Identify Lao farmers' goals and their ranking using Best-Worst Scaling Experiment and Scale-adjusted Latent Class Models

Damien Jourdain[a,b,c*], Juliette Lairez[d,e], François Affholder[e,f]

* Corresponding author.
e:mail: damien.jourdain@cirad.fr
tel: +27 79 800 01 63
ORCID: 0000-0001-8707-5221

[a] CIRAD, UMR G-EAU, Hatfield 0028, South Africa
[b] G-EAU, Univ Montpellier, AgroParisTech, CIRAD, IRD, IRSTEA, Montpellier SupAgro, Montpellier, France
[c] Centre of Environmental Economics and Policy for Africa (CEEPA), Faculty of Agricultural Economics, University of Pretoria, Hatfield 0028, South Africa
[d] CIRAD, UPR AIDA, Vientiane, Laos.
[e] AIDA, Univ Montpellier, CIRAD, Montpellier, France.
[f] CIRAD, UPR AIDA, F-34398 Montpellier, France.

Abstract

In order to better design more sustainable farming systems, and prepare for the development of multi-criteria farm decision model, we investigate how farmers rank their main goals when making decisions. First, we identified the main goals used by farmers through in-depth interviews with randomly selected farmers in which we used small games to elicit the main goals they are using to make farm-level decisions. Then, we developed a Best-Worst Scaling experiment, in which farmers have to declare the « most » and the least « important » goals they use when making decisions. The experiment was conducted with 120 farmers. We first derive a ranking of the goals according to the population average, which showed the importance of rice self-sufficiency and transmission of farm capital. We then use a scale-adjusted latent class analysis. We identified four groups of homogenous preferences among farmers. The use of differentiated scale, a measure of choice inconsistencies, suggested different levels of certainty about the ranking, and the presence of more inconsistencies when asking the least important goal. While a large group focuses only on rice self-sufficiency, and farm transmission, we also identified a group of optimizers, and risk-averse farmers. Farmers of each group are likely to behave differently with regard to sustainable innovations. We also showed that some socio-economic variables describing the farms and the households influenced the probabilities for farmers to belong to one of the four classes. Overall, we showed that best-worst scaling

experiments provide a rich set of information about the diversity of rankings. It also provides the set of tools to evaluate the consistency and quality of respondents' choices.

Acknowledgements : We beneficiated from fruitful discussions with and data made available by the Sustainable Agricultural Land Management Initiative of the Xieng Khouang province in which the NUDP-Eficas project (Eco-Friendly Intensification and Climate Resilient Agricultural Systems in Lao PDR) supports farmers to develop more sustainable and resilient production systems.

Funding: This study was funded by the French Development Agency (AFD) and the European Union (EU) within the EFICAS project Landscape Management and Conservation Agriculture Development for Eco-Friendly Intensification and Climate Resilient Agricultural Systems in Lao PDR.

1 Introduction

A transformation of current agricultural production systems is needed to create healthy, equitable, resilient and sustainable food systems (FAO, 2010, Garnett, et al., 2013). This is particularly urgent in the developing world where agriculture systems are practiced in fragile environments (Tilman, et al., 2011). The adoption of new agricultural practices based on sound agronomic principles – e.g., diversification of cropping systems (rotations, intercropping), terracing, etc. – has the potential to significantly increase the sustainability and resilience of the current agricultural production systems while maintaining the overall productivity of the land (Frison, et al., 2011, Gaudin, et al., 2015). However, adoption of these practices is slow and it is important but difficult to predict the acceptability of these new practices by farmers under different environmental and policy scenarios (Knowler and Bradshaw, 2007, Knowler, 2015).

Farm household modeling can help predict how farmers will respond to technology supply (Hazell and Norton, 1986). However, many agricultural models have focused solely on farmers' attitudes toward income and risk. (Maatman, et al., 1998, Van den Berg, et al., 2007, Laborte, et al., 2009, Affholder, et al., 2010). In fact, understanding farmers' potential attitudes towards new technologies is a typical multi-criteria decision analysis (MCDA) because farmers have multiple and sometimes conflicting goals (social, economic, environment) and have to fulfil current goals without compromising future generation (Featherstone and Goodwin, 1993, Greiner, et al., 2009, Wilson, et al., 2013, Alrøe, et al., 2016, Kandakoglu, et al., 2019). Ex ante multi-criteria assessment taking into account a larger set of farmer's goals would improve our understanding of the trade-offs they are making when choosing technologies. In turn, this would help tailor the technologies to farmers' preferences and adjust policies to promote them.

A multi-attribute decision framework is needed to identify the most important criteria for farmers when choosing their cropping systems (Riesgo and Gomez-Limon, 2006, Carpani, et al., 2012, Craheix, et al., 2016, López, et al., 2018). Many empirical techniques have been developed to understand people's ranking (choice) of alternatives when they have to consider different and conflicting criteria (for a precise defininition of ranking, we refered to Colorni and Tsoukiàs, 2021). Multi-attribute utility theory (MAUT) is a commonly advocated method (Keeney and Raiffa, 1976, Velasquez and Hester, 2013, Lavik, et al., 2020) to identify farming strategies that best meet farmers' multiple goals. The MAUT method has three major steps: (i) evaluation of the performance of each alternative in terms of its utility, (ii) identification of the weights statistically representing the decision-maker's priorities for each criterion, and (iii) aggregation and ranking of the various alternatives. Once elicited, the MAUT can also easily be integrated into mathematical programming models representing farmers' selection of activities (e.g., Gómez-Limón and Riesgo, 2004).

One of the critical steps in formulating the problem is to find the adequate weights associated to each objective. However, an overview of the literature on the multi-criteria evaluation of cropping systems suggests three main gaps.

First, many existing multi-criteria evaluations of cropping systems do not take into full account the diversity of farmers' criteria and rankings. Among the ones that do, Riesgo and Gomez-Limon (2006) estimated farmer's goal weighting that better fit actual farmers' behavior using multiple goal programming. This method is appealing since it does not ask farmers directly for their rankings but infer them from their real-life choices. However, it makes strong assumptions about farmers' rationality, and it is not clear how farmers' constraints interfere with the weighting calculations: farmers in developing countries are highly constrained and it is possible that the choices reproduced by the model are the result of these constraints and not of the inferred weights. Given the difficulties to obtain weights directly from farmers, many other studies select the criteria using a review of the literature or using scientists' point of view. (e.g., Ditzler, et al., 2019, Alary, et al., 2020, Senanayake, et al., 2021).

Second, studies mainly use consensual views of stakeholder groups but do not take into account the diversity of preferences among farmers (e.g., Xavier, et al., 2020, Iocola, et al., 2021). Since ex-ante assessment of technologies can rely on farm models of representative farmers, we are looking for a typology of farmers based on the respective weights for the different goals.

Third, there is an abundant literature and many empirical techniques to infer decision-makers weighting of criteria. However, most rely on strong assumptions about the rationality of the respondents and the consistency of their choices. Consequently, their use for development studies poses a number of additional challenges. First, such studies are conducted in areas where communication is often an issue. Researchers may not master the local languages and need to rely on local translators. Moreover, some of the goals presented may be mastered differently by different farmers, e.g., because of differences in the education level. Therefore, it is necessary to be able to account for inconsistent responses.

This paper proposes to address these three gaps. We will investigate what are the farm-level goals used by farmers to manage their farms, and how do farmers rank these different goals when deciding which activity they will select. As we anticipate some diversity in farmers ranking, we also want to characterize this diversity in a way that will be amenable for further use in multi-criteria assessment. Finally, we want to investigate if some of the respondents were less consistent when making the choices presented to them.

In the reminder of the paper, we will review the methods that can address these challenges. After selecting Best–Worst Scaling (BWS) stated preference choice experiment as potentially addressing these challenges, we present the methodology and analysis applied to Lao farmers of the province of Xieng Khouang. Heterogeneity in producer ranking of these objectives was further explored with a scale-adjusted latent class logistic regression. The possibility of respondents being less assured/consistent in their choice is addressed specifically with the estimation of a “scale” factor in our analysis of the choices.

2 Theoretical background

This paper is interested in methods to elicit weights required in full aggregation approaches, with a special interest for MAUT weights. As we conducted the research with Lao farmers, it was important to evaluate how the potential mistakes or inconsistent responses would be treated in the methodology.

