its attributes and the levels that these take. These attributes are the source of satisfaction that users of the good (service) derive from its use. From CEs, we can infer information about which attributes significantly influence choice, their ranking, the marginal willingness-to-pay (WTP) for a change in any significant attribute or the WTP for a programme which changes more than one attribute simultaneously. These are interpretations that are made possible by combining random utility theory with limited dependent variable econometrics (Hanley, et al., 1998).
A closely related method used mainly in the marketing literature is Conjoint Analysis (CA). Its use has grown since its popularization for practical analysis of rank ordered consumer judgment data (Orme, 2010). CA involves decomposing a designed set of multi-attribute alternatives into part-worth utilities (Louviere, 1988). CA, has been applied to provide calculations for royalty damages in patent infringement cases (Sidak & Skog, 2016) and in valuing cow and product attributes (Tano, et al., 2003; Makokha, et al., 2007; Fadiga & Makokha, 2014). CA traces its origins in psychology mainly in applications looking into ways towards representing the behavior of rankings observed as an outcome of systematic factorial manipulation of independent factors mathematically—theory of Conjoint Measurement (Louviere, et al., 2010).

In all, among goals of any SP study is the maximization of validity and reliability of the results (Johnston, et al., 2017). However, whereas many researchers do apply CE and CA in their work, most fail at least to report how the choice tasks are chosen for inclusion into the study (Louviere, et al., 2010). Some authors have described as questionable the rigor with which the first two stages—attribute development and the choice of levels of these attributes for discrete choice experiments are undertaken Coast & Horrocks, (2007) or are poorly reported (Coast, et al., 2012). Others have attempted to bridge this gap with suggestions on how these stages in CE design can be handled or provide some experiences gathered during the design stage (Kløjgaard, et al., 2012; Johnson, et al., 2013; Abiiro, et al., 2014; Helter & Boehler, 2016; Mangham, et al., 2009). Most, if not all the above articles are undertaken in the sub-field of health economics. The reasoning might be because health economics has had a long tradition in using CE in policy analysis (Louviere & Lancsar, 2009).

In this paper, we provide an account of the process undertaken in developing and choosing among various attributes of domestic biogas digesters in Kenya for later inclusion in a CE. There could be over ten different biogas digester designs in Kenya, each with its own set of attributes. Since the lack of awareness of the technology is a leading reason for non-adoption in Kenya (Mwirigi, et al., 2009; Muriuki, 2014), the study is timely as it can reveal the important attributes that technology developers and information providers can consider when promoting for widespread adoption. Enhancing these attractive attributes would be one means towards improving adoption of the technology.

This is because biogas production from animal manure, residues and on-farm produced or imported energy crops is a very promising option for generating renewable energy. Simultaneously, it reduces GHG (including CH_4) emissions both directly, through manure management, or by offsetting CO_2 emissions from fossil fuels and woodfuel while at the same reducing indoor pollution from smoke. It has been estimated that in Africa, biogas production has the potential to reduce deforestation due to woodfuel demand by between 6% and 36% in 2010 and between 4% and 26% in 2030 (Subedi, et al., 2014). African countries have been advised to develop innovative policy frameworks that support research, public awareness, standardization in addition to sustaining credit and subsidies to make these technologies affordable to poorer sections of the population (Mwirigi, et al., 2014).With a technology potential of over 170,000 biogas units in Kenya, (World Bank, 2016), there is potential for reducing reliance on other biomass energy sources even by half (Somanathan & Bluffstone, 2015). However, while introduced in Kenya in the 1950s, the adoption of biogas technologies has been marginal (Mwirigi, et al., 2009).

Nyumba et al. (2017) review studies relevant to biodiversity conservation and find that most studies utilizing stated preference methods do so alongside focus group discussions. In their study on milk and meat quality attributes, (Fadiga & Makokha, 2014) use a rapid market assessment to extract attributes for their study. Another CE study created hypothetical water source alternatives (Cook, et al., 2016) indicating the versatility of SP in creating conditions (attributes) that need to be studied even in their absence.

The paper is structured thus: A methodology section outlines the ranking exercises undertaken, whose results are presented in a subsequent section. A discussion concludes by consolidating evidence from three groups of respondents and highlights strengths and weaknesses identified in this study.

Materials and Methods

Establishing attributes and levels

The study began with an unstructured literature review mainly from published and unpublished manuscripts on the subject of domestic biogas. The literature comprised of theses, refereed journal papers, product catalogs from various biogas assembly companies, discussion papers, reports and other sources of grey literature. This choice of literature was grounded on studies incorporating any aspect of the attributes which helped with the definition of levels of the attributes. This literature helped give the research team a better idea about the possible attributes (and their levels) of a biogas plant. An initial listing of 17 attributes was generated from the literature (Table 1). This literature review was also strengthened through a two-pronged approach where consultations with specialists through unstructured interviews were made.

Focus group discussions

Three Focus Group Discussion sessions (FGDs), one each with farmers (FDG1: n=12 participants), biogas plant installation experts (trained masons and technicians) and entrepreneurs (FDG2: n=15 participants), researchers and policymakers (FDG3: n=8 participants) were conducted in July 2018 following the literature scan.

