

Running head: NATURALISTIC DECISION MAKING THROUGH ARGUMENTATION

Running head: NATURALISTIC DECISION MAKING THROUGH  
ARGUMENTATION

**Naturalistic Decision Making Through Argumentation: Resolving Labour Disputes**

Sivanes P. Ramiah and Adrian P. Banks\*

<sup>1</sup> University of Surrey, UK

Word count: 7515

\*Requests for reprints should be addressed to Adrian P. Banks, School of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, UK (e-mail: a.banks@surrey.ac.uk).

## Abstract

Naturalistic Decision Making research has been successful in describing fast, intuitive decision making by experts in complex, time pressured situations. However, less is known about how experts make naturalistic decisions when they think analytically. Previous research suggests that this process involves constructing a story of the situation to explain and simulate the consequences of decisions, but we hypothesise that explicit argumentation is valuable when it is necessary to demonstrate the accuracy of a decision and to explicitly resolve inconsistent and uncertain information. This study investigates analytic naturalistic decision making by contrasting experts and novices resolving labour disputes. Fifty expert and novice labour officers resolved labour cases whilst thinking aloud. The verbal protocols were analysed using Toulmin's argument model. Expert decision makers made more accurate decisions, used more warrants (in other words, reasoned justifications) during the decision process, and increased use of warrants was associated with accurate decision making. In addition, a decision aid was developed in the form of a checklist that encouraged decision makers to explicitly consider components of the argument. This increased the frequency of warrants and the decision accuracy of both novice and expert decision makers. Together, these findings imply that accurate and expert analytic decision making in this domain involves a process of argumentation and in particular the use of explicit warrants linking evidence to conclusions.

*Keywords:* Naturalistic Decision Making; Toulmin; Argumentation; Analytic thinking; Reasoning

### Practitioner Points

- Analytic Naturalistic Decision Making in this domain involves argumentation
- Superior decision accuracy arises from exploring more reasoned justifications
- A checklist encouraging all components of the argument to be made explicit improves decision accuracy in both novice and expert decision makers

The aim of Naturalistic Decision Making (NDM) research is to understand how experienced people make decisions in meaningful, real-world contexts (Klein, 2008; Lipshitz, Klein, Orasnu, & Salas, 2001). Research in this field has been very successful in discovering new and important decision making processes in a number of domains such as fire ground commanders (Klein, Calderwood, & Clinton-Cirocco, 1986; 2010); neonatal nurses (Crandall & Getchell-Reiter, 1993); and military intelligence analysis (Roth et al., 2010). These domains are typical of those studied within the NDM framework. They share the factors that Orasnu and Connolly (1993) identified as characteristic of naturalistic decision making contexts. These are: ill-structured problems; uncertain, dynamic environments; shifting, ill-defined, or competing goals; action/feedback loops; time stress; high stakes; multiple players; and organizational goals and norms. A second characteristic of NDM theories is that they describe how expert decision makers draw upon their experiences when making decisions (e.g. Salas & Klein, 2001).

A common finding is that NDM often involves an intuitive decision making process (Kahneman & Klein, 2009; Klein, 2008). For example, the most influential model of naturalistic decision making is the Recognition-Primed Decision Model (RPD) (Klein, Calderwood & Clinton-Cirocco, 1986; 2010). This arose from research on fireground commanders who are faced with the task of making fast decisions. A key insight was that they do not generate and assess multiple alternatives. Instead, they primarily recognise patterns of causal factors in a situation that have been learnt over time and these suggest the most typical course of action in that situation. This provides a good account of extremely rapid yet effective decision making. It is an effective method of decision making because of both features of NDM – characteristics of the decision making domain and expertise of the decision makers. The time pressured domain does not allow for a slower, analytic decision making process and the expertise ensures that the rapidly retrieved course of action is an appropriate one.