2.1 Weights for full aggregation approaches

Techniques to elicit weights enter into three broad categories: rating techniques, pair-comparisons techniques and choice-based techniques.

The rating techniques, popularized from the work of Likert (1932), ask the respondents to rate each object separately, for example on scales of the type "not important-very important". The technique is easy to apply (Grover, et al., 2020) but it has been criticized because (a) it does not force respondents to consider trade-offs between alternative objects and results in very positive answers for all the objects also known as the acquiescent response bias (Billiet and McClendon, 2000) (b) rankings are personal and cannot be compared across individuals, (c) scales may suffer cultural biases (Steenkamp and Ter Hofstede, 2002), and finally (d) the hypothesis of equidistance between the levels of a scale of intervals may be inadequate (Mühlbacher and Kaczynski, 2016, Koban and MacDonald Gibson, 2017). These weaknesses can be partially overcome but at the cost of complexity of the questions asked to respondents (e.g., Lodge, 1981, Marder, 1997, Bartikowski, et al., 2006). The rating technique DIRECT (Bottomley and Doyle, 2001, van Calker, et al., 2006), and SWING (Von Winterfeldt and Edwards, 1986) also rely on the direct declaration of weights. However, both of these methods do not take into account the possibility of tradeoffs, and a recent comparison of rating methods with choice-based methods showed that rating discriminate more poorly between attributes (Muhlbacher, et al., 2016, Whichello, et al., 2020). Overall, the rating techniques do not have a problem of consistency since they do require respondents to compare the criteria. However, the difficulties to compare results across individuals and to consider tradeoffs make these approaches impractical.

Among the pairwise comparisons, the analytical hierarchy process (AHP) method and its extensions has been extensively used (Saaty, 1977, Duke and Aull-Hyde, 2002, Goméz-Limon and Atance, 2004, Mühlbacher and Kaczynski, 2016, Ajibade, et al., 2021). During the comparisons, the respondents are requested to rate the level of importance of one criterion over the second one. The $(n \times (n-1)/2)$ comparisons are reported in a symmetric comparison matrix, whose eigenvalues serve as a basis for the calculations of the respondent weights. The technique can be cumbersome when the number of items to rank is large. Besides, the techniques relies on comparisons along a given scale, e.g., attribute a_1 is twice as important as a_2. However, due to limited rationality or errors, respondents can be inconsistent, e.g. a_1 is deemed twice as important as a_2, and a_2 twice as important as a_3, but a_1 is as important as a_3 (Brunelli, 2017). One possible strategy used in AHP, especially when evaluating group rankings, is to detect respondents who were too inconsistent and to discard them as “non-rational”. Another possible strategy is to detect what comparisons are the most contradictory (Ergu, et al., 2011) and guide the decision maker to obtain sufficiently consistent preferences (Pereira and Costa, 2015). In both cases, we feel it undermines the respondents' initial assessments and imposes a certain "rationality" on them. The limited capacity to take into account inconsistencies without influencing respondents response also made these approaches impractical.

The choice-based approaches, or choice experiments, do not rely on direct ratings but infer the strength of the preferences from how often respondents choose one object over other known objects, where each object is composed of a number of attributes. On the one hand, this approach has four distinct advantages. Firstly, respondents do not provide ratings; therefore there is less possibilities of different interpretation of the scale by different respondents and this increases the comparability of weightings across respondents (Louviere, et al., 2015). Secondly, the use of experimental designs allows us to ask much less than the $(n \times (n-1)/2$ comparisons. Individuals are given successive choices to be made among different combinations of objects and asked which of the options available in the different sets they would prefer (Louviere and Woodworth, 1983). Experimental designs are used to construct a manageable number of choices, usually 6 to 10 choice scenarios per respondents, and still allow for estimation of the preferences (See Street and Street, 1986 for additional information about experimental plans, Rose and Bliemer, 2009). Thirdly, weights inferred from the choices are interpreted as marginal values of the respondents random utility (McFadden, 1973, McFadden, 1974) so they are fully compatible with the MAUT approach which also reasons in terms of utility. Fourthly, since choice experiment is based upon random utility theory, it fully embraces the possibility that respondents may make mistakes (McFadden, 1973, Swait and Louviere, 1993). On the other hand, choice experiments require large (> 100) surveys and mostly provide average population (or classes) weights. However, as the ultimate use of the results will be integrated in model of representative farm household, this is not a major concern.

Given these important advantages, we used a choice based approach. However, there is a large diversity of choice-based approaches (Hensher, et al., 2015). MCDA practitioners have used two important developments.

First, when the number of attributes to consider is high, full-profile comparisons are cognitively more difficult to perform, which may lead the respondents to give inconsistent responses, or to use simplifying heuristics by rating only based on a few important attributes while disregarding the rest (Scarpa, et al., 2013, Hensher, 2014). Partial-profile DCE is often used in that case. The adaptive PAPRIKA method (Hansen and Ombler, 2008) has been widely used (for recent examples, see Jakubczyk, et al., 2021, Payini, et al., 2022). One of the reasons for its popularity is that it produces results even with small samples. While non-adaptive DCE designs will typically require a large sample of respondents to divide the necessary comparison questions, the adaptive PAPRIKA method can derive the full set of weights from just one respondent, while still keeping the choices easy to compare and number of questions manageable. To achieve this efficiency, the method utilizes the factor independence assumption and the transitive conservation of relationships (i.e., if A > B and B > C, then A > C) to implicitly solve a large proportion of the necessary discrete choice comparisons based on the participant's prior response (Hansen and Ombler, 2008). However, the adaptive nature of the method can actually inflate the impact of respondent's initial errors. This approach may not be warranted when possibility of misunderstandings by respondents is high.

Second, in order to gain more information per respondents, Finn and Louviere (1992) suggested that asking for both best and worst choices provides additional reliable information about the overall ranking of the objects. The method is often referred as the Best-Worst Scaling (BWS) approach where the terms "best" and "worst" are indicators that define the extremes of a latent subjective continuum (e.g., the most and least important, the most and least difficult, etc.). BWS is also underpinned by a theory about the processes that individuals might follow in providing best and worst data (Marley and Louviere, 2005). BWS allows obtaining information with a smaller sample, but it does not reduce the cognitive burden of the respondents as it require comparisons of all the attributes.

Among the possible forms of BWS, the "Object BWS", subsequently OBWS, corresponds to studies where the items being ranked, called objects, are not described by attributes that vary (Louviere, et al., 2015). The purpose of the OBWS is to rank each object on a latent subjective scale. The OBWS is particularly well fitted to answer our questions, as we want to obtain the relative importance of farmers' goals when choosing household activities, where the goals are defined in broad terms and not differentiated by different levels of achievement of these goals.

During an OBWS experiment, respondents are presented with several subsets of objects, and for each subset are required to make two choices: the best and the worst choices (or the most and least important goals, in our case). The number of objects per set and the number of subsets presented to respondents is designed to gather information about the relative preferences for each respondent in a consistent and efficient manner. For OBWS, the number of objects per set and the number of set presented to respondent is usually designed using a balanced incomplete block design (BIBD) (Street and Street, 1986).

OBWS data can be analyzed for either groups or individuals. For groups, one calculates the total number of times that each object of interest is chosen as best and worst across all comparison sets pooled for all the individuals of the group and derives some indicators of preferences (e.g., the best counts minus worst counts). Researchers who wanted to understand the diversity of rankings used various cluster analytic methods on the individual scores (Ochieng' and Hobbs, 2016), latent class models (Loureiro and Arcos, 2012, Louviere, et al., 2013) or random parameter models (Erdem and Rigby, 2013).

Given the additional information gain, we opted for a BWS approach.

2.2 Scale-adjusted Latent Class Models of OBWS

The latent class models are based upon a utility-based theoretical framework (Marley and Louviere, 2005). With a standard latent class model, we can assign each individuals a probability of belonging to one of the C classes of homogenous preferences. It uses two sub-models to calculate the probability that an individual will choose a specific alternative. The first sub-model estimates the probability that each individual will belong to classes, while the second sub-model estimates the class probabilities of choosing one alternative conditional on the preference parameters of each class.