Some of the items in the FGD checklist included the skill requirement and post-installation support available for the systems as well as quality standardization including the ease of installing and managing biogas systems. The costs of installing and maintaining the different systems, their durability, output, sizes available as well as aesthetics and translocatability were other discussion points that were recorded during the FDGs. These attributes are descriptors of the three major types of biogas digesters commonly used in Kenya—tubular, fixed dome and floating drum. A moderator and two note takers were present at all these FGD sessions. The moderator used a FGD script, while the note takers used FGD note taking sheets to record proceedings. Participants were all advised of a number of points before the sessions began. Whereas no recording was made, they were informed by the moderator of note takers attending the sessions and that what each participant contributed was neither right nor wrong and that all their opinions were important and would be held in confidence. These FDGs clarified some attributes and also added a list of attributes to the list e.g. aesthetics, ability to identify defects, availability of skilled installers and ability to finance these through microcredit plans. The FDGs added to the initial list of 17 attributes derived from literature to a total of 22 attributes for consideration (Table 1).

Table 1: **Possible attributes and levels of domestic biogas plants**

	Attribute	Fixed dome	Floating Drum	Tubular
1[ac]	Durability / Expected lifespan (years)	12-15	15-20	2-5
2[ef]	Scalability	None	medium	Scalable
3[a]	Range of digester volume (m3)	6-91	5-70	5-20
4[abcdf]	Installation cost (Ksh/10m3)	102,000	92,500	65,000
5[c]	Maintenance cost	Moderate	High	Low
6[a]	Preferred substrate (feedstock)	Fibrous/non fibrous	Fibrous/non fibrous	non fibrous
7[a]	Daily gas output (m3 gas/m3 Vd)	0.2-0.5	0.3 - 0.6	0.3-0.8
8[c]	Ease of translocation	Difficult	Difficult	Easy
9[ce]	Aesthetics	neutral	unpleasant	pleasant
10[c]	Skill requirement at installation	Great	Great	not great
11[e]	Moving parts maintenance	none	Drum corrosion	none
12[c]	Labour required for construction	High	Moderate	Low
13[ad]	Gas pressure	high	high	low
14[c]	Days to commissioning	21	7-21	2
15[a]	Energy payback period (months)	24	26	17
16[ad]	Operational reliability	95%	92%	40%
17[a]	Gas pressure	High	Constant	Low
18[ce]	Identification of defects	Difficult	Easy	Intermediate
19[e]	Number of skilled installers	many	few	several
20[e]	Financeable through microcredit?	low	medium	high
21[cg]	Need for secure land tenure	high	medium	Low
22[c]	Escapes UV rays	NA	NA	Yes

Sources: [a]Nzila, et al., (2011) [b]Muriuki, (2014) [c]IFAD, (2012) [d]Mwirigi, et al., (2009) [e]FDG/consultations, [f]World Bank, (2016), [g]Mwirigi, et al., (2014)
1 USD = Ksh 100.5 in October 2018

Preliminary reduction of attribute tally

The 22 attributes in Table 1 were then passed over to a technical team at the Kenya Biogas Programme who wilted these attributes down to nine, after a careful consideration of the merits of the attributes themselves and the ease with which each could be described to, and observed by non-experts. Other reasons for dropping some attributes were that they were highly subjective. For instance, the attribute *aesthetics* is fraught with a-lot of subjective judgements over what is good or poor aesthetically. Others e.g. *number of skilled artisans* which was picked up mentioned during the FGDs was dropped since this attribute was not precisely a physical attributes of different plants per se. ability to finance the biogas plant through micro-credit and the requirement of secure land title were thought to reflect the same concept. The remaining number (nine) attributes was also dictated by the maximum number of attributes usable with the online conjoint software available to the researchers. The remaining nine attributes that were under consideration include: *durability, installation cost, maintenance cost, daily gas output, ease of translocation, operational reliability, gas pressure, consistency of gas output* and *ability to identify defects in gas holder*. Since the research team felt that nine attributes would still be too many to include in a choice experiment Alpizar et al., (2001), the next stage involved prioritizing the list of attributes so that the most important candidate attributes could feature in a CE.

Ranking exercise 1: University students

In keeping with the findings that self-selected student subjects are an appropriate subject pool for the study of social behaviour Exadaktylos, et al., (2013), contact with colleagues at one public university was initiated. Using a non-technical instruction prepared for use by a research assistant, students were made aware of the purpose of the assignment that they were about to undertake. The instruction described the PAPRIKA method as well as its implementation in meenymo. PAPRIKA[1] is one of the Multiple Criteria Decision-Making (MCDM) method which involves the decision-maker pairwise rank potentially all undominated pairs of all possible alternatives. Since the publishing of the method (Hansen & Ombler, 2009) and launch of 1000minds, the PAPRIKA method has been used in a host of applications ranging from breeding (Slagboom, et al., 2016; Martin-Collado, et al., 2015; Byrne, et al., 2016), environmental management (Chhun, et al., 2013; de Olde, et al., 2017), marketing research (Wijland, et al., 2016; Lee, et al., 2015), land use and urban planning, among other instances which present the problem of deciding among attributes. In our study, each choice question was based on two hypothetical alternatives from the list of 9 attributes defined on just two alternatives at a time and involving a trade-off (see sample screenshot on top-left of Figure 1).

