The Systems Engineering Tool Box

Dr Stuart Burge

"Give us the tools and we will finish the job"

Winston Churchill

Pugh Matrix (PM)

What is it and what does it do?

The Pugh Matrix (PM) is a type of Matrix Diagram [1] that allows for the comparison of a number of design candidates leading ultimately to which best meets a set of criteria. It also permits a degree of qualitative optimisation of the alternative concepts through the generation of hybrid candidates.

The Pugh Matrix is easy to use and relies upon a series of pairwise comparisons between design candidates against a number of criteria or requirements. One of its key advantages over other decision-making tools such as the Decision Matrix is its ability to handle a large number of decision criteria.

Why do it?

Many decisions often concern a number of interwoven factors or criteria for which humans struggle to handle the complexity resulting in inconsistent and irrational decisions. The Pugh Matrix provides a simple approach to taking these multiple factors into account when reaching a decision. By exploiting people's innate ability to make a pairwise comparison allows for subjective opinions about one alternative versus another to be made more objective. The Pugh Matrix also allows for simple sensitivity analysis to be performed, thereby providing some information as to the robustness of a particular decision.

Where and when to use it?

Fundamentally a Pugh Matrix can be used whenever there is the need to decide amongst a number of alternatives. Although specifically developed by Stuart Pugh[1] to help in selecting between a number of design

---

[1] Stuart Pugh was Professor of Design at Strathclyde University in Glasgow. He wrote the very influential book "Total Design" [2] is which he introduced the decision matrix describe here.

alternatives, the tools has in recent years be used a general purpose decision making aid because of its ease of use.

Who does it?

An individual or team can use a Pugh Matrix. It is important to emphasise, however, that the quality of the outcome is dependent upon the experience of team or individual.

How to do it?

Overview

The basic concept of a Pugh Matrix is both simple and elegant. Figure 1 shows a completed Pugh Matrix that has been used to evaluate and select from a number of design alternatives.

	Design Concept A	Design Concept B	Design Concept C	Design Concept D	Design Concept BC	Design Concept BD
Criteria 1	S	+	S	+	+	+
Criteria 2	S	-	S	+	S	+
Criteria 3	S	S	S	+	S	+
Criteria 4	S	-	+	+	+	+
Criteria 5	S	-	+	+	+	+
Criteria 6	S	-	S	-	S	-
Criteria 7	S	+	S	-	+	+
Criteria 8	S	+	S	-	+	+
Criteria 9	S	-	S	-	S	-
Criteria 10	S	S	-	S	S	S
TOTAL +	0	3	2	5	5	7
TOTAL -	0	5	1	4	0	2
TOTAL SCORE	0	-2	1	1	5	5

Figure 1: Example of a completed Pugh Matrix

Figure 1 shows a completed Pugh Matrix for four candidate design concepts called A, B, C and D which can be found along the top of the matrix. These concepts have been evaluated against 10 criteria. In constructing a Pugh Matrix, one design concept, in this example "Design

Concept A" is selected as the "baseline". This baseline is score as "S" against all of the criteria. The other candidate design concepts are then compared in a pairwise fashion against Design Concept A for each of the criteria. If a candidate design concept is:

- better than the baseline a "+" is entered in the appropriate cell
- worse than the baseline a "-" is entered in the appropriate cell
- the same than the baseline a "S" is entered in the appropriate cell

Hence, in figure 1:

- Design Concept B is better than Design Concept A baseline for criteria 1
- Design Concept B is worse than Design Concept A baseline for criteria 2
- Design Concept B is the same as Design Concept A baseline for criteria 3

The overall evaluation is made by adding the "+" and "-" for each design concept.

The Pugh Matrix can also be used to perform qualitative optimisation by combining the candidate concept designs to form hybrid candidates. Figure 1 shows two such hybrids "Concept BC" and "Concept BD".

The Pugh Process

The process for constructing a Pugh Matrix comprises five steps. This assumes that alternative candidate design options (or decision options) have been determined.

Step 1: identify and clearly define the criteria for selection. Typically when using a Pugh Matrix to select between a number of candidate design options the design requirements can be used either in part or in whole. Ideally the design requirements should reflect both the user-customer as well as other key stakeholders including internal stakeholders. The robustness and validity of the outcome is fundamentally dependent on an appropriate set of criteria/requirements. Rushing this step usually results in a non-robust outcome that is challenged and overturned.