The probability of observing the respondent classifying the alternative j as most important attribute forms the basis for the construction of the likelihood function of the model to be estimated:

$$Pr(y_{i,j,t} = 1) = \sum_{c=1}^{C} Pr\left(y_{i,j,t} = 1 \middle| i \in c\right) \times Pr(i \in c) \quad (1)$$

Both sub-models use a multinomial logistic formulation. The probability π_c that respondent *i* belongs to class *c* is represented by a multinomial logit:

$$\Pr\,(i \in c) = \pi_c = \frac{exp(\theta_c'.Z_i)}{\sum_{c'=1}^{C} exp\,(\theta_{c'}'.Z_i)}; \quad c = 1, \dots, C \quad (2)$$

where Zi is a vector of observable characteristics of individuals related to class membership (Greene and Hensher, 2003).

The probability that an individual i belonging to a specific class $c \in \{1, \dots, C\}$ will choose one alternative $j \in \{1, \dots, J\}$ proposed in choice situation t is written as (Greene and Hensher, 2003):

$$Pr\left(y_{i,j,t} = 1 \middle| i \in c\right) = \frac{exp\ (\lambda.\beta_c'.X_{i,j,t})}{\sum_{j'=1}^{J} exp\ (\lambda.\beta_c'.X_{i,j',t})} \quad (3)$$

where $y_{i,j,t}$ is an indicator variable that takes the value 1 when respondent i chooses the alternative j and 0 otherwise, $X_{i,j,t}$ is a vector describing the attributes of the choice situation, and β_c is a vector of utility parameters specific to class c

The response variance is inversely related to λ, known as the scale parameter, and is interpreted as a measure of the lack of certainty of respondents when making choices (Swait and Louviere, 1993). Under standard latent class approach, each class c has its own unique preference parameters which are estimated under the assumption $\lambda = 1$ for all respondents.

To account for the heterogeneity in the certainty of responses, Magidson and Vermunt (2007) proposed a scale-extended model latent class model:

$$Pr\left(y_{i,j,t} = 1 \middle| i \in c, s\right) = \frac{exp\ (\lambda_s.\beta_c'.X_{i,j,t})}{\sum_{j'=1}^{J} exp\ (\lambda_s.\beta_c'.X_{i,j',t})}$$

where c is a preference class, and s is a scale class. As an example, if we have two classes of preferences and two scale classes, we have four joint classes with associated probabilities (Table 1).

Table 1: Example of scale-extended latent class model with two preference classes and two scale classes: allowing for the estimation of four classes of weights.

	Scale classes	
Preference classes	$\lambda_1 = 1$ (for identification purpose)	λ_2
$\beta_1 = (0.1, 0.5, 0.4)$	$\beta = (0.1, 0.5, 0.4)$	$\beta = (0.1\ \lambda_2, 0.5\ \lambda_2, 0.4\ \lambda_2)$
$\beta_2 = (0.3, 0.1, 0.6)$	$\beta = (0.1, 0.5, 0.6)$	$\beta = (0.1\ \lambda_2, 0.5\ \lambda_2, 0.6\ \lambda_2)$

This formulation allows the distinction between heterogeneity in the preferences (the weights attributed to each objective) and the uncertainty about these weights.

To summarize, the analysis of OBWS will allow us to elicit farmers' weights for the different objectives using a small number of simple tasks. The OBWS data are more amenable to comparisons across respondents, as they do not require any ratings. Finally, the Scale-adjusted Latent Class Models should improve our capacity to separate classes of weights since they allow for the possibility of less consistent answers across the different tasks. The latent scale approach assign a probability of belonging to the different classes, and therefore allows a discussion of the socio-economic factors related to the identified classes of weight.

3 Designing the Best–Worst Scaling experiment

3.1 The Lao case study

Xieng Khouang (XKH) province is located in the northern part of the Lao People's Democratic Republic. During the 2000s land use changed quickly. Hybrid maize cultivation replaced traditional upland rice, gardens, orchards and also expanded on forests and fallows areas (Castella, et al., 2012). These changes are a direct consequence of the increased demand for meat products in South-east Asia and the resulting increased demand for maize for animal feed industry. Vietnamese agricultural traders introduced hybrids cultivars of maize in the region, hence favouring the rapid replacement of the slash-and-burn-based upland rice cultivation with hybrid maize crops. Mechanical ploughing and herbicide use began to be common practices linked to the increased benefits of improved hybrid maize (Castella et al. 2012). The costs of ploughing services could therefore be covered by the sale of maize; this encouraged the abandonment of rainfed rice and of the fallow system traditionally used to manage weeds in the uplands. All of these transformations appear to be intensification and a simplification of the cropping systems.

Kham district was selected as a typical example of the XKH's agricultural intensification and its consequences. Kham district is characterised by relatively fertile soils, good accessibility and a microclimate conducive to various commercial crops such as pepper, vegetable, chili, maize. The simplification of the landscape and this rapid agricultural transition generated negative environmental impacts such as soil erosion, siltation of the lowland, weeds invasion, and water contamination with herbicides. The inadequate use of inputs or ploughing service may also have serious economic impacts like household indebtedness (Jobard, 2010). Climate change will further increase farmers' challenges in the future: average temperatures are likely to rise (+2°C), rainfall is likely to be more intensive leading to more erosion and nutrients leaching, and the frequency of extreme weather events is likely to rise.

We selected 6 villages of the Kham basin (Dokham, Laeng, Le, Houat, Xay and Nadou) constrasted in terms of ecological zone, road accessibility and village size.

3.2 Elicitation of farmers goals

The first step of the research was to obtain reliable information about the main goals used by farmers. We opted for an indirect elicitation of farmers' goals based on individual interviews conducted with 20 farmers in four villages. We used serious games to reveal farmers' objectives as they can reveal more salient information than direct household interviews (Cash et al., 2003). For a full description of the game structure and methodology, the work of Lairez, et al. (2020) may be consulted. Based on these interviews, we collected 20 goals that were combined into 7 goals because they were formulated in very similar terms (Table 2).

Table 2 : Final list of objectives potentially guiding agricultural household choices

	Objective	General Description
TR	Having a transmissible farm capital	Having enough farm assets to transmit to the next generation
HI	Having high incomes punctually (even discontinuous)	Select activities that allow high income at least once in a year
RI	Having regular incomes (even small)	Select activities that provide regular income (daily, weekly, monthly)
LC	Minimize cash out	Select activities that need less "investment" or cash-out during the year.
LL	Reducing work & efforts	Select of set of activities that would reduce the need for important quantity of work, or to reduce the hardship/effort of the work
RR	Reduce risk by diversifying	Select a set of activities so that when the income of one activity is not good, I will have other activities to compensate. It could be with in/ off-farm activities."
SU	Being self-sufficient in rice	Being able to produce rice in sufficient quantities to feed the household for one year (without having to buy outside)

3.3 Experimental design

To construct the sets of goals to be evaluated by farmers, a balanced incomplete block design (BIBD) was used. A BIBD is an experimental design which produces fixed set sizes and ensures equal occurrence and co-occurrence of objects across all comparison sets (Street and Street, 1986). This helps minimizing the chance that respondents make unintended assumptions about the objects based on the aspects of the design. The design retained ensured that each of the seven scenarios contained three goals, each goal appeared three times, and co-appeared once. The final sets (Table 3) were obtained using the R package Algdesign (Wheeler, 2004).

From May to July 2017, we interviewed 120 households selected from the 6 studied villages. The selection of the households followed a procedure equivalent to the one described in section 3.2. At the beginning of each interview, the goals were presented and discussed with the farmer until they felt comfortable with the concepts. To ensure mutual understanding, each goal was illustrated by some drawing, and each drawing had a caption describing it in a few words (in Lao). Then, we opened successively, seven envelopes each containing 3 goals, according to the BIBD design. The order of the envelopes and order of the different goals presented on a table to the farmer were random to avoid possible biases linked to these orderings.

Table 3 : Scenarios presented to farmers using a Balanced Incomplete Block Design

	Goal 1	**Goal 2**	**Goal 3**
Set 1	HI	LC	RR
Set 2	RI	LC	LL
Set 3	TR	HI	RI
Set 4	TR	LL	RR
Set 5	TR	LC	SU
Set 6	HI	LL	SU
Set 7	RI	RR	SU

TR: Having a transmissible farm capital; HI: Having high incomes punctually (even discontinuous); RI: Having regular incomes (even small); LC: Minimize cash out; LL: Reducing work & efforts ; RR: Reduce risk; SU: Being self-sufficient in rice

For each set, the farmers were asked first to choose the most important goal (the “best”) guiding their farming household decisions for the coming five years (choice of activities, technical choices, investments, etc.). Once the farmer identified the most important goal, it was removed and the farmer had then to identify the least important goal (the worst). This procedure somehow forced a sequential best-worst type of choice process.