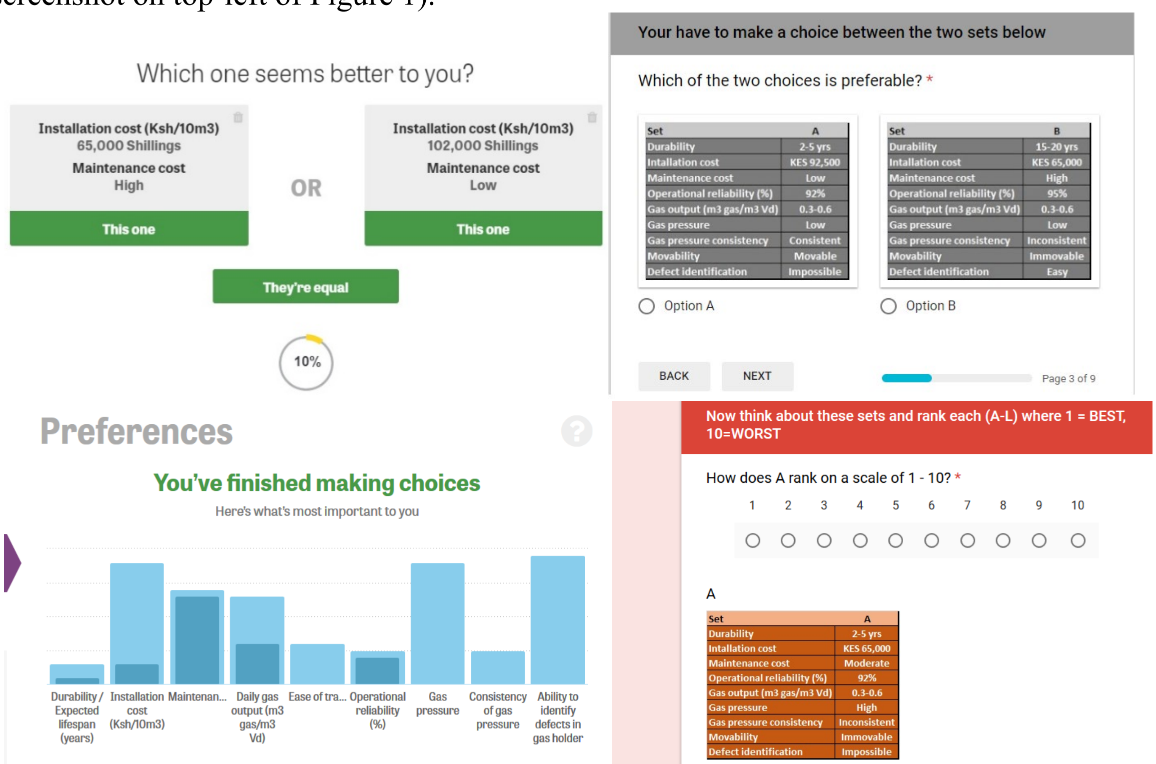

***Figure 1*: Sample screenshots of candidate profiles in the meenymo (upper-left) and googleforms (upper-right) that the result of the ranking by individual student (lower-left) and the ranking screen in googleforms (lower-right)**

[1] Besides PAPRIKA, other tools include DEFINITE; MEM - Multiplex Electionis Methodus; MCDA Package for R.; FITradeoff A Flexible and Interactive Tradeoff elicitation procedure for multicriteria additive models; 1000Minds; BENSOLVE Free MatLab implementation of Benson's algorithm to solve linear vector optimization problems; DEXi; GUIMOO, a Graphical User Interface for Multi Objective Optimization; MACBETH-Measuring Attractiveness by a Categorical Based Evaluation TecHnique; modeFRONTIER

Since the invitation was open to students who wished to contribute data, only 49 students from the faculty of agriculture and 10 students from the faculty of engineering from one of the public universities in Kenya were able to volunteer time to complete the ranking exercise (Table 2). This represents roughly 35 percent of the group. This exercise was undertaken within a span of 1 week during the month of September 2018. As with other software tools, results for each individual are easily summarized without the need for extra much data entry; the only recording required for the research assistant being the transcription of the result for each respondent from the screen (see sample screenshot bottom-left on Figure 1).

Table 2: **Sample sizes of different groups**

Group	Group details	No. invited	No. responding	Response rate (%)	
Students[a]	AgEng	40[c]	10	25.0	34.7
	AgEcon	90	29	32.2	
	AgBus	40	20	50.0	
FDG participants	FDG1	7[e]	1	14.3	36.8
	FDG2	17[f]	7	41.2	
	FDG3	14[g]	6	42.9	
KALRO	KALRO[b]	345	63	18.2	18.8
	DRI	17	7	41.2	
Total			143		

Notes: [a]It had been planned to have another set of students from Nnamdi Azikiwe University, Awka but during the period, a strike took away the possibility of engaging students there. [c] Due to logistical difficulties, the research assistant snowballed to get this group of students. [b]This group has a collection of researchers from 9 of 16 KALRO research institutes (food crops, agricultural mechanization, non-ruminants, industrial crops, horticulture, sheep and goats, beef, sugar, and veterinary sciences). [e,f,g] The numbers do not total the list of FDG participants since some emails were randomly assigned to receive more than one invitation.