Much NDM work has focused on domains such as the military and the emergency services and in these domains recognition-primed decisions have been shown to be very common (e.g. Kaempf, Klein, Thordsen, & Wolf, 1996; Klein et al., 2010). However, not all real-world contexts share their characteristics. In particular, many business and organizational contexts do not require decisions to be made within seconds. There may be time pressure, but there is still time to analyse a decision rather than respond intuitively. Second, a recognition strategy can only be effective in a predictable environment with feedback about the success of a decision, so that the best course of action for a given situation can be learnt over time (Kahneman & Klein, 2009; Shanteau, 1992a). But if the situation is novel then there is no typical course of action. The decision maker must use other strategies to make a decision. Many decisions in business and organisational settings are novel and so will not be suited to a recognition strategy. Therefore, more so than typical domains of NDM research and application, a slower analytic decision process is possible and potentially valuable in many business and organizational contexts.

Whilst there are some NDM studies in domains that have greater opportunity for analytic decision making such as day traders (e.g. McAndrew, Gore, & Banks, 2009; McAndrew & Gore, 2013) these studies are less common and there is less understanding of how analytic naturalistic decisions are made. The aim of this study is to develop this understanding by comparing current theoretical explanations of analytic NDM with an alternative account that we propose based on decision making through argumentation. Second, we will apply our account practically by developing and testing a decision aid to support analytic decision making. We will do this in a domain where there is an opportunity for analytic NDM: resolving labour disputes.

### **Analytic Naturalistic Decision Making**

In many naturalistic situations people make decisions analytically in which they cannot readily evaluate one course of action over another because of the complexity of the situation or lack of information about it (Green & McCloy, 2003; Lipshitz, 1994; Shafir, Simonson, & Tversky,

1993). Therefore they cannot or do not apply classical theories of choice (e.g. Edwards, 1961). NDM research has shown some of the alternative decision processes that are used. The main analytic decision making process proposed in NDM models involves story building. This occurs when the situation does not clearly match a familiar pattern (Klein, 2008). The decision maker links the information available into a coherent account that explains the situation. This narrative explanation of the facts aims to provide a plausible explanation of the events that led to the current situation. It also allows a mental simulation of the future consequences of following a course of action that is used to evaluate the effectiveness of that action.

A number of studies have shown that decision makers construct stories of events during the decision making process. Kaempf et al. (1996) found that 12% of decisions in naval command-and-control environments depended on story building. Armenis and Neal (2008) found that senior managers making employee promotion decisions constructed stories about the consequences of decisions in 81% of cases. Lipshitz and Shulimovitz (2007) found that negative feelings were aroused in loan officers making credit decisions if a client's story was unpersuasive, suggesting a more emotional evaluation of the coherence of a story.

A story building model has also been proposed by Pennington and Hastie (1986, 1988, 1992) to describe juror decision making. According to this model, jurors use the evidence to construct a story of what happened. The story is an explanation in terms of the causes and intentions of the people and events and how these unfold over time. The story is then used as a basis for choosing the verdict. Hence these theories propose that decision makers construct a chronological narrative of events to guide their decisions.

However, building a chronological story is only one way of forming and testing a coherent account of a situation. Schum (1993) referred to this approach of developing a chronological story of legal evidence as a temporal structure. He contrasts this with a relational structure in which events and information are combined in novel ways to form reasons or justifications for a conclusion. This might involve bringing together temporally disparate or diverse facts rather than

only linking them chronologically. Thus relational structures represent explicit arguments in which evidence is linked to conclusions. This might include some narrative elements, but is not limited to it. Weinstock (2011) likened the contrast of temporal and relational structures to Bereiter and Scardamalia's (1987) distinction between 'knowledge-telling' and 'knowledge-transforming'. Whereas in knowledge-telling people construct a representation that simply tells the story of what happened, in knowledge-transforming people select, reorganise, integrate, and synthesise information in alternative ways to meet a goal of drawing conclusions about the situation.

Schum proposed that temporal and relational structures are used by legal professionals. Explicit argumentation has also been proposed in other areas of expert decision making. Cohen, Freeman, and Wolf (1996) propose an NDM model in which people assess uncertain situations by developing arguments. Their recognition/metarecognition model suggests that people recognise arguments linking evidence (e.g. gunboat turning towards ownship) and conclusions (e.g. therefore the gunboat is hostile). These are then critiqued to improve the accuracy of the argument and the resultant decision. However, whilst this model is popular, there is little published empirical research testing it and so the prevalence of this form of argumentation is not known.