3.4 Data analysis

3.4.1 Analysis of the average ranking

To evaluate the average ranking, the choices from all the respondents are pooled into one set and the following indicators are calculated for each goal: (a) *B*: the number of times it was mentioned as the most important, (b) *W*: the number of times it was mentioned as the least important, (c) the standard score: *SS = (B-W)/(N*3)* (where *N* is the number of surveys, and 3 reflects the fact that each objective is presented three times to respondents), (d) the analytical best–worst: $ABW = Log(1 + SS)/Log(1 - SS)$ (Marley, et al., 2016), and finally (e) the ratio scores $RS = \sqrt{B/W}$ scaled by a factor such that the numerically highest *RS* takes the value 100. *ABW* is anticipated to provide better fits to the aggregate choices (Lipovetsky and Conklin, 2014; Marley et al., 2016), and *RS* to provide an indicator of the *relative* importance of the different objectives (Louviere et al., 2015).

3.4.2 Heterogeneity of rankings

The indicators described in the previous section were also calculated for each farmer, and used to evaluate the heterogeneity of rankings for each goal (univariate analysis) and the presence of homogenous groups of rankings in the population. The standard deviation of the individual BWS scores provide an indication of the extent to which choices made by farmers during the BWS experiment were consistent or whether those choices exhibited heterogeneity. To establish the extent of heterogeneity, the individual standard deviation to individual mean of B–W score is utilized (Stdev/Mean). High absolute ratios of Stdev/Mean signal greater heterogeneity, while absolute ratios that are close to zero indicate high levels of agreement with respect to the importance of the goal.

The scale-adjusted latent class model presented in the theoretical framework was further adapted to use the best as well as the worst choices. As the farmers were first asked to choose the most important objective among three possible choices, then asked to choose the least important objective among the remaining two objectives, we aggregated the two successive choice situations, the first one where the chosen objective (among three) had a positive impact on the utility, and the second one (the least important among the remaining two) where the chosen objective have a negative impact on the utility (Louviere, et al., 2015).

We tested the existence classes with different scale parameters, meaning that some respondents may be more certain about their choices than others, even if they have the same weights. Finally, we tested models where the scale parameters are potentially different when considering the "most important" and the "least important" choices. The last type of model would correspond to a situation where farmers, in general, would have a different level of certainty for the two types of choices to be made.

As described in Equation 2, we also tested for possible associations between the classes and structural variables describing the farm or the farm's current activities (Table 4). In the first case, we wanted to test whether some of the farmers' objectives could be associated with their socio-economic conditions or life-stage. Some correlations would mean that typologies of the structure of farms would be helpful to predict farmers' objectives. In the second case, we tested possible correlations between what farmers stated as important objectives and what they actually do. For example, it would be expected that farmers including the development of activities generating regular income would actually develop these activities more than farmers in the other groups.

Table 4: Covariates associated with preference classes and related hypotheses

No	Covariates	Hypotheses made
1	Age	Younger household heads are less concerned about transmissibility; and may be more concerned with labour and cash constraints
2	Family Labour (FL)	Households with large work force are less concerned with the labour constraints.
3	Family Size (FS)	More mouth to feed should increase the concern for food sufficiency
4	FL / FS	Households with large work force are less concerned with the labour constraints.
5	Cultivated area (CA)	Higher cultivated area should reduce the concern for self-sufficiency
6	Paddy Area (PA)	Household with larger paddy area less concerned with food self-sufficiency issues.
7	FL/CA	Higher family labour per cultivated area may reduce concerns for labour requirement
8	MAIZE /CA	% of maize in the cultivated area
9	MAIZE/FL	Maize area per family labour
10	MAIZE/FS	Maize area per mouth to feed
11	Cattle	No of cattles
12	Cattle / FS	No of cattle per mouth to feed
13	Tot income / FS	Income per mouth to feed
14	Weaving income	Income generated by weaving activities

The scale-adjusted latent class model of the joint best and worst data was run using Latent GOLD 5.1 (Vermunt and Magidson, 2016).

4 Results

4.1 Average rankings

The average rankings are presented in Table 5. The results indicate that farmers place a very high value on being rice self-sufficient through their farming activities alone. Although most farms on the area are now connected to markets and can rely on them for household food supply, they still want to be self-sufficient in terms of rice production. The second most important goal was the transmission of a farm capital to the next generation. As such, it suggests that farmers are concerned about the long-term sustainability of the activities they develop.

Another important goal was the development of activities that could provide a regular income to ensure a steady supply of cash to the household. However, this objective was ranked at a much lower importance on the 0-100 SRS scale. Similarly, reducing risks through diversification of activities, generating higher cash income and reducing labor and cash requirements appeared to be less important goals, on average, for decision making. These results should be interpreted with caution and do not mean that farmers do not take into account these lower ranking goals. Given the underlying choice model we used, the results should be interpreted in terms of the trade-offs that farmers are likely to make when considering new activities, e.g. an activity that would reduce the household rice self-sufficiency would have to demonstrate very significant improvements in terms of reduced labour (or cash) requirements before being considered by the farmers.

Table 5: Sample average ranking indicators of the different goals

Goals	B[a]	W[b]	SS[c]	ABW[d]	SRS[e]
SU: Self-sufficient rice production	220	44	0.49	1.07	100.00
TR: Transmittable farm capital	199	49	0.42	0.89	90.12
RI: Regular income (even small)	113	113	0.00	0.00	44.72
RR: Reduce risk by diversifying	89	139	-0.14	-0.28	35.79
HI: High income (even if irregular)	71	124	-0.15	-0.30	33.84
LL: Low Labour requirements	75	172	-0.27	-0.55	29.53
LC: Low cash requirements	73	199	-0.35	-0.73	27.09

[a] B: number of times a goal was mentioned as the most important
[b] W: number of times a goal was mentioned as the least important
[c] SS: Standard score = (B-W)/(N*3)
[d] ABW: Analytical best–worst =(Log(1+SS))/(Log(1-SS))
[e] SRS: Standard ratio scores where RS=√(B/W) is scaled

The mean and standard deviations of individual (B-W) indicators are presented in Table 6. The results suggest a relative consensus for the objectives SU and TR as very important, and LC (low cash requirement) as being less important and greater variability for the other objectives (RI, RR, HI, and LL) among different households. This suggests that different groups of preference may be present and will mainly be contrasted by the importance they attach to these four goals.

Table 6: Mean and standard deviations of individual (B-W)[a] indicators

Objectives	M mean of (B-W)[a]	S standard deviation of (B-W)[a]	S/M
Self-sufficient rice production	1.47	1.47	1.00
Transmittable farm capital	1.25	1.40	1.12
Regular income (even small)	0.00	1.67	>1000
Reduce risk by diversifying	-0.42	1.49	-3.57
High income (even if irregular)	-0.44	1.53	-3.46
Low labour requirements	-0.81	1.48	-1.84
Low cash requirements	-1.05	1.42	-1.35

[a] B is number of times a goal was mentioned as the most important, and W is number of times a goal was mentioned as the least important

4.2 Scale-adjusted latent class analysis

We ran 16 different models varying the number of preference classes from two to five, the number of scale classes (0 or 2), and the possibility of different scale factors when treating the best and the worst scenarios. Standard procedures (Log Likelihood, plausibility and significance of classes) and information criteria scores were used to select the optimal number of classes within the data. In particular, the Akaike Information Criteria 3 (AIC3) was given special attention since results of Andrews and Currim (2003) suggest that it is a better criterion than BIC and AIC for determining the number of latent classes in choice models. Based on these criteria, we selected a model with four preferences classes, two scale classes and a different scale between the best and worst choices. The results of this model are presented in Table 6.

The log-scale factor for preferences was significantly different from zero.[1] The scale classes differentiate farmers give the same weights to the different objectives but differ in the consistency of their answers. The results suggest that the probability to give less consistent answers was 30%. Lower consistency could be related the difficulty of making choices or to a misunderstanding about the objects to be compared.