Ranking exercise 2: Kenya Agricultural and Livestock Research Organization (KALRO) scientists & FGD participants

An online tool was developed in GoogleForms and which respondents could interact with and make their choices among sets of all nine attributes was presented in different attribute level configurations (profiles). The alternatives were given generic labels (option A, B, C….) rather than their alternative specific form (e.g. Floating Drum) as the former has been said to increase a respondent's attention to attributes and therefore suitable to investigate the trade-offs between attributes. In this case, a short description of each attribute as presented to the students was made available in the introductory section to this online tool. The questionnaire comprised of six multiple choice questions and 12 profiles to rate on a linear scale presented as choice buttons (see screenshot top & bottom-right in Figure 1 for a sample). The scale was end-anchored with 1=BEST and 10=WORST and presented as a series of numbers spaced out across the page. An early study has indicated that a scale with ten response categories can yield reliable scores and has greater discriminating power as opposed to one with fewer response categories (Preston & Colman, 2000).

The profile method (as opposed to the two-factor trade-off method) was chosen since it was believed that time demands for scientists might be high for them to engage in a trade-off study. Since there are nine factors, four at two levels and five at three levels (denoted $2^4 3^5$), there are $2\times2\times2\times2\times3\times3\times3\times3\times3 = 3{,}888$ combinations in a full-factorial design which would be very large and intractable if used in an experiment (Alpizar, et al., 2001). Hensher, (2006) studied design dimensions (number of choice sets, number of attributes and range of attribute levels) on choices and didn't find evidence of statistically significant systematic differences due to any design dimensions. Using the list of attributes and levels, we generated an orthogonal linear D-optimal design using STATA (ver 13) comprising 48 choice profiles. To reduce the burden on respondents, the 48 choice profiles, each featuring nine attributes[2] were blocked into four groups, each containing 12 choice profiles and respondents randomly assigned into the 4 groups/blocks. We chose to go with the full profile since in a study studying national parks, the number of attributes didn't seem to matter in terms of individual choices and aggregate choice shares (Pullman, et al., 1999). In the end, each respondent was confronted by a choice over 6 pairs of profiles and a ranking exercise of the 12 profiles—18 decisions in total.

A list of 345 different email addresses of participants at KALRO in addition to 17 from the Dairy Research Institute as well as 38 from the FGDs participants were recipients of the invitation to undertake the online survey. A summary of this distribution is shown in Table 2. The email invitation was accompanied by a letter that gave a brief background and reason for the exercise that they had been invited to participate in. The body of the email contained a link to the project website and further elaboration of the task and an estimated last date by which time the survey would be closed. A reminder was sent to all email addresses one day before the deadline (2 weeks post invitation). This survey was open from 5/11/2018 to 15/12/2018. A number of the design principles for online surveys outlined by Dillman, et al., (2009) were used in constructing the survey instrument including the accompanying communication.

In total, 63 responses were received from the KALRO researchers, seven from DRI and 14 from the FDG participants making a total of 84 individual rankings. This represents a response rate of 18.2% from KALRO, 36.8% from the FGDs and 41% from DRI. The seemingly higher response rate from DRI (Table 2) reflects the personal request made to some of the researchers unlike reminders offered via email to the other groups—KALRO researchers and FDG participants. Given the generally low response rates to online surveys, the response achieved from the scientists and FGD participants can be considered satisfactory.

Checking for Consistency

A high correlation between ratings by different judges reduces the probability of type-II errors, as the low noise supports our ability to detect relationships that actually exist (Hallgren, 2012). To reduce the probability of these kind of errors, we opt to check that the consistency of rankings made by the different groups shown in Table 2 is consistent. The interest here is in the reliability of the mean ratings provided by all judges (8 groups). The hypothesis here is that the different groups maintain similar ranking consistency. We undertake to check whether the different groups consistently placed the attributes in a similar order of preference. The interest is not in the degree that groups agreed in the absolute values of their rankings but in the consistency, warranting a consistency type ICC (Hallgren, 2012; Feng, 2014). We use the benchmark correlation of 0.7 considered as an acceptable level of internal consistency of items during the early stages of research (Nunnally, 1978). The issue of what benchmark to use has been revisited by various authors (Banerjee, et al., 1990; Cicchetti, 1994; Lance, et al., 2006; LeBreton & Senter, 2008)[3].

[2] Mazotta and Opaluch (1995) cited in Alpizar et al. (2001 p19) indicate that including more than 4-5 attributes in a choice set could lead to quality deterioration due to task complexity.

[3] We do also acknowledge that according to Nunnally (1978, pp245-246), the 0.8 cut-off for basic and applied research is not nearly high enough.

More formally, consider *n* targets (replications) that are randomly sampled from a population of interest. Each target is rated independently by a set of *k* raters (judges), yielding one rating per target and rater. It is of interest to determine the extent of the agreement of the ratings; otherwise, without any congruence in the ratings, the tool used cannot be said to yield reliable results. We use a two-way mixed-effects model since each target is rated by the same set of *k* independent raters. The "Two-Way Mixed" model is so called since 1) it models both an effect of judge and of the target—two effects and 2) assumes a random effect of target but a fixed effect of rater (Landers, 2015).