Step 2: Use one candidate design option as the baseline and core all criteria/requirements as 'S' (some people prefer to use an O) for this baseline. If appropriate, a good choice is to, use the previous design for the baseline because it exists and therefore its performance should be reasonably well known.

Step 3: Compare each candidate design option against the baseline design, criteria by criteria (or requirement by requirement) and decide a "pair-wise score with:

S = same
+ = better
- = worse

It is also possible to add extra levels of discrimination by using:

++ = **much** better
-- = **much** worse

Some people use a 1 to 5 scale where the baseline/same is a 3 with 1 and 2 being much worse and worse respectively, and 4 and 5 being better and much better respectively

Step 4: For each candidate design option the total score can be calculated by summing the number of +'s and –'s. The highest ranked score is the "winner" but use common sense - DON'T just select "highest" ranked concept.

Step 5: Having scored each candidate design option consider hybrids by combining where possible the best from each alternative. This is form of qualitative optimisation.

Step 6: Make the decision and record reasons behind decisions. Quite often with a Pugh Matrix there is no clear "winner" but there is often a clear "loser" in such cases perform a sanity check (does the decision make sense) and remove the losing option. At this point the criteria/requirements can be weighted to give better differentiation. Typically the weighting is on a 1 to 5 scale with 1 the lowest and 5 the highest weighting. If there is still no clear winner, the matrix is basically saying that there is not enough information to discriminate between the options. In such cases it will be necessary to:

- refine the criteria/requirements
- use an alternative selection approach
- perform more work to gain the information to be able to select between the options

It is also recommended at this point to undertake a sensitivity analysis. This can be performed in many ways that include:

- flexing the importance numbers by ± 1 and monitoring the ranking of the candidates
- removing criteria/requirements from the assessment and monitoring the ranking of candidates

The following simple and familiar example illustrates the application of a Pugh Matric to decide amongst a number of options. The options are:

4-slot Electric Toaster

Toasting Fork and Fire

Electric Conveyor

Gas Grill

In this illustrative example the scenario is of a user attempting to select amongst the options for domestic use.

The first step is to determine a suitable set of criteria. In this example the criteria will be solely user oriented. The following were identified:

Primary	Secondary
Good Toast Quality	Even Toasting
	Good Taste
	Repeatable
	Quick
Have Capacity	Large Range of Bread Products
	Multiple Slices/Units
Long Life	Reliable
	Durable
	Low
	Maintenance
	Affordable
	Attractive
	Safe
	Good Size
Easy to Use	Easy to use Controls
	Easy to Load
	Easy to Remove Toast

The use of primary and secondary criteria is often useful; particularly if there are a large number of criteria. It may be possible in such situations initially to use just the primary criteria in order to rapidly de-select the weaker options. The remaining options can then be re-evaluated using the full list of primary and secondary criteria.

It should also be remembered that as the evaluation proceeds to introduce new criteria as the understanding of the selection problem grows.

Step 2 selects one option as the baseline which scores an "S" for all criteria as shown in Figure 2. Some users prefer a "0" to an S and some even use "3" to score the baseline (score of 1 and 2 are therefore worse and 4 and 5 better).

		Pugh Concept Selection Matrix	Wieght	Electric 4-slot	Toasting Fork & Fire	Electric Conveyor	Gas Grill	
Selection Criteria	Good Toast Quality	Even Toasting		S				
		Good Taste		S				
		Repeatable		S				
		Quick		S				
	Capacity	Large Range of Bread Products		S				
		Multiple Slices/Units		S				
	Long Life	Reliable		S				
		Durable		S				
		Low Maintenance		S				
	Physical Attributes	Affordable		S				
		Attractive		S				
		Safe		S				
		Good Size		S				
	Easy to Use	Easy to use Controls		S				
		Easy to Load		S				
		East to RemoveToast		S				

TOTAL +		0	
TOTAL -		0	
TOTAL SCORE		0	
WEIGHTED SCORE			

Figure 2: Initial Pugh Matrix for the Toaster Selection Problem

Note that weightings have yet to be given to the criteria. On a personal note, even if the weightings already exist, I tend to conduct the first assessment without using the weightings, the second with. This provides a "rough and ready" robustness assessment.