[1] For model identification, the log-scale factor associated with the first preference class was set to 0.

The log-scale factor for the worst choices was significantly less than zero[2]. This suggests that the choice of the least important goal is only exp(-.291) = .74 times as consistent as the choice of the most important goal. This suggests that farmers had more difficulty in choosing the least important goal than the most important goal. On this aspect, the interviewers felt that the farmers tended to take more time to express their least important goal, but they did not record these times systematically.

In order to present a comparable ranking between the classes (e.g. on a scale of 0 to 1), and to avoid conveying the idea that a negative coefficient would mean a "negative importance" when it only means a less important goal, the literature suggests commenting on the probability:

$$P(j|c) = \frac{\exp{(\beta_{cj})}}{\sum_{j'=1}^{k} \exp{(\beta_{cj'})}}.$$

$\mathbf{P(j|c)}$ can be interpreted as the estimated choice probabilities by members of class c to choose the objective j as the best objective. They are presented in Table 7.

Table 7: Log likelihood estimates of the scale-adjusted latent class model of the joint best and worst data

Attributes	**CL1**	**s.e.**	**CL2**	**s.e.**	**CL3**	**s.e.**	**CL4**	**s.e.**
Class size	43%		20%		19%		18%	
SU	1.64	0.255	1.70	0.370	1.01	0.256	0.30	0.215
TR	1.16	0.199	2.05	0.426	0.59	0.213	0.55	0.208
RI	-0.01	0.148	-0.01	0.238	-1.68	0.400	1.56	0.333
RR	-0.29	0.162	-1.61	0.362	-0.04	0.221	0.48	0.242
HI	-0.70	0.177	-0.37	0.254	0.81	0.213	-0.76	0.284
LL	-0.39	0.166	-1.47	0.413	-0.71	0.276	-0.48	0.262
LC	-1.39	0.242	-0.29	0.235	0.01	0.229	-1.63	0.387
Scale Models		Overall	s.e.	Wald	p-value			
Worst log-scaling factor		-0.291	0.075	15.138	1.0E-04			
Preference log-scaling Factor		-1.037	0.258	16.222	5.6E-05			
Log-likelihood	-1226.44	p= 0.000						
AIC3	2650.87							

SU: Self-sufficient rice production; TR: Transmittable farm capital ; RI: Regular income (even small); RR: Reduce risk by diversifying; HI: High income (even if irregular), LL: Low Labour requirements; LC: Low cash requirements

[2] For model identification, the log-scale factor associated with the most important choices was set to 0.

Table 8: Ranking presented as probabilities

Attributes	CL1	CL2	CL3	CL4
Class size	43%	20%	19%	18%
Self-sufficient rice production	**0.38**	**0.30**	**0.26**	0.13
Transmittable farm capital	**0.25**	**0.41**	**0.19**	0.16
Regular income (even small);	0.10	**0.08**	0.04	**0.37**
Reduce Risk by Diversifying	0.08	0.03	**0.11**	0.15
High income (even if irregular)	0.06	**0.06**	**0.22**	0.06
Low Labour requirements	0.09	0.04	0.07	0.07
Low cash requirements	0.04	**0.08**	**0.12**	0.04

We can make general comments on the rankings. Firstly, the SU goal is considered the most important goal for two classes (63%) and the second most important for the class 2 (20%) where it is considered only with TR. This is consistent with the calculated average rankings. Second, the TR goal is considered the most important for class 2 (20%) and the second most important for the other classes. Again, this result is consistent with the overall high average BW score for this goal. A third set of goals is related to the regularity of the flow of cash incomes (RI, HI) and to the reduction of risks taken (RR). These goals were, on average, ranked as less important than the two previous goals, but their ranking varied across classes. Therefore, these goals are very useful to differentiate farmers. Farmers in classes 1 and 2 do not consider any of these 3 goals as important, while Class 3 and 4 clearly differentiate farmers with opposite views about RI. Finally, low labor or low investment requirements were often ranked as less important than the other goals.

The results also allow the identification of contrasted preference classes. The parameters of farm and household characteristics influencing the probabilities of the classes are presented in Table 8 to facilitate their interpretation. This table does not include the variables PA/CA (paddy areas expressed as a percentage of cultivated areas), and PA/FS (paddy area per mouth to feed) as their coefficients were not significantly different from zero in the probability model of the latent class, and were removed from the analysis. For these two variables, we conducted an ANOVA to compare their means across classes, where each farmer was assigned to a class, i.e. the one with the highest unconditional probability. Only the variable PA/CA had different means across classes (Table 10) but at the 90% probability level. We also compared the proportion of farmers using fertilizers or herbicides on their paddy and maize fields across classes to assess the possible relationships between goals and production intensification (Table 11). We compared the proportion of farmers with grazing areas across classes to assess the relationship between goals and the choice of raising grazing animals (Table 11). Finally, we also assessed the relationship between goals and the choice of having significant weaving activities (Table 11). For this last criterion, we classified farm households as having significant activities when that activity represented more than 10% of the household income.

For the first class, the two most important goals are SU and TR. The other goals with a lower relative ranking are RI, LL, and RR. Farms of this group are associated with lower percentage of paddy and higher maize areas. With smaller paddy areas (mainly a result of their constrained access to land and water), these farmers ranked rice production as an important goal. These farmers could increase their rice production by using more external inputs on their paddy fields. Farmers in this group have relatively less family labor per cultivated area and they also considered the reduction of labor requirement among the secondary goals. Table 11 suggests that a higher proportion in farmers of this class use herbicides compared to all other classes (related to the labor goal) and fertilizers compared to Class 2 & 3 (related to the rice sufficiency goal). In addition, many farmers in this class have a high proportion of their income coming from weaving activities. This class is representative of the farmers who have made a transition from upland rainfed rice to maize, which has allowed them to increase their income and intensify their paddy rice through less labor-intensive maize crop.

The second class listed TR and SU as its most important goals. Other goals, although considered much less important, are probably also taken into account, especially RI and LC. These farmers are older, have larger families and workforces than the other groups, and devote a larger percentage of their cultivated area to paddy cultivation. They keep more cattle than groups 1 & 4. This class is representative of farmers who have fewer problems with rice production (partly because of their larger share of paddy fields), but who do not rely on maize production (or are gradually reducing it) and who replace maize with grassland to raise livestock (Table 9). A higher proportion of farmers in this group use herbicides on maize[3] even if they have more workforce available for manual weeding. They also did not indicate the reduction of labor was an important objective, which could be interpreted as a willingness to prioritize the work in rice fields (SU important) at the time of the peak of work at the beginning of agricultural season rather than for manual weeding of maize fields.

Farmers in class 3 farmers can be described as "constrained optimizers". They consider a larger number of distinct goals, with a better balance between the different objectives. SU is

[3] One should note however, that all the groups have low proportion of herbicides users.

considered an important goal. However, they also consider the production of high income generating activities as their second most important goal. TR / RR/ LC are also likely to be considered. They have less paddy and maize land than the other groups, but a higher number of cattle than class 1. In short, these farmers are trying to find activities that produce high incomes while keeping in mind their constraints (rice sufficiency, cash constraints, risk).

Finally, farmers in the fourth class could be described as “risk-averse farmers”. Their main goals are the production of regular incomes (RI) and the diversification of activities (RR). Other important goals are the transmission of farm assets and the reduction of labor requirements. Although SU is not the highest priority, a higher percentage of their cultivated area is devoted to rice. However, the land they cultivate are smaller and use of fertilizers which seem to indicate they have solved the issue of rice production for the household. This class comprises younger farmers; they keep more cattle than the classes 1 and 3, and also have weaving revenues (although to a lesser extent than class 1); these activities are compatible with their main goals (RI, and probably TR and LL).