Results and Discussion

Ranking exercises

After several weeks granted to undertake the ranking exercises, only about 20 percent (70 of 362 emails) of the researchers had provided their rankings and 37 percent (14 of 38 emails) of the FGD participants, an average response rate of 21 percent. In all, there were 84 completed rankings from the googleforms tool and 60 from the group of students which together form part of the data discussed in the next paragraphs. This makes a sample size of 143. In a study such as this, which is mainly investigational trying to seek a simple ranking of attributes, we think that this size is adequate for the task as it is higher than the thirty or sixty respondents suggested by (Orme, 2010).

Ranking exercise 1: Students

Using the PAPRIKA method, the group of students collectively ranked the attributes in decreasing order of importance thus; gas output, ability to translocate, installation cost, maintenance cost, durability, gas pressure, operational reliability, consistency of gas output and ability to identify defects (Table 3). This ranking is recovered by simply converting the relative rank of each attribute (ranked from 1-9) into a percentage by dividing each row (representing a respondent) and columns representing attributes by the row total (=45). The figure 45 is the maximum rating that any respondent can give to the 9 attributes (9+8+7+6+5+4+3+2+1=45). In certain instances, respondents were indifferent with respect to certain attributes (tied rankings). In that case, assuming the leading ranked attributes tied, then the next (3rd) ranked attribute is given a rank of 4, rather than 3. This ensures that the relative ranking of attributes is maintained. The average of this rating (%) is then used to rank the different attributes as presented in Table 3.

***Table 3*: Ratings (%) of different attributes by class of student**

Attribute	Class			All
	1: AgEng	2: Ag.Bus	3: AgEcon	
N	10	20	29	59
Gas production	16.7	14.9	14.5	15.0
Translocation	17.7	14.8	13.8	14.8
Installation cost	10.1	15.4	14.0	13.8
Maintenance cost	13.7	11.9	12.5	12.5
Durability	12.9	9.2	10.0	10.3
Reliability	8.6	9.5	9.1	9.2
Pressure	8.1	8.8	9.8	9.2
Consistency of gas	7.0	9.3	8.4	8.4
Defect identification	5.3	6.2	7.9	6.9

Note: This table is from the 2 attribute trade-off task and contains results from 29 agricultural economics, 20 agribusiness management and 10 engineering students.

For the data from students, rankings are transformed into a scale of 1-10[4] and the Interclass Correlation (ICC) calculated to check for consistency among students. This transformation was done in order to obtain data that is comparable to that from the KALRO and FDGs conjoint study. This yields an ICC(C,k) value of 0.94, ($F_{(8, 464)}$ = 19.04, p<0.001, 95% [CI 0.88-0.98]. The conclusion here is that the 59 students consistently rated the different attributes. For AgEng students, the ICC(C,k) is 0.89, ($F_{(8, 72)}$ = 9.66, p<0.000, 95% CI [0.75, 0.97]. For the AgBus students class, the ICC(C,k) calculated yields 0.88, ($F_{(8, 152)}$ = 8.16, p<0.000, 95% CI [0.72, 0.97] while AgEcon class shows an ICC(C,k) estimated as 0.84, ($F_{(8, 224)}$ = 6.4, p<0.000, 95% CI [0.64, 0.96]. These results imply some respectable level of consistency since in all three cases, the ICC(C,k) was higher than the 0.7 benchmark.

A series of two-sample Wilcoxon rank-sum (Mann-Whitney) tests are run to test the hypotheses: all classes rate the attributes in a similar manner i.e. $U1_{ij} = U2_{ij},\ (for\ i = 1..3; j = 1 \dots 9)$; where U1 and U2 are the two classes being compared. All comparisons save for 3 cases out of 28 do we reject the null hypothesis (p<0.05). Installation cost for the engineering class was significantly lower (Mdn=5) than for the agribusiness class (Mdn=7.75), *Z= 2.659, p=0.0078,* r^5*=0.48*. Conversely, the ease of translocation was rated significantly higher by the engineering class (Mdn=8.5) compared to the agribusiness class (Mdn=7), *Z=-1.974, p=0.0484, r=0.36*. The engineering class rated installation cost significantly differently (Mdn=5) than the agricultural economics class (Mdn=7), *Z= 2.342, p=0.0192, r=0.37*.

Ranking exercise 2: KALRO researchers & FGD participants

From the ranking data collected using the online tool presented in googleforms, 84 respondents (out of an expected 400) were able to complete the survey. This represents a response rate of 21% which is not unexpected since online surveys have been noted to suffer low response rates. These respondents were able to rank 12 different profiles from 1-10 (1=best, 10=worst). Participants answered the question "Now think about these sets and rank each (A-L) where 1 = BEST, 10=WORST" denoting a 10-point scale (forced choice). Ideally, one would have expected researchers to use 12 ranks but they were forced to do with the limitations with the online tool chosen[6]. This exercise was conducted for 40 straight days following reminders on day 20 following the initial invitation to the exercise. This survey period was ended when it was clear that no more responses were expected. Ranks ranged from 1-10 and preliminary analysis seems to indicate that the median rating for all choice profiles was 5. This can be reflective of the fact that since the scale (best-worst) can be seen as a two-polar scale, a rating of 5 may be an answer to valid rankings: "neutral," "equal/both," or other interpretations such as "no opinion/don't know," "unsure," or "neither". Some researchers have shown that when this middle option is offered, it is more likely to be chosen (Moors, 2008).