In the third step, the pairwise evaluation takes place and the results for the toaster selection is shown in figure 3.

		Pugh Concept Selection Matrix	Wieght	Electric 4-slot	Toasting Fork & Fire	Electric Conveyor	Gas Grill		
Selection Criteria	Good Toast Quality	Even Toasting		S	- -	S	-		
		Good Taste		S	+ +	S	S		
		Repeatable		S	- -	+	-		
		Quick		S	+	S	S		
	Capacity	Large Range of Bread Products		S	+	S	+		
		Multiple Slices/Units		S	- -	+	+		
	Long Life	Reliable		S	+	-	S		
		Durable		S	+	S	S		
		Low Maintenance		S	- -	-	S		
	Physical Attributes	Affordable		S	+	-	+		
		Attractive		S	+	-	-		
		Safe		S	- -	-	- -		
		Good Size		S	- -	-	-		
	Easy to Use	Easy to use Controls		S	-	S	+		
		Easy to Load		S	-	+	-		
		East to RemoveToast		S	- -	+	-		

	Electric 4-slot	Toasting Fork & Fire	Electric Conveyor	Gas Grill		
TOTAL +	0	8	4	4		
TOTAL -	0	18	6	-8		
TOTAL SCORE	0	-8	-2	-4		
WEIGHTED SCORE						

Figure 3: Initial Pugh Evaluation

The evaluation performed in Figure 3 indicates that with a total score of 0, the 4-slot electric toaster is the winner. The electric conveyor is a close second scoring -2, and the gas grill third at -4. In a distinct fourth place is the toasting fork and fire! At this point it is worth examining and debating why one concept is better than another. Moreover, performing a sanity check – "does the outcome make sense?" Quiet often it does not and this can be a consequence of not having the complete or correct evaluation criteria. In this instance it does "feel" about right although the score for the gas grill was close to either of the electric toasters. Indeed, debate led to the

consideration of a further criterion "the toaster should be automatic". It is a simple matter of including this criterion, assigning scores and re-calculating the total score as shown in figure 4.

		Pugh Concept Selection Matrix	Wieght	Electric 4-slot	Toasting Fork & Fire	Electric Conveyor	Gas Grill		
					Design Concepts				
Selection Criteria	Good Toast Quality	Even Toasting		S	- -	S	-		
		Good Taste		S	+ +	S	S		
		Repeatable		S	- -	+	-		
		Quick		S	+	S	S		
	Capacity	Large Range of Bread Products		S	+	S	+		
		Multiple Slices/Units		S	- -	+	+		
	Long Life	Reliable		S	+	-	S		
		Durable		S	+	S	S		
		Low Maintenance		S	- -	-	S		
	Physical Attributes	Affordable		S	+	-	+		
		Attractive		S	+	-	-		
		Safe		S	- -	-	- -		
		Good Size		S	- -	-	-		
	Easy to Use	Easy to use Controls		S	-	S	+		
		Easy to Load		S	-	+	+		
		East to RemoveToast		S	- -	+	-		
		Automated		S	- -	S	- -		

	Electric 4-slot	Toasting Fork & Fire	Electric Conveyor	Gas Grill		
TOTAL +	0	8	4	4		
TOTAL -	0	18	6	-9		
TOTAL SCORE	0	-10	-2	-5		
WEIGHTED SCORE						

Figure 4: Toaster example with the inclusion of a new criterion.

Figure 4 shows the inclusion of the additional criterion has not changed the ranking of the options. If anything, it has make the case for the "fire and toasting fork" option less attractive - but before dropping the toasting fork it well worth assessing if it has any strengths that could, in a design situation, be "transferred" to the competing options.

At this point, even with the elimination of the "fire and fork" option there are still potential candidates (although the gas grill options is looking weak). To help discriminate between the options the criteria can be weighted. Figure 5 shows a Pugh Matrix for the remaining Toaster options with weighted criteria.