Table 9: Coefficients of the farm and farm household information in the Class Probability model

Covariates	Class1	s.e.	Class2	s.e.	Class3	s.e.	Class4	s.e.
Age	-0.16	0.07	0.29	0.12	-0.01	0.06	-0.12	0.06
Family labour (FL)	-0.30	2.00	3.60	2.38	**-5.15**	**1.65**	1.86	2.50
Family size (FS)	-2.17	1.35	0.99	1.34	2.84	1.37	-1.66	1.42
Cultivated area (CA)	3.35	2.15	-0.03	1.73	0.46	1.78	-3.78	1.59
FL/CA	-4.44	5.55	-0.18	4.87	**11.86**	**4.45**	-7.24	7.48
FS/CA	4.02	2.41	3.03	2.51	**-9.31**	**3.42**	2.26	3.15
Maize Area /CA	4.28	8.59	4.07	9.37	14.67	9.94	-23.01	8.52
Maize Area /FS	18.58	14.76	-23.28	14.91	-12.22	13.47	16.92	10.25
Cattles	0.14	0.79	-1.45	0.64	-0.22	0.42	1.53	0.54
Cattles/FS	-9.32	4.06	12.87	4.25	2.33	2.37	-5.87	2.88
Income/FS	0.06	0.17	0.11	0.19	-0.54	0.22	0.37	0.16
Weaving income	0.61	0.23	0.01	0.14	-0.55	0.26	-0.06	0.17
Note: When using Latent Gold, the sum of the coefficients for one variable across classes is set to zero to obtain model identification								

Table 10: Comparison of means of the variable Paddy area / Cultivated Area across classes using ANOVA

Class	1	2	3	4
Mean†	0.29^a	0.42^b	0.31^a	0.38^b
St. Error	0.16	0.25	0.16	0.13
F value= 3.412 Pr(>F) = 0.0199				

† Letters indicate statistically different groups by HSD Tukey at 90%

Table 11: Test of association between farm activities and classes

	Class 1	Class 2	Class 3	Class 4	p value†
Farmer used herbicides in paddy fields	0.25	0.05	0.1	0.14	0.147
Farmer used herbicides in maize fields	0.25	0.65	0.68	0.5	0.001
Farmer used chemical fertilizers in paddy fields	0.74	0.59	0.62	0.86	0.18
Farmer used chemical fertilizers in maize fields	0.58	0.56	0.45	0.53	0.811
Farm has pasture fields	0.25	0.65	0.68	0.5	0.001
Farm household has significant weaving activities††	0.19	0.04	0	0.05	0.042

† Fisher's exact test for count data
†† Significant weaving activities when they represent more than 10% of the household income

5 Discussion and conclusion

The analysis presented in this paper is a first step in understanding farmers' goals and their relationship to current activities. We developed a methodology to identify and rank the goals farmers pursue when taking decisions based on a BWS experiment. The main strength of BWS is that farmers were not asked to rate different goals directly, but to identify the most and the least important goals for carefully designed sub-sets of goals. The methods ensure that rankings are comparable across respondents and allow the classification of farmers according to their goals ranking. The methodology was simple to establish and could provide rapidly some first insights about the diversity of farming styles useful for researchers and extension workers.

The analysis involved a random sample of farmers of the Xieng Khouang province in Laos. The results suggest farmers' are very concerned about the sustainability of their farms. This could be associated with recent integration to markets that may not deliver on all its promises or with the environmental issues emerging from not poorly-mastered intensive cropping systems (Lairez, et al., 2020). We found that self-sufficiency in rice and transmission of productive assets to the next generation are the most important and consensual goals in the sampled population. The importance given to risk avoidance and cash flows was more disparate in the sample and largely defined the differences found between the latent classes. Finally, and contrary to our expectations, most farmers placed relatively low importance on the goals of reducing labor and cash requirements. The methods also allowed to identify groups of farmers with homogenous “goal profile” that could prove useful when tailoring innovations to different farming styles.

The analyses provided an overview of the different rankings of farm management goals in the study area. The results also provide interesting insights for the development of sustainable farm technologies or practices in the region. First, we noted that rice self-sufficiency was ranked highly by many farmers (especially for Class 1, 2 and 3 farmers). Any changes proposed to these farmers should take into account the impact on household rice production, the main staple food in the study area. Proposing alternative activities or strategies that would reduce rice production is likely to be ineffective, unless large gains in other goals are expected. In contrast, projects such as terracing sloping area to increase the rice production are more likely to attract their interest. This is consistent with other studies in mountainous areas of Southeast Asia (e.g., Jourdain, et al., 2014). We also noted that the transmission of farm capital was often ranked highly. We associated this goal with farmers’ concern for the sustainability of their activities.

While a large share of farmers considered these two goals to be very important, we found a diversity of rankings for the other goals. This is the case for the goals related to the attitude towards risks. The presence of different attitudes towards risk will be important in terms of technologies proposed to farmers of the region. While the presence of risk-averse farmers was not surprising, we also found that a fairly significant proportion of the sample did not consider risk avoidance to be their primary objective. This is important, as risk is often seen as slowing the development of new technologies especially when little is known about their effects (Feder, 1980, Greiner, et al., 2009). Farmers who are less concerned about the possibility of losses may be easier to convince to test and co-develop more sustainable practices in the early stages of development.

Finally, most farmers placed relatively low importance on the goals of reducing labor and cash requirements. This result is unexpected as other studies suggest that farmers in mountainous areas of Southeast Asia face high labor or cash constraints (e.g., Affholder, et al., 2010, Alexander, et al., 2018). However, it should probably be interpreted with caution, as it does not mean these goals are not considered at all by farmers, but only that they would be given less importance. Additional surveys with farmers are needed to better evaluate their access to labor and credit markets, as this would influence their attitudes towards these two constraints. For example, the relatively low importance of labor constraints may also be related to the relative difficulty for farmers to find off-farm opportunities in the region, making labor relatively available compared to the regions considered in the studies cited above.

In terms of methodology, we have shown that some respondents were less consistent in ranking the different goals. This indicates that some of the farmers had more difficulties to respond than others. The modelling assumptions link this inconsistency to the scale parameter of the utility function. A higher level of inconsistencies is usually associated with lower understanding of the tasks at hand. These difficulties could stem from the respondents (e.g., low education levels). Although not reported here, we did not find an association of this scale parameter with socio-demographic variables such as education level. The fact that surveys were conducted using translators (English to Lao to English) might explain partly these difficulties. Despite all the precautions made to describe the concepts (prior discussion and agreement on the concepts used, use of pictograms, etc.), some ambiguities could persist with at least some of the farmers. Difficulties in interacting with respondents are present in most development studies as researchers often interact with people of different languages and cultural backgrounds. However, the problem is seldom explicitly recognized and evaluated. Therefore, we consider it an advantage that our methodology allows us to assess the uncertainty of respondents' choices.

Finally, the methodology is entirely based on the farmers' stated reactions to a list of goals. However, relying solely on stated preferences can lead to a number of biases known to practitioners of choice experiments: social desirability bias, hypothetical bias, etc. A "cheap-talk" script was discussed before the choice questions to encourage truthful answers. The results suggest that the class rankings are reliable as we found a number of logical links between the goal rankings and the activities. However, further research could also compare this ranking method based on "stated" comparisons of goals with other methods that infer these weights from observed decisions and farm modeling (Sumpsi, et al., 1997, Riesgo and Gomez-Limon, 2006, Silva, et al., 2015). Other stated-preference methodologies such as Q-methodology (Alexander, et al., 2018) could also be used to compare these results.