In order to have comparable data as that from data 1 above, we need to transform the profile ranks into ranks for specific attributes. To do so we run a conjoint analysis. We also consider that it would be safer to compute importances for respondents individually and then average them, rather than computing importance from average utilities when summarizing attribute importance for groups (Orme, 2010). We run the conjoint analysis and extract the relative importance for each respondent in the 8 groups/blocks. The conjoint analysis is implemented using the marketing research GUI available in SAS® (ver 9.0)

[4] The maximum rank is 9 in the original data since these ranks represent the complete set of attributes as ranked using the meenymo© online tool.

[5] A measure of effect size, r is calculated by dividing Z by the square root of N i.e. (r =Z /√N).

[6] GoogleForms provides survey developers options to offer respondents: checkboxes, choice grids, linear scales, and drop down menus, checkboxes, or even multiple choice answers. We chose to go with the linear scale since it comes closest to a slider which some researchers' state eases the burden on respondents in such hypothetical tasks.

(Kuhfeld, 2009). From here, we extract part-worth utilities for each attribute and thus, their weights, which represent the relative importance of the attributes. This procedure repeated separately for each of the 8 blocks and is the source of the results displayed in the columns labelled 1-4 in Table 3.

The result for these two groups (Table 3) shows a generally similar ranking as that from the students, save for the ability to translocate the biogas plant which for students ranks second after gas production. Whereas in this study we are not overly concerned with the absolute ranking (or rating), this switch in position points to the wisdom of seeking multiple points of view. For the consistency between ratings from the conjoint experiment with 8 groups, the ICC(C,k) is calculated and yields a value of 0.91, $F_{(8, 56)} = 12.49$, $p<0.001$, 95% CI [0.80, 0.97]. This result suggests that the results can be considered consistent), thus showing a reliable result from the conjoint experiment.

Table 4: **Ratings (%) of different attributes for KALRO researchers and FGD participants**

	Rating score								
	KALRO researchers[a]				FDG participants				**Average**
	Block[a]				Block[a]				
	1	2	3	4	1	2	3	4	
N	17	13	19	21	4	3	4	3	**84**
Gas production	24.4	15.1	25	20.1	22.7	16.7	26	21.2	**20.7**
Installation cost	16.5	19.2	25.6	13.2	18.7	20	28.7	12.5	**18.7**
Durability	9.5	16	19.6	24	5.2	14.9	16.8	23.5	**15.5**
Maintenance cost	14.7	16.8	9.1	18.2	16.9	16.1	11.3	21.4	**15.2**
Reliability	7	22.3	13	0	6.2	23.4	12.4	0	**10.4**
Translocation	11.3	0	0	11.2	15.2	0	0	17	**7.7**
Pressure	10.2	10.7	7.6	0	11.5	8.9	4.8	0	**7.0**
Consistency of gas	6.5	0	0	6.4	3.6	0	0	2	**3.0**
Defect identification	0	0	0	6.8	0	0	0	2.4	**1.8**

Note: The total score along the column is a maximum of 100 representing the total weight given to all nine attributes with a few exceptions due to rounding.
[a] Each of the 4 blocks contains 12 different randomly assigned choice profiles partitioned at the design stage.

We can test whether the FGDs and KALRO researchers were able to rank the choice sets similarly. We test the hypothesis $U1_{ij} = U2_{ij},\ (for\ i = 1...4; j = 1 \ldots 12)$; where there are 4 blocks and 12 choice sets per block. A series of the two-sample Wilcoxon rank-sum (Mann-Whitney) tests was ran and indicated that in each case (for all but one choice set), researchers and FGD participants did not rank the different choice sets significantly differently ($p \leq 0.05$). This is with the exception of one choice set where in block 4, one choice set was rated (Mdn=6) by the scientists but (Mdn=3) by the FDG participants, *Z=-2.296, p=0.0217, r=0.47.*

<u>*Combining ranking exercise 1 & 2*</u>

The final task involved check for consistency between the 9 groups now treating the data from the ranking from study 1 (students) and 2 (researchers & FGD participants). This correlation yields the result is a correlation ICC(C,k) of 0.917, $F_{(8, 64)} = 12.18$, $p<0.001$. This result provides evidence of a reasonably robust result, leading us to conclude that students, FDG participants, and KALRO researchers consider attributes in the sequence presented in *Table 2* and *Table 3*.

Conclusion

At the present stage in the overall project under which this study was undertaken, the objective was not to estimate welfare (WTP) values. This paper however contributes to the existing practice of making choices among attributes for inclusion in a CE. Here, we have demonstrated the mixed approach to obtaining and prioritization of attributes for use in a CE. This paper adds to the literature about pre-experimental design issues for discrete choice experiments, giving a relatively more robust justification for choosing one attribute and not the other based on responses from at least three different groups. In this study, *gas production* is an important attribute of any biogas plant. So too is the *installation cost* and *plant durability*.

Furthermore, with the two groups (KALRO scientists and FGD participants) showing no appreciable difference in their ranking, this increases the confidence about the robustness of the results. This is strengthened by the similarity in ranking as done by the students who used a different approach (PAPRIKA). One important practical result for this work is that these results already stand as a rough benchmark upon which data from the actual CE can be evaluated—at least as far as the relative preference for domestic biogas attributes goes.