		Pugh Concept Selection Matrix	Wieght	Electric 4-slot	Electric Conveyor	Gas Grill	Design Concepts	
Selection Criteria	Good Toast Quality	Even Toasting	2	S	S	-		
		Good Taste	3	S	S	S		
		Repeatable	3	S	+	-		
		Quick	3	S	S	S		
	Capacity	Large Range of Bread Products	2	S	S	+		
		Multiple Slices/Units	4	S	+	+		
	Long Life	Reliable	1	S	-	S		
		Durable	3	S	S	S		
		Low Maintenance	3	S	-	S		
	Physical Attributes	Affordable	2	S	-	+		
		Attractive	5	S	-	-		
		Safe	3	S	-	- -		
		Good Size	4	S	-	-		
	Easy to Use	Easy to use Controls	5	S	S	+		
		Easy to Load	4	S	+	+		
		East to RemoveToast	4	S	+	-		
		Automated	4	S	S	- -		

		Electric 4-slot	Electric Conveyor	Gas Grill		
TOTAL +		0	4	4		
TOTAL -		0	6	9		
TOTAL SCORE		0	-2	-5		
WEIGHTED TOTAL +		0	15	17		
WEIGHTED TOTAL -		0	17	32		
WEIGHTED SCORE		0	-2	-15		

Figure 5: Pugh Matrix with Weighted Criteria

Weighting the selection criteria can often provide an extra level of discrimination when making decisions. It can also provide a form of "robustness" assessment – "does the ranking of the options change?" What is clearly important when using weightings in a Pugh Matrix is the validity of the weightings! It is possible to select criteria weightings to influence the ranking outcome. There are many approaches that can be adopted to arrive at a sensible set of weightings that range from absolute scales, through to the use of order winners and the Analytic Hierarchy Process (AHP). A brief description of these is given in Appendix A

Returning to the example shown in Figure 5, the use of weightings clearly now eliminates the "gas grill" option. It also shows that the two remaining options are very close. This outcome is typical when using a Pugh Matrix – it can often result in the elimination of weaker options, but rarely identifies a winning option cleanly. Fundamentally, the Figure 5 is saying that there is insufficient information to make the decision between the two remaining options and further work is necessary.

What Goes Wrong: The limitations of the Pugh Matrix

Incorrect, incomplete and inadequate selection criteria. The "quality" of the decision using a Pugh Matrix is fundamentally related to the "quality" of the selection criteria. This quality has three aspects:

Incorrect selection criteria will obviously lead to the wrong decision. Typically incorrect criteria occur because the team using the Pugh Matrix use their opinions when identifying the criteria rather than those of the stakeholder(s). It is important to consider validating the criteria in some way or using criteria that have already been validated as representing the views of the stakeholder(s).

Incomplete selection criteria are obviously where we do not have a complete set of selection criteria. This can be because we have:

- not considering all appropriate stakeholders, just using "user-based" criteria thereby ignoring potentially important aspects such as technology readiness, manufactability, ease of implementation etc.
- deliberately ignored criteria because we thing they are not important or don't fit our opinions

It is important to put the time and effort into determining the criteria and to consider some form of validation.

Inadequate selection criteria are where we attempt to use poorly defined criteria that can have multiple interpretations. For example "low cost" is an inadequate criteria because unless it is clearly defined it can have multiple meanings including, purchase cost, running costs, development costs etc.

Care must be exercised when developing suitable criteria to ensure that the team has a clear, consistent and, where necessary, documented set of selection criteria.

Granularity of pairwise scale. One of the strengths of the Pugh Matrix is also one of its weaknesses – the low granularity of the pairwise scale. The simple scale does permit a rapid evaluation of

options, but the low granularity can give poor decision robustness. It is possible to get a different ranking of options by changing the baseline option – particularly if the +, S, - scale is used. To some extent this can be moderated by using the ++, +, S, -, -- scale. Some users even recommend using a baseline of 5, with the numbers 4, 3, 2, and 1 representing levels of worsening satisfaction and 6, 7, 8, and 9 representing levels of increasing satisfaction.

Outcome can also be " validated" by performing simple robustness assessments. There is also "engineering judgement" – the outcomes "feel" right. Indeed, outcomes that are unexpected should be carefully reviewed as to why they have occurred.

Wrong expertise and insufficient experience in teams. Like a great many Systems Engineering tools, the Pugh Matrix is really only a vehicle to help extract the knowledge and experience from the team. The wrong team can still follow the process and arrive at a result – but the result may not robust.

Success Criteria

The following list represents a set of criteria that have been found to be useful when using a Pugh Matrix.