Under what circumstances would an argument-based decision process be favoured over a story building process? The construction of arguments is more common when decision makers are motivated to explicitly justify or persuade that a decision is correct (e.g. Mercier & Sperber, 2011) or to explicitly resolve an uncertain issue (e.g. Cohen et al., 1996; Schum, 1993). This contrasts with the story building process in which one version of the story is selected because it is intuitively more plausible (e.g. Pennington & Hastie, 1986). Therefore the difference lies in the use of explicit reasons to support a decision, resolve uncertainty or contrast alternative possibilities.

Overall, previous research suggests that analytic NDM involves the construction of a story of the situation that is used to explain the events that led to the current situation and simulate the consequences of a course of action. However, a coherent account of a situation need not take the form of a chronological sequence of events. Events could be linked through diverse arguments to inform a decision. This may be more common in domains where explicit argumentation to support a decision or to resolve uncertainties is necessary or is an organisational norm. These domains have been less frequently studied within NDM (Gore, Banks, Millward, & Kyriakidou, 2006). We will explore analytic naturalistic decisions in a domain like this in order to discover if argumentation-based decision making is more common under these circumstances, and if so, how arguments are used.

### **Overview of the Study**

In this study we examine naturalistic decisions that are made in a context that allows the opportunity for analysis and deliberation. The domain we study is resolving labour disputes. Labour disputes display many of the characteristics of naturalistic decision making. Labour officers are required to resolve disputes using evidence about the case that is complex and uncertain, and they draw upon their experience to establish the correct course of action in each unique case. We will compare expert and novice labour officers within the Department of Labour in Malaysia to establish the differentiating characteristics of expert decision makers.

Participants will be asked to decide on the correct course of action in a series of realistic labour dispute scenarios. Whilst they are doing this they will think aloud, allowing us to capture their decision processes within a verbal protocol (Ericsson & Simon, 1980). This method has been used successfully before to analyse naturalistic decision making in intelligence analysis (Roth et al., 2010) and in the development of explanation-based decision making by Pennington and Hastie (1986). One concern about the use of verbal protocols is that requiring participants to think aloud could influence their thought processes. For example, Nisbett and Wilson (1977)

found that participants often gave inaccurate explanations for their behaviour because they did not have access to their own cognition. However, more recent research shows that the validity of the approach is determined by the verbalisation instructions used (Ericsson, 2002; Ericsson & Simon, 1993). The least interference occurs if participants are simply asked to report what their current focus of attention is, rather than explain or justify their behaviour. This instruction captures what is on the mind of the participant but does not change it. This is the instruction that we adopt and therefore we believe that our verbal protocols will not have greatly influenced the decision process.

The protocols will be coded to assess the frequency of different components of the decision process in the different experimental groups to establish if they are more consistent with a story building or argumentation-based approach. To do this we will apply Toulmin's model of arguments (Toulmin, 1958). If decisions are made based on arguments, this coding scheme will extract the components of those arguments. If, on the other hand, decisions are made through constructing narrative accounts of the evidence, this coding scheme will extract the events used within the story and the decision that follows from it but no further components of explicit argument will be found.

Toulmin's model of arguments is an influential account of practical reasoning; an approach to reasoning that explores how reasons are used in deciding how to act (e.g. Erduran, Simon, & Osborne, 2004; Hitchcock, & Verheij, 2007; Prakken, 2005). He describes the essential elements of an argument as the relationship between a claim that is being made, the evidence that supports the claim, and the explanation why that claim is justified by the information presented. In a labour dispute, the claim is the decision or conclusion that has been made that resolves the case. An example of a *claim* in this context is 'the worker is eligible for his termination benefit'. The information or evidence that supports the claim is referred to as *data*. An example of data is 'the employment of the worker has been terminated without the benefit agreed in the worker's employment contract'. The justification that the data supports the claim is referred to as a



*warrant*. An example of a warrant is ‘if he worked for more than a year, then the worker is eligible for termination benefit’. In addition to these essential components of an argument, there are three auxiliary components. If the warrant is fallible, it can be reinforced with *backing* from more general statements. An example of a backing statement in a labour dispute is to support the warrant with reference to the law, e.g. ‘the worker is protected under the Employment Act’. Sometimes claims are questioned with a *rebuttal* e.g. ‘unless the employer proved that the worker was employed for less than one year’. Finally, claims can be moderated with a *qualifier* e.g. ‘probably’ or ‘almost certainly’.