Acknowledgments
This work has been supported through support from the National Research Fund (grant NRF/R/2016/045). Carol Michuki and Teresiah Njeri from KBP are acknowledged for making arrangements to meet with a highly engaging group of farmers. The authors welcome the useful input from Hillary Ndambiri who moderated the FGD sessions and James Kinyanjui for his assistance in note taking and presenting the tool to University students. Simon Mugi was instrumental in arranging logistics and communication throughout the study period. Participants at the FGDs who took time off from their busy schedules to attend are also greatly appreciated. Eng. Mike Wasonga is also appreciated for giving the team a more nuanced look at the long list of attributes and levels. All respondents who took part in the exercise are also thanked for making their choices known. All remaining errors are those of the authors.

Ethics statement
All participants voluntarily opted to the study. Their anonymity is preserved by assigning respondents a random numerical code, which would identify them in the data. No association has been made between their real names and the results as such personal data such as names, sex etc was not collected.

References
Abiiro, G. A., Leppert, G., Mbera, G. B., & Robyn, P. J. (2014). Developing attributes and attribute-levels for a discrete choice experiment on micro health insurance in rural Malawi. *BMC Health Services Research*, 14(235).

Alpizar, F., Carlsson, F., & Martinsson, P. (2001). Using choice experiments for non-market valuation. (Working papers in Economics #52): Department of Economics, Gothenburg University.

Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1990). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics*, 27(1), 3-23.

Byrne, T., Santos, B., Amer, P., Martin-Collado, D., Pryce, J., & Axford, M. (2016). New breeding objectives and selection indices for the Australian dairy industry. *Journal of Dairy Science*, 99(10), 8146-8167. doi:10.3168/jds.2015-10747

Chhun, S., Thorsnes, P., & Moller, H. (2013). Preferences for management of near-shore marine ecosystems: A choice experiment in New Zealand. *Resources*, 2(3), 406-438. doi:10.3390/resources2030406

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.

Coast, J., & Horrocks, S. (2007). Developing attributes and levels for discrete choice experiments using qualitative methods. *Journal of Health Services Research & Policy*, 12(1), 25-30. doi:10.1258/135581907779497602

Coast, J., Al-Janabi, H., Sutton, E. J., Horrocks, S. A., Vosper, J. A., Swancutt, D. R., & Flynn, T. N. (2012). Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics*, 21(6), 730-741. doi:10.1002/hec.1739

Cook, J., Kimuyu, P., Blum, A. G., & Gatua, J. (2016). A simple stated preference tool for estimating the value of travel time in rural Africa. *Journal of Benefit Cost Analysis*, 7(2), 221–247. doi:10.1017/bca.2016.13

de Olde, E. M., Moller, H., Marchand, F., McDowell, R. W., MacLeod, C. J., Sautier, M., . . . Legun, K. A. (2017). When experts disagree: the need to rethink indicator selection for assessing sustainability of agriculture. *Environment, Development and Sustainability*, 19(4), 1327–1342.

Dillman, D. A., Smyth, J. D., & Christian, L. (2009). Internet, mail and mixed-mode surveys: The tailored design method (3rd ed.). Hoboken, NJ: John Wiley.

Exadaktylos, F., Espín, A. M., & Brañas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports*, 33(1213), 6pp. doi:10.1038/srep01213

Fadiga, M. L., & Makokha, S. N. (2014). Consumer valuations of the quality and safety attributes of milk and meat in Kenya. *African Journal of Agricultural and Resource Economics*, 9(2), 1-15.

Feng, G. C. (2014). Mistakes and how to avoid mistakes in using intercoder reliability indices. Methodology: European *Journal of Research Methods for the Behavioral and Social Sciences*. doi:10.1027/1614-2241/a000086

Hallgren, K. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.

Hanley, N., MacMillan, D., Wright, R., Bullock, C., Simpson, I., Parisson, D., & Crabtree, B. (1998). Contingent valuation versus choice experiments: Estimating the benefits of environmentally sensitive areas in Scotland. *Journal of Agricultural Economics*, 49(1), 1-15.

Hansen, P., & Ombler, F. (2009). A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *Journal of Multi-Criteria Decision Analysis*, 15, 87-107. doi:10.1002/mcda.428

Helter, T. M., & Boehler, C. E. (2016). Developing attributes for discrete choice experiments in health: a systematic literature review and case study of alcohol misuse interventions. *Journal of Substance Use*, 21(6), 662–668. doi:10.3109/14659891.2015.1118563

Hensher, D. (2006). Revealing differences in willingness to pay due to the dimensionality of stated choice designs: An initial assessment. *Environmental and Resource Economics*, 34, 7–44. doi:10.1007/s10640-005-3782-y

IFAD. (2012). Flexi Biogas systems: inexpensive, renewable energy for developing countries. Rome, italy: International Fund for Agricultural Development.

Johnson, R.F., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D.A., Bresnahan, B.W., Kanninen, B., Bridges, J.F.P. (2013). Constructing experimental designs for discrete-choice experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value in Health*, 16, 3-13. doi:10.1016/j.jval.2012.08.2223

Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A.W., Hanemann, M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., Vossler, C.A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), 319-405.