6 References

Affholder, F., Jourdain, D., Quang, D. D., Tuong, T. P., Morize, M., & Ricome, A. (2010). Constraints to farmers' adoption of direct-seeding mulch-based cropping systems: A farm scale modeling approach applied to the mountainous slopes of Vietnam. *Agricultural Systems*, 103 (1), 51-62.https://doi.org/10.1016/j.agsy.2009.09.001

Ajibade, F. O., Ajibade, T. F., Idowu, T. E., Nwogwu, N. A., Adelodun, B., Lasisi, K. H., Opafola, O. T., Ajala, O. A., Fadugba, O. G., & Adewumi, J. R. (2021). Flood-prone area mapping using GIS-based analytical hierarchy frameworks for Ibadan city, Nigeria. *Journal of Multi-criteria Decision Analysis*, 28 (5-6), 283-295.https://doi.org/10.1002/mcda.1759

Alary, V., Messad, S., Aboul-Naga, A., Osman, M. A., Abdelsabour, T. H., Salah, A.-A. E., & Juanes, X. (2020). Multi-criteria assessment of the sustainability of farming systems in the reclaimed desert lands of Egypt. *Agricultural Systems*, 183 (102863
Alexander, K. S., Parry, L., Thammavong, P., Sacklokham, S., Pasouvang, S., Connell, J. G., Jovanovic, T., Moglia, M., Larson, S., & Case, P. (2018). Rice farming systems in Southern Lao PDR: Interpreting farmers' agricultural production decisions using Q methodology. *Agricultural Systems*, 160 (1-10
Alrøe, H. F., Moller, H., Læssøe, J., & Noe, E. (2016). Opportunities and challenges for multicriteria assessment of food system sustainability. *Ecology and Society*, 21 (1)
Andrews, R. L., & Currim, I. S. (2003). A Comparison of Segment Retention Criteria for Finite Mixture Logit Models. *Journal of Marketing Research*, 40 (2), 235-243.10.1509/jmkr.40.2.235.19225
Bartikowski, B., Chandon, J.-L., & Gierl, H. (2006). Calibration internationale des échelles sémantiques. *Décisions Marketing*Jul-Dec), 207-219
Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural equation modeling*, 7 (4), 608-628
Bottomley, P. A., & Doyle, J. R. (2001). A comparison of three weight elicitation methods: good, better, and best. *Omega*, 29 (6), 553-560
Brunelli, M. (2017). Studying a set of properties of inconsistency indices for pairwise comparisons. *Annals of Operations Research*, 248 (1), 143-161.https://doi.org/10.1007/s10479-016-2166-8
Carpani, M., Bergez, J.-E., & Monod, H. (2012). Sensitivity analysis of a hierarchical qualitative model for sustainability assessment of cropping systems. *Environmental Modelling & Software*, 27 (15-22
Castella, J.-C., Jobard, E., Lestrelin, G., Nanthavong, K., & Lienhard, P. (2012). *Maize expansion in Xieng Khouang province, Laos: what prospects for conservation agriculture?* Paper presented at the The 3rd International Conference on Conservation Agriculture in Southeast Asia, Hanoi, 10-15/12/2012
Colorni, A., & Tsoukiàs, A. (2021). Rating or sorting: Terminology matters. *Journal of Multi-criteria Decision Analysis*, 28 (3-4), 131-133.https://doi.org/10.1002/mcda.1733
Craheix, D., Angevin, F., Doré, T., & De Tourdonnet, S. (2016). Using a multicriteria assessment model to evaluate the sustainability of conservation agriculture at the cropping system level in France. *European Journal of Agronomy*, 76 (75-86
Ditzler, L., Komarek, A. M., Chiang, T.-W., Alvarez, S., Chatterjee, S. A., Timler, C., Raneri, J. E., Carmona, N. E., Kennedy, G., & Groot, J. C. (2019). A model to examine farm household trade-offs and synergies with an application to smallholders in Vietnam. *Agricultural Systems*, 173 (49-63
Duke, J. M., & Aull-Hyde, R. (2002). Identifying public preferences for land preservation using the analytic hierarchy process. *Ecological Economics*, 42 (1–2), 131-145.http://dx.doi.org/10.1016/S0921-8009(02)00053-8
Erdem, S., & Rigby, D. (2013). Investigating heterogeneity in the characterization of risks using best worst scaling. *Risk Analysis*, 33 (9), 1728-1748.10.1111/risa.12012
Ergu, D., Kou, G., Peng, Y., & Shi, Y. (2011). A simple method to improve the consistency ratio of the pair-wise comparison matrix in ANP. *European Journal of Operational Research*, 213 (1), 246-259
FAO (2010). Climate Smart Agriculture - Policies, practices and financing food security, adaptation and mitigation, available.
Featherstone, A. M., & Goodwin, B. K. (1993). Factors influencing a farmer's decision to invest in long-term conservation improvements. *Land Economics*, 69 (1), 67-81
Feder, G. (1980). Farm size, risk aversion and the adoption of new technology under uncertainty. *Oxford Economic Papers*, 32 (1980), 263-283
Finn, A., & Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy & Marketing*, 11 (2), 12-25
Frison, E. A., Cherfas, J., & Hodgkin, T. (2011). Agricultural biodiversity is essential for a sustainable improvement in food and nutrition security. *Sustainability*, 3 (1), 238-253.10.3390/su3010238
Garnett, T., Appleby, M. C., Balmford, A., Bateman, I. J., Benton, T. G., Bloomer, P., Burlingame, B., Dawkins, M., Dolan, L., Fraser, D., Herrero, M., Hoffmann, I., Smith, P., Thornton, P. K., Toulmin, C., Vermeulen, S. J., & Godfray, H. C. J. (2013). Sustainable Intensification in Agriculture: Premises and Policies. *Science*, 341 (6141), 33-34.10.1126/science.1234485
Gaudin, A. C. M., Tolhurst, T. N., Ker, A. P., Janovicek, K., Tortora, C., Martin, R. C., & Deen, W. (2015). Increasing crop diversity mitigates weather variations and improves yield stability. *PLOS ONE*, 10 (2).10.1371/journal.pone.0113261

Goméz-Limon, J. A., & Atance, I. (2004). Identification of public objectives related to agricultural sector support. *Journal of Policy Modeling*, 26 (8-9), 1045-1071

Gómez-Limón, J. A., & Riesgo, L. (2004). Irrigation water pricing: differential impacts on irrigated farms. *Agricultural Economics*, 31 (1), 47-66

Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37 (8), 681-698

Greiner, R., Patterson, L., & Miller, O. (2009). Motivations, risk perceptions and adoption of conservation practices by farmers. *Agricultural Systems*, 99 (2), 86-104

Grover, A. K., Chopra, S., & Krejci, C. C. (2020). A multi-criteria decision analysis for the public distribution system of food grains in Indian Punjab: Towards decentralized food policies in developing countries. *Journal of Multi-criteria Decision Analysis*, 27 (3-4), 236-251.https://doi.org/10.1002/mcda.1698

Hansen, P., & Ombler, F. (2008). A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *Journal of Multi-criteria Decision Analysis*, 15 (3-4), 87-107.https://doi.org/10.1002/mcda.428

Hazell, P. B. R., & Norton, R. D. (1986). *Mathematical programming for economic analysis in agriculture*. New York: Macmillan Publishing Company.

Hensher, D. A. (2014). Process heuristics in choice analysis: An editorial. *Journal of Choice Modelling*, 11 (June 2014), 1-3

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied choice analysis*. Cambridge, U.K.: Cambridge University Press.

Iocola, I., Colombo, L., Guccione, G. D., De Vita, P., Palumbo, M., Ritunnano, V., Sciacca, F., Virzì, N., & Canali, S. (2021). A multi-criteria qualitative tool for the sustainability assessment of organic durum wheat-based farming systems designed through a participative process. *Italian Journal of Agronomy*, 16 (1)

Jakubczyk, M., Niewada, M., Plisko, R., Władysiuk, M., Jachimowicz, M., Pruszko, C., Wepsięć, K., & Lis, J. (2021). What matters in treating non-oncological rare diseases?—Eliciting experts' preferences in Poland with PAPRIKA. *Journal of Multi-criteria Decision Analysis*.https://doi.org/10.1002/mcda.1754

Jobard, E. (2010). Assessing the socio-economic impacts of conservation agriculture adoption in Xieng Khouang province, Lao PDR. Vientiane: IRD-AgroParisTech, MSc Thesis, pp. 36.

Jourdain, D., Boere, E., van den Berg, M., Dang, Q. D., Cu, T. P., Affholder, F., & Pandey, S. (2014). Water for forests to restore environmental services and alleviate poverty in Vietnam: A farm modeling approach to analyze alternative PES programs. *Land Use Policy*, 41 (2014), 423-437.http://dx.doi.org/10.1016/j.landusepol.2014.06.024

Kandakoglu, A., Frini, A., & Ben Amor, S. (2019). Multicriteria decision making for sustainable development: A systematic review. *Journal of Multi-Criteria Decision Analysis*, 26 (5-6), 202-251. https://doi.org/10.1002/mcda.1682

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: preferences and value trade-offs*. New York, NY, USA: John Wiley & Sons.

Knowler, D. (2015). Farmer Adoption of Conservation Agriculture: A Review and Update, in M. Farooq and K. H. M. Siddique eds., *Conservation Agriculture*. (pp. 621-642). Heidelberg: Springer.