If participants construct a chronological narrative of the events there will be a high frequency of data statements in their protocols as they refer to evidence about events in the scenario. They will also make claims, which are their decision based on the narrative. However, the other codes refer to components of an argument that are not frequently used in constructing a narrative account and so will not be common. Figure 1 presents an example of a story that is built from the evidence in a scenario to illustrate the coding of events as data in a chronological sequence.

-----

Insert Figure 1 about here

-----

If participants make a decision based on arguments, there will be data statements and claims in the protocols as they refer to evidence and report their decisions. In addition, there will be more warrants and backing statements. These statements provide the reasons that link the evidence to the conclusion in arguments and so will be more frequent in argument-based decisions. If the argument process is highly sophisticated, as Schum (1993) suggests, participants will also use rebuttals and qualifiers to explicitly construct counter-arguments and moderate the strength of conclusions. Figure 2 presents an example of an argument for a decision linking

evidence to a conclusion. The key difference in this structure compared to Figure 1 is that the evidence is not simply retold within a narrative (i.e. employee offered job 18/2/2004) but that this fact is interpreted to focus on the salient feature – that the employee has been in work for more than one year. This is then combined with the relevant warrant – employees who work for more than one year are eligible for termination benefit – to present an explicit argument justifying the decision. This illustrates the coding of a simple argument applying the components from Toulmin’s argument model. The argument could be elaborated with more data, rebuttals, etc.

-----

Insert Figure 2 about here

-----

Our first aim is to compare expert and novice labour officers to establish the differentiating characteristics of expert decision makers. We hypothesise that expert decision makers will make more accurate decisions than novices, i.e. their claims will be correct, reflecting their expertise in this domain. It has been shown that experts do not rely on more information than novices (Shanteau, 1992b) and so we hypothesise that there will be no difference in the frequency of data statements between expert and novice decision makers. Finally, the critical hypothesis to contrast the story building and argument-based decision making account depends on the frequency of warrant and backing statements. If the story building account is correct, these statements will not be used frequently by either group and there will be no differences between them. However, if expert decision making in this domain does rely on argumentation, the expert officers will make more frequent use of warrants and backing statements than novices. Furthermore, across both conditions, the frequency of warrants and backing statements will correlate with decision accuracy if these components of arguments are used in good decision making. Finally, if expert

decision making in this domain relies on sophisticated arguments, the expert officers will make more frequent use of rebuttals and qualifiers.

Our second aim is to improve decision making in this domain using a decision aid. We will develop a decision aid that asks questions that will improve decision making. The questions will be presented in the form of a checklist. Checklists have often been used in organizational interventions (e.g. Bacon, Fulton, & Malott, 1983; Doll, Livesey, McHaffie, & Ludwig, 2007) and their value in improving decision making in applied settings has been popularised by Gawande (2010). Checklists are simple and robust, increasing the ease with which they can be applied in practice and making them an ideal aid for naturalistic decision making.

We will construct the decision aid using a focus group of subject matter experts (SMEs) to discover the key information used and questions asked in making a decision in this domain. From this we will develop a series of questions that encourages decision makers to follow the process advocated by the SMEs. Thus the questions chosen were not taken from an imposed theoretical framework, they emerged from the consensus that the SMEs formed. To investigate if the SMEs advocate a role for argumentation in their set of key questions we will map the questions onto Toulmin's argument model. We predict that if decisions are made through a process of argumentation, questions in this aid will map onto the components of Toulmin's argument model. If argumentation is not used, then they will not.

We hypothesise that novice decision makers using the decision aid will make more accurate decisions than those without. We further hypothesise that, if the decision aid does encourage argumentation, the decision aid will increase the frequency of data, warrant, and backing statements as the checklist would encourage all of these components of the argument to be made explicit. However, we hypothesise that the decision aid will not influence the accuracy or decision process of expert decision makers as they already perform at a high level in this task without the aid.