- Team size between 4 and 8
- Team constitution has expertise and experience in the system of interest but can (and perhaps should) include members with limited experience and expertise
- Use an experience independent facilitator
- Plan for 2-3 hour's effort.
- Define clearly what we are trying to do
- Have validated and weighted selection criteria available (the "customer requirements" from Quality Function Deployment 1 for example)
- Spend time to check-out the team's understanding of the criteria and if necessary clearly define (and document) the agreed understanding (very useful in subsequent design reviews)
- Perform robustness checks by flexing weightings or pairwise decisions
- Document any debate

References

[1]   Burge S E (2006) "Matrix Diagram" from the Systems Thinking Tool Box www.Burgehugheswalsh.co.uk
[2]   Pugh, S. (1991). Total Design: Integrated Methods for Successful Product Engineering. Addison-Wesley. ISBN 0201416395
[3]   Chambers

Appendix A: Overview of approaches for weighting selection/evaluation criteria

The criteria weightings can significantly affect the quality and accuracy of the outcome from a Pugh Matrix. There are a number of approaches that can be used to help determine these which include:

> Absolute Scale
> Sales Driven Scale
> Analytic Hierarchy Process

A brief overview of these options is presented below

Absolute Scale

The absolute scale uses the following definitions to weight the criteria:

> 1 minor importance to the stakeholder
> 2 moderate importance to the stakeholder
> 3 important to the stakeholder
> 4 very important to the stakeholder
> 5 extremely importance to the customer

The absolute scale is perhaps the easiest to apply but does lead to only the upper numbers being used (3 to 5). This is because if the stakeholder has articulated a requirement then it is hardly going to be rated as a 1. There is also a tendency for the scale to over-weight certain "givens" and this is where the sale driven scale can be useful.

Sales Driven Scale

An alternative to the absolute scale is the sales driven scale based on the concept of "Order Winners" and "Order Qualifiers"

- Order Winners: those aspects of a product that WIN it orders in the market place. If we are better than our competition we will win on these aspects. These are sometimes called Unique Selling Points (USP)

- Order Qualifiers: those aspects of a product that KEEP it in the market place. We only need to be as good as the competition in these aspects. Bettering the competition does not win any further orders. These are sometimes called 'givens'. However, these can be *order-losing* sensitive. That is, we must be as good as the competition; if we are worse we will not win ANY orders.

The affect of order winners and qualifiers is shown pictorially in figure A1.
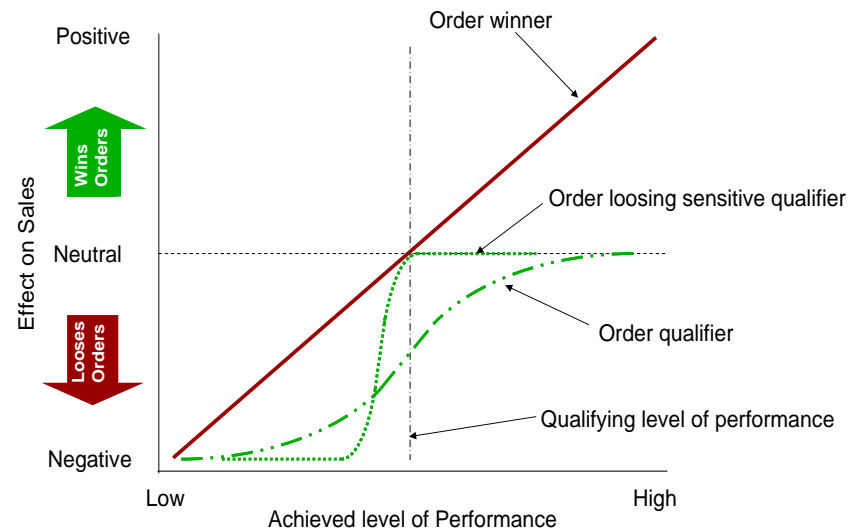

Figure A1: The Affect of Order Winners and Qualifiers on Sales (adapted from [3])

Order qualifying criteria are 'givens' that the customer expects to be there at a certain level. Exceeding this level is unlikely to bring in more sales and failure to deliver to this level leads to customer dissatisfaction. On the 1 to 5 scale, such customer requirements are rated as a 1 or 2.