Knowler, D., & Bradshaw, B. (2007). Farmers' adoption of conservation agriculture: A review and synthesis of recent research. *Food Policy*, 32 (1), 25-48

Koban, L. A., & MacDonald Gibson, J. (2017). Small-unit water purifiers for remote military outposts: A new application of multicriteria decision analysis. *Journal of Multi-Criteria Decision Analysis*, 24 (3-4), 146-161.https://doi.org/10.1002/mcda.1606

Laborte, A. G., Schipper, R. A., van Ittersum, M. K., van den Berg, M. M., van Keulen, H., Prins A.G., & Hossain, M. (2009). Farmers' welfare, food production and the environment: a model-based assessment of the effects of new technologies in the northern Philippines. *NJAS*, 54 (4), 345-372

Lairez, J., Lopez-Ridaura, S., Jourdain, D., Falconnier, G. N., Lienhard, P., Striffler, B., Syfongxay, C., & Affholder, F. (2020). Context matters: Agronomic field monitoring and participatory research to identify criteria of farming system sustainability in South-East Asia. *Agricultural Systems*, 182 (102830.10.1016/j.agsy.2020.102830

Lavik, M. S., Hardaker, J. B., Lien, G., & Berge, T. W. (2020). A multi-attribute decision analysis of pest management strategies for Norwegian crop farmers. *Agricultural Systems*, 178 (102741.10.1016/j.agsy.2019.102741
Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*, 140 (5-55
Lodge, M. (1981). Magnitude scaling: quantitative measurement of opinions., available.
López, J. C. L., Carrillo, P. A. Á., & Valenzuela, O. A. (2018). A multicriteria group decision model for ranking technology packages in agriculture, *Soft computing for sustainability science*. (pp. 137-161): Springer.
Loureiro, M. L., & Arcos, F. D. (2012). Applying best–worst scaling in a stated preference analysis of forest management programs. *Journal of Forest Economics*, 18 (4), 381-394
Louviere, J., Lings, I., Islam, T., Gudergan, S., & Flynn, T. (2013). An introduction to the application of (case 1) best-worst scaling in marketing research. *International Journal of Research in Marketing*, 30 (3), 292-303.DOI 10.1016/j.ijresmar.2012.10.002
Louviere, J. J., Flynn, T. N., & Marley, A. (2015). *Best-worst scaling: Theory, methods and applications*: Cambridge University Press.
Maatman, A., Schweigman, C., Ruijs, A., & van der Vlerk, M. H. (1998). Modeling farmers' response to uncertain rainfall in Burkina Faso: a stochastic programming approach. *Operations Research*, 50 (3), 399-414
Magidson, J., & Vermunt, J. K. (2007). *Removing the scale factor confound in multinomial logit choice models to obtain better estimates of preference.* Paper presented at the Sawtooth Software Conference, Santa Rosa, California, USA, October 17-19, 2007
Marder, E. (1997). *The laws of choice : Predicting customer behavior*. New York: Free Press.
Marley, A., Islam, T., & Hawkins, G. (2016). A formal and empirical comparison of two score measures for best–worst scaling. *Journal of Choice Modelling*, 21 (December 2016), 15-24
Marley, A. A., & Louviere, J. J. (2005). Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology*, 49 (6), 464-480
McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior, in P. Zarembka ed., *Frontiers in Econometrics*. (pp. 105-142). New York: Academic Press.
McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, 3 (4), 303-328
Mühlbacher, A. C., & Kaczynski, A. (2016). The Expert Perspective in Treatment of Functional Gastrointestinal Conditions: A Multi-Criteria Decision Analysis Using AHP and BWS. *Journal of Multi-criteria Decision Analysis*, 23 (3-4), 112-125.https://doi.org/10.1002/mcda.1562
Muhlbacher, A. C., Zweifel, P., Kaczynski, A., & Johnson, F. R. (2016). Experimental measurement of preferences in health care using best-worst scaling (BWS): theoretical and statistical issues. *Health Econ Rev*, 6 (1), 5.10.1186/s13561-015-0077-z
Ochieng', B. J., & Hobbs, J. E. (2016). Incentives for cattle producers to adopt an E. Coli vaccine: An application of best–worst scaling. *Food Policy*, 59 (2016), 78-87.10.1016/j.foodpol.2015.12.004
Payini, V., Mallya, J., & Piramanayagam, S. (2022). Indian women consumers' wine choice: a study based on conjoint analysis. *International Journal of Wine Business Research*.10.1108/ijwbr-05-2021-0031
Pereira, V., & Costa, H. G. (2015). Nonlinear programming applied to the reduction of inconsistency in the AHP method. *Annals of Operations Research*, 229 (1), 635-655
Riesgo, L., & Gomez-Limon, J. A. (2006). Multi-criteria policy scenario analysis for public regulation of irrigated agriculture. *Agricultural Systems*, 91 (1), 1-28
Rose, J., & Bliemer, M. (2009). *NGENE User Manual & Reference guide*. Sydney, Australia: Choice Metrics.
Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15 (3), 234-281.https://doi.org/10.1016/0022-2496(77)90033-5
Scarpa, R., Zanoli, R., Bruschi, V., & Naspetti, S. (2013). Inferred and Stated Attribute Non-attendance in Food Choice Experiments. *American Journal of Agricultural Economics*, 95 (1), 165-180.10.1093/ajae/aas073
Senanayake, S., Pradhan, B., Huete, A., & Brennan, J. (2021). Proposing an ecologically viable and economically sound farming system using a matrix-based geo-informatics approach. *Sci Total Environ*, 794 (148788.10.1016/j.scitotenv.2021.148788

Silva, E., Marta-Costa, A. A., & Berbel, J. (2015). The Objectives and Priorities for the Azorean Dairy Farmers' Decisions, *The Agricultural Economics of the 21st Century*. (pp. 137-156): Springer.
Steenkamp, J.-B. E., & Ter Hofstede, F. (2002). International market segmentation: issues and perspectives. *International Journal of Research in Marketing*, 19 (3), 185-213
Street, A. P., & Street, D. J. (1986). *Combinatorics of experimental design*: Oxford University Press, Inc.
Sumpsi, J., Amador, F., & Romero, C. (1997). On farmers' objectives: A multi-criteria approach. *European Journal of Operational Research*, 96 (1), 64-71.https://doi.org/10.1016/0377-2217(95)00338-X
Swait, J., & Louviere, J. (1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research*, 30 (3), 305-314.10.2307/3172883
Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108 (50), 20260-20264.10.1073/pnas.1116437108
van Calker, K. J., Berentsen, P. B. M., Romero, C., Giesen, G. W. J., & Huirne, R. B. M. (2006). Development and application of a multi-attribute sustainability function for Dutch dairy farming systems. *Ecological Economics*, 57 (4), 640-658.10.1016/j.ecolecon.2005.05.016
Van den Berg, M. M., Hengsdijk, H., Wolf, J., Van Ittersum, M. K., Guanghuo, W., & Roetter, R. P. (2007). The impact of increasing farm size and mechanisation on rural income and rice production in Zhejiang province, China. *Agricultural Systems*, 94 (841-850
Velasquez, M., & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International journal of operations research*, 10 (2), 56-66
Vermunt, J. K., & Magidson, J. (2016). Upgrade manual for Latent GOLD 5.1, available.
Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*: Cambridge University Press.
Wheeler, B. (2004). The AlgDesign Package.
Whichello, C. L., Smith, I. P., Veldwijk, J., de Wit, G. A., Rutten-van Molken, M. P., & de Bekker-Grob, E. W. (2020). Discrete choice experiment versus swing-weighting: A head-to-head comparison. *Patient Preferences throughout the Medical Product Lifecycle*, 102
Wilson, P., Harper, N., & Darling, R. (2013). Explaining variation in farm and farm business performance in respect to farmer behavioural segmentation analysis: Implications for land use policies. *Land Use Policy*, 30 (1), 147-156.https://doi.org/10.1016/j.landusepol.2012.03.006
Xavier, J. H. V., Gomes, M. C., Anjos, F. S. d., Scopel, E., Silva, F. A. M. d., & Corbeels, M. (2020). Participatory multicriteria assessment of maize cropping systems in the context of family farmers in the Brazilian Cerrado. *International Journal of Agricultural Sustainability*, 18 (5), 410-426