## Method

### Organizational Setting

The study was carried out in the Department of Labour in Peninsula Malaysia in 2011. The labour department ensures that workers' rights and welfare are protected without jeopardising the interests of employers. This study focuses on how labour officers within the department resolve disputes between workers and employers. Labour disputes are defined as the failure of employers or employees to fulfil the provisions and requirements set out under the Employment Act 1955 or its regulations. The jurisdiction of the Labour Court in resolving labour disputes in the private sector involves an important decision process and is a core activity in the Department of Labour.

Although there is a legal framework that must be applied, this does not mean that the decision is made simply by applying a series of rules to reach a decision. The cases that labour officers resolve are complex and uncertain. It is in the nature of a dispute that the solution is not resolvable in a straightforward manner. The cause of the uncertainty includes ambiguous, conflicting or missing information. For example, one scenario presented a dispute over non-payment of an employee's salary. This included the statement that 'In the contract of service it was stated that the employer will bear the cost of school fee and buy return air tickets to all family members once a year without quoting the amount.' The labour officer must then address questions such as: can the money be claimed instead of an air ticket? Why has this particular size of claim been chosen and is it valid? To what extent does this reflect current ticket prices, or prices in previous years? All of the labour cases contain uncertain points. Furthermore, by definition, the disputing employer and employee present conflicting accounts of these uncertain points. Hence there is not a clear set of facts that the law unambiguously maps onto. The essence of resolving a labour dispute is to reach a decision based on uncertain and conflicting information.

## Participants

Fifty labour officers from the Department of Labour in Malaysia participated in the study. Participants were allocated to the expert group on the basis of the number of cases that they had worked on, their length of service in the department, and their job level. The expert officers had a mean age of 42.7 years ( $SD = 7.04$ ), had worked on a mean of 890.80 cases ( $SD = 779.92$ ) throughout their career at the time of the study, and had served for a mean of 15.80 years ( $SD = 7.60$ ). There were 12 men and 13 women in this group. The novice group were trained and qualified in resolving labour disputes and were knowledgeable about the context and the procedures, but had not yet worked on many real cases. The novice officers had a mean age of 29.68 years ( $SD = 3.79$ ), had worked on a mean of 21.84 cases ( $SD = 37.06$ ) throughout their career at the time of the study, and had served for a mean of 2.24 years ( $SD = 0.88$ ). There were 14 men and 11 women in this group. The officers job grades ranged from S27 (assistant labour officer) through to S51 (chief assistant director). Assistant labour officers possess a higher secondary certificate and can be promoted after 10-15 years of service. Higher grades possess a degree in various subjects typically HR, economics or law. The frequency of grades represented in the expert group were S52 = 1, S48 = 1, S44 = 12, S41 = 6, S38 = 4, S32 = 1. The frequency of grades represented in the novice group were S41 = 15, S27 = 10.

As a manipulation check, we assessed the differences in the number of cases resolved in the different experimental conditions. The expert group had resolved more cases than the novice group ( $F(1, 46) = 31.60, p = .001, \eta_p^2 = .41$ ), demonstrating that they were indeed more experienced. However there was no difference in the number of cases resolved in groups that completed the task with and without the decision aid ( $F(1, 46) = 0.64, p = .43, \eta_p^2 = .01$ ), demonstrating that differences between these groups were not attributable to differences in experience.

## **Materials**

Six scenarios were developed for use in the study by the first author, a subject matter expert. These were based on actual cases filed at the department and are in accordance with the requirement of the Employment Act 1955. The scenarios were comprised of a description of the situation and evidence from the complainant and the defendant. Three of the cases involved a claim that employment was terminated without payment of wages and three involved a claim that employment was terminated without complying with required regulations. Given the evidence provided, it was possible to reach a legally correct decision in all scenarios. The correct decisions in each scenario were determined by the first author in consultation with other subject matter experts from the Department of Labour. An example scenario is:

### *Description of the case*

Nova Express Sdn. Bhd. [a courier service] charged for non payment of termination benefit @ 39 days RM2, 179.71; balance of notice in lieu RM880.00; and medical leaves @ 5 days RM274.20 to Moorty Govindasamy.