Some qualifying criteria can be "sensitive". In such cases failure to reach the perceived minimum standard will result in a dramatic loss of orders. These criteria are typically aspects, like safety, where a minimum level is expected (due to legislation for example) and failing to meet that level will halt customer purchases – but exceeding the minimum level will not realise any more sales. These are called *order losing sensitive qualifiers*. Such criteria are rated between 3 or 4.

Order winners are those aspects in which if we are better than our competitors will result in an increase in sales. These are rated 4 or 5.

We have to think carefully about order winners and qualifiers because they do change over time. For example when the Japanese car industry first started selling overseas, it competed on price – price was the order winner. It was possible in the late 60's and early 70's to purchase a Japanese car for 20% less than its competitors; companies like Nissan (Datsun), Honda, and Toyota increased their sales at the expense of more established companies like Ford's and General Motors. The competitors had no choice but to lower their prices to compete – and in the UK several automotive manufactures went bankrupt. This lowering of prices, made purchase price become an order qualifier. In fact, this was also riven by the fact that the Japanese automotive industry realised that the order winner was now quality and reliability and began to improve in this area and once again their sales increased. Japanese motorcars are now known to have the best reliability and quality [1] – even more of UK automotive manufacturers

went bankrupt. Quality and Reliability are of course now givens – i.e. order qualifiers.

In summary the Sales Driven scale is:

    1 Very much expected by the customer as a given
    2 Generally expected by the customer
    3 Expected by the customer but failure to achieve will significantly reduce sales
    4 Either a very sensitive order qualifier or a gradual order winner
    5 A sensitive order winner. Changes in performance have a significant affect on sales

Using the sales driven scale is more difficult because the stakeholder usually cannot provide the information directly. The stakeholder has to be questioned about each requirement as well as an internal interpretation of the requirements.

Analytic Hierarchy Process (AHP)

The Analytic Hierarchy Process provides an effective mechanism for group or team decision-making by imposing a discipline on the group's thought processes. It helps the group assign numerical values to subjective judgments and translate these into importance ratings. What makes the approach useful is that the importance ratings are on a true ratio scale thereby allowing the team to say how much more important one item is than another.

In the context of the Pugh Matrix, AHP determines the criteria weightings though the pairwise comparison of the criteria. It also allows for the checking of teams consistency in this pairwise comparison providing a degree of validation of the weightings. Figure A2 shows an example AHP output (this is the output from a software package called Qualica [2]).
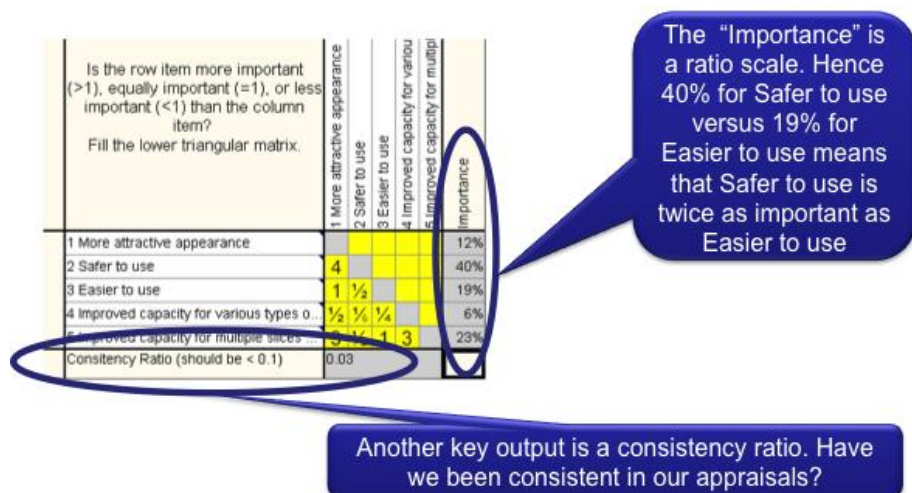

Figure A2: An Example AHP output from Qualica[TM]

In Figure A2, the Analytic Hierarchy Process has been used to determine the importance ratings of 5 requirements:

1. More Attractive Appearance
2. Safer to use
3. Easier to use
4. Improved capacity
5. Improved capability for multiple slices