Kløjgaard, M. E., Bech, M., & Søgaard, R. (2012). Designing a stated choice experiment: The value of a qualitative process. *Journal of Choice Modelling*, 5(2), 1-18.
Kuhfeld, W. F. (2009). Marketing research methods in SAS: Experimental design, choice, conjoint, and graphical techniques (SAS 9.2 ed.). Cary, NC, USA: SAS Institute Inc.
Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, 9(2), 202-220. doi:10.1177/1094428105284919
Landers, R. (2015). Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS. *The Winnower*, 2(e143518.81744), 1-4. doi:10.15200/winn.143518.81744
LeBreton, J., & Senter, J. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, 11(4), 815-852. doi:10.1177/1094428106296642
Lee, P. Y., Lusk, K., Mirosa, M., & Oey, I. (2015). An attribute prioritization-based segmentation of the Chinese consumer market for fruit juice. *Food Quality and Preference*, 46, 1-8. doi:10.1016/j.foodqual.2015.06.016
Louviere, J. (1988). Conjoint analysis modelling of stated preferences: A review of theory, methods, recent developments and external validity. *Journal of Transport Economics and Policy*, 22(1), 93-119.
Louviere, J. J., & Lancsar, E. (2009). Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Economics, Policy and Law*, 4, 527–546. doi:10.1017/S1744133109990193
Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling*, 3(3), 57-72.
Makokha, S., Karugia, J., Staal, S., & Kosura-Oluoch, W. (2007). Valuation of cow attributes by conjoint analysis: A case study of western Kenya. *African Journal of Agricultural and Resource Economics*, 1(2), 95-113.
Mangham, L., Hanson, K., & McPake, B. (2009). How to do (or not to do)...Designing a discrete choice experiment for application in a low income county. *Health Policy and Planning*, 24, 151-158. doi:10.1093/heapol/czn047
Martin-Collado, D., Byrne, T., Amer, P., Santos, B., Axford, M., & Pryce, J. (2015). Analyzing the heterogeneity of farmers' preferences for improvements in dairy cow traits using farmer typologies. *Journal of Dairy Science*, 98(6), 4148-4161. doi:10.3168/jds.2014-9194
Moors, G. (2008). Exploring the effect of a middle response category on response style in attitude measurement. *The Journal of General Psychology*, 42(6), 779–794. doi:10.1007/s11135-006-9067-x
Muriuki, S. W. (2014). Analysis of biogas technology for household energy, sustainable livelihoods and climate change mitigation in Kiambu county, Kenya. Kenyatta University, Kenya: (Unpublished doctoral dissertation).
Mwirigi, J. W., Makenzi, P. M., & Ochola, W. O. (2009). Socio-economic constraints to adoption and sustainability of biogas technology by farmers in Nakuru Districts, Kenya. *Energy for Sustainable Development*, 13, 106-115.
Mwirigi, J., Balana, B., Mugisha, J., Walekhwa, P., Melamu, R., Nakami, S., & Makenzi, P. (2014). Socio-economic hurdles to widespread adoption of small-scale biogas digesters in Sub-Saharan Africa: A review. *Biomass and Bioenergy*, 70, 17-25.
Nyumba, T. O., Wilson, K., Derrick, C. J., & Mukherjee, N. (2017). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution* (9), 20-32.
Nzila, C., Dewulf, J., Spanjers, H., Tuigong, D., Kiriamiti, H., & van Langenhove, H. (2011). Multi criteria sustainability assessment of biogas production in Kenya. *Applied Energy*, 93, 496–506.

Orme, B. K. (2010). Getting started with conjoint analysis: Strategies for product design and pricing research (Second ed.). Madison, Wisc: Research Publishers LLC.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1-15.

Pullman, M. E., Dodson, K. J., & Moore, W. L. (1999). A comparison of conjoint methods when there are many attributes. *Marketing Letters*, 10(2), 125–138.

Sidak, G., & Skog, J. O. (2016). Using conjoint analysis to apportion patent damages. *The Federal Circuit Bar Journal*, 25(4), 581-620.

Slagboom, M., Kargo, M., Edwards, D., Sørensen, A., Thomasen, J., & Hjortø, L. (2016). Organic dairy farmers put more emphasis on production traits than conventional farmers. *Journal of Dairy Science*, 99(12), 9845–9856. doi:10.3168/jds.2016-11346.

Somanathan, E., & Bluffstone, R. (2015). Biogas: Clean energy access with low-cost mitigation of climate change. Washington D.C.: World Bank Policy Research Working Paper 7349.

Subedi, M., Matthews, R. B., Pogson, M., Abegaz, A., Balana, B. B., Oyesiku-Blakemore, J., & Smith, J. (2014). Can biogas digesters help to reduce deforestation in Africa? *Biomass and Bioenergy*, 70, 87-98.

Tano, K., Kamuanga, M., Faminow, M. D., & Swallow, B. (2003). Using conjoint analysis to estimate farmer’s preferences for cattle traits in West Africa. *Ecological Economics*, 45, 393-407.

Wijland, R., Hansen, P., & Gardezi, F. (2016). Mobile nudging: Youth engagement with banking apps. *Journal of Financial Services Marketing*, 21(1), 51–63.

World Bank. (2016). Promoting biogas as sustainable clean cooking fuel for rural households in Kenya project. Washington D.C.: World Bank.