### *Evidence from the complainant*

I was offered a job as lorry driver on 18/2/2004. My last drawn salary was RM1700 per month. I was not given an appointment letter or confirmation letter. I had only few payslips as evidence which I insisted my employer to give me. I was informed verbally (somewhere in August) by my employer to find a new job within two to three months because the company could not afford to continue the business.

On the 2/10/2006, my wife called the employer and informed him that I was admitted into hospital due to chest pain. My wife also told him that I will return to work after discharge from hospital (I need to go back work because all my children are still studying and my wife is not working). I was discharged on 6/10/2006 and

the medical leave given until 7/10/2006. The next day (8/10/2006) was Sunday which was my rest day. I have only taken 2 days medical leave so far for this year. I went to work as usual on Monday (9/10/2006); my employer said that I was terminated with a reason and that the company did not have any job to provide to me. I handed over my medical certificate for 2-7/10/2006 on 9/10/2006. It was not accepted by my employer and I was paid a compensation of RM1500.

However, a new driver was employed on the 5/10/2006 (informed by the worker). I found that it was very unfair to me and the amount paid was not according to labour law.

*Evidence from the defendant*

As one of the directors, I agree that the company was not doing well. However, the worker had been informed verbally that his last working day will be on the 1<sup>st</sup> September 2006. He agreed to accept RM1500 as compensation and walked off without any disapproval.

After reading each case the labour officers decided how it should be resolved.

**Development of the decision aid.** The decision aid was developed through a focus group with four SMEs and the first author (also an SME). The aim was to develop a set of questions that could form a checklist to ensure that the decision maker has addressed all of the key points in the case and had followed a sound decision making process to reach a good decision. To do this, the SMEs discussed two questions. ‘What are the most important questions that you ask yourself when making a decision?’ and ‘What is the most important information to look at when making a decision?’. The SMEs generated a number of possible answers to these questions that were discussed before finally agreeing upon a set of six questions. We assessed if these questions can be mapped onto Toulmin’s argument model, and it is notable that they each emphasise the need

to identify different components of the argument model, i.e. data (Q3 & Q5), warrant (Q2), backing (Q1), and rebuttals (Q4) and to form an explicit representation of the case (Q6).

The questions were:

1. Determine whether the complainant covers under section 69 or 69B of the Employment Act 1955.
2. Reveal which sections and provisions of law that the claims made by the complainant are based on.
3. Consider the verbal statement presented by the complainant, defendant and the witnesses (if any).
4. Is the verbal statement of one party challenged by the other party?
5. Is there any supporting (written) evidence related to the verbal evidence produced?
6. Do you think the inability to visualise the trial process influenced your judgment or decision?

## **Procedure**

Half of the participants completed the task without the decision aid and half completed it with the aid. Those with the aid were given the six questions and were asked to consider each of the questions as they made their decisions. Participants were tested individually. Data were recorded with a digital voice recorder. First, think aloud instructions were explained to participants followed by warm up think aloud problems such as ‘what is the sixth letter after B?’. Participants then read through the scenarios and thought aloud as they reached a decision. They were instructed to verbalise any thoughts that came to mind, but not to explain their decisions. If the participants fell silent, the experimenter encouraged them to continue thinking aloud with prompts such as ‘just say aloud what you are thinking?’. Without the decision aid, the data collection for expert officers lasted between 35:29 and 63:21 minutes whereas for the novice



officers data collection lasted between 31:47 and 56:49 minutes. With the decision aid, data collection for expert officers lasted between 23:13 and 61:03 minutes whereas for novice officers data collection between 24:31 and 80:54 minutes.

## Results

### Protocol analysis

The verbal protocols were transcribed and coded by the first author, a subject matter expert. A randomly selected subset of 20% of the protocols was rated by a second subject matter expert who was blind to the experimental hypotheses. Cohen's Kappa was 0.93 indicating a high level of inter-rater agreement (Landis, & Koch, 1977).

The verbal protocols were coded to determine the frequency of data, warrants, claims, backing, qualifiers, and rebuttals. However, after initial coding it became clear that rebuttals and qualifiers were not used by these participants. These elements are not considered by Toulmin to be essential components of an argument. Therefore only the three essential elements of an argument – data, warrant, and claim – and backing were coded. The elements of the argument were identified within the transcript using the method proposed by Fisher (2004). He describes specific words that signal either reasons or claims. Common words or phrases that often signal reasons are: Because...; For...; Since... etc. Within Toulmin's model data, warrants, and backing are all part of the reasons given for a claim. So having identified a reason, it was further coded as data, backing or warrant. Data statements were identified because, in this study, data are derived from evidence presented in the scenario. So descriptions of the evidence were coded as data statements. Backing statements were identified because, in this study, backing for the argument is derived from the Employment Act. So references to the act were coded as backing statements. Warrants were coded as reasons that justify or support a conclusion or claim. According to Fisher, common words or phrases that often signal conclusions in arguments are: Therefore...;

So...; Hence... etc. Within Toulmin's model conclusions are referred to as claims. Having identified a claim, it was further divided into correct claims and incorrect claims, based on the legal interpretation of the scenario.

### Frequency of Argument Model Codes

Table 1 presents the mean frequency of the codes in each condition.

-----

Insert Table 1 about here

-----

**Claims.** Claims are the decisions made that resolves the case. An example of a claim is: “so the employer should pay notice in lieu, termination benefit and balance of annual leave”. An overall measure of accuracy was calculated by dividing the number of correct claims by the total number of claims made, both correct and incorrect, to assess the proportion of correct claims made. Expert decision makers made a significantly higher proportion of correct claims than novices  $F(1,46) = 6.26, p = .02, \eta_p^2 = .12$ . Decision makers using the decision aid made a significantly higher proportion of correct claims than those without the aid  $F(1,46) = 46.32, p = .001, \eta_p^2 = .50$ . There was no significant interaction between level of experience and use of the aid on the proportion of correct claims made  $F(1,46) = 1.83, p = .18, \eta_p^2 = .04$ . These findings support our predictions that expert decision makers make more accurate decisions than novices, and that the decision aid increased the accuracy of decisions. However, unexpectedly, there was no interaction of experience with use of the aid. That is, the decision aid did not solely improve accuracy of novice decision makers; it improved the expert decision makers' accuracy too.

**Warrants.** Warrants refer to the justifications given for a decision. An example of a warrant is: “because the worker was not given the chance to defend, they conducted a verbal domestic

inquiry and no evidence was submitted to court”. Expert decision makers used significantly more warrants than novices  $F(1,46) = 11.48, p = .001, \eta_p^2 = .20$ . Decision makers using the decision aid used significantly more warrants than decision makers without the aid  $F(1,46) = 134.58, p = .001, \eta_p^2 = .75$ . There was no significant interaction between level of experience and use of the aid on the frequency of warrants  $F(1,46) = 0.16, p = .70, \eta_p^2 = .003$ . As predicted, expert decision makers refer to more warrants than novices, and the decision aid increases the use of warrants in decision making.

Not only were warrants more frequent with experience and when using the aid, but the variety of warrants increased too. The mean number of different types of warrant used in each of the cases without the aid was 1.00 ( $SD = 1.22$ ) for novices and 4.17 ( $SD = 3.37$ ) for experts. The mean number of different types of warrants used in each of the cases with the aid was 13.33 ( $SD = 3.56$ ) for novices and 13.00 ( $SD = 1.90$ ) for experts.

**Backing.** Backing statements refer to general statements that reinforce the warrant, in this case references to the part of the Employment Act that supports the warrant. An example of a backing statement is: “according to Regulations 6 (Termination and Layoff Benefits, 1980) if the worker worked for more than 12 months continuously, he/she is eligible for termination benefit”. There was no significant difference in the frequency that expert decision makers made backing statements compared to novices  $F(1,46) = 1.53, p = .22, \eta_p^2 = .03$ . Decision makers using the decision aid made significantly more backing statements than decision makers without the aid  $F(1,46) = 4.10, p = .05, \eta_p^2 = .08$ . There was no significant interaction between level of experience and use of the aid on the frequency of backing statements  $F(1,46) = .06, p = .81, \eta_p^2 = .001$ . Our hypothesis that expert decision makers would make more backing statements was not supported, but the decision aid did increase the use of backing statements as we predicted.

**Data.** Data refer to the evidence from the scenario that is repeated in the protocol. An example of a statement coded as data is: “the worker is a production operator with a monthly salary of RM1300”. There was no significant difference in the frequency that expert decision

makers referred to data compared to novices  $F(1,46) = 0.57, p = .46, \eta_p^2 = .01$ . Decision makers using the decision aid referred to less data than decision makers without the aid  $F(1,46)=8.43, p = .01, \eta_p^2 = .16$ . There was no significant interaction between level of experience and use of the aid on frequency of referring to data  $F(1,46) = 0.30, p = .59, \eta_p^2 = .006$ . As predicted, expert decision makers did not use more data in their decisions than novice decision makers. However, opposite to our prediction, the decision aid reduced the use of data.

**Correlations.** Table 2 presents the correlations between the frequencies of each code. The frequency of data correlated positively with frequency of backing statement. That is, participants who mentioned the evidence more also mentioned the act more. However, neither of these factors was associated with the proportion of correct conclusions drawn. The largest correlation was between the frequency of warrants and the proportion of correct conclusions. That is, those participants who gave more reasons and justifications for their decisions also made more accurate decisions.

-----

Insert Table 2 about here

-----

## Discussion

The aim this study was to explore how analytic naturalistic decisions are made. The domain we studied was resolving labour disputes. We found that decisions made by more expert labour officers were more accurate than novice officers and that they were more likely to use warrants, that is justifications, during the decision process. The number of justifications considered correlated positively with decision accuracy. Our second aim was to develop and test a decision aid. We found that those using the decision aid made more accurate decisions and that they were

more likely to use warrants and backing statements during the decision process. These results demonstrate that, in this domain, making decisions by considering arguments that link evidence to conclusion with explicit reasoning is characteristic of accurate, expert decision making.

### **Naturalistic Decision Making Through Argumentation**

The main analytic decision making process proposed in NDM models involves constructing a story of the situation to explain and simulate the consequences of decisions. The contribution of this study to NDM is the finding that explicit argumentation is valuable when it is necessary to demonstrate the accuracy of a decision and to explicitly resolve inconsistent and uncertain information. Experts use argumentation-based decision making more than novices and using arguments is associated with increased accuracy of decision making. Whilst argumentation has been suggested elsewhere in the NDM literature (e.g. Lipshitz, 1993), there has been little empirical research demonstrating its role in expertise and accurate decision making.

Furthermore, the study indicates how arguments are used. Whilst an increase in warrants was associated with good performance, there were no rebuttals or qualifiers. Rebuttals and qualifiers are not an essential part of an argument, according to Toulmin, but they do indicate a certain type of argument. Rebuttals are found when competing arguments are used to undermine a conclusion. They reflect a use of an alternative perspective to challenge a decision. Qualifiers indicate how a conclusion can be moderated. Together, these argument components would suggest a sophisticated argument process in which the decision maker weighs alternative arguments, leading to a decision based on an evaluation of different perspectives. However, this is not what we found. We found that warrants were used to support a conclusion, but contrasting arguments were not used to rebut it. This suggests an argument process in which reasons are used to support a position but not to challenge it. This may imply a pattern of myside bias in which people generate more arguments in favour of their position than opposing it (Baron, 1995), but the accuracy of the decisions suggests there is no bias. Instead, experts

seemingly focus on the most valid arguments that support the correct decision without explicitly weighing alternatives.

Not only were more warrants used by experts, but the range of warrants used in each scenario was greater. It was not the case that experts simply said the same things more frequently. As warrants have emerged as an important part of the argument-based decision process, a limitation of this study is that different types of warrants were not coded separately and future research could explore them more to find how they differ. Green and McCloy (2003) found that the strength of arguments influenced the conclusions drawn by naïve participants. It is likely that this would also be the case for experts and so it would be informative to discover how argument strength is assessed and how this relates to the type of warrant.

The importance of warrants to justify a decision is not predicted by story building accounts. These theories do not predict that explicit reasons to link events to a conclusion are central to accurate, expert decisions. Instead they rely on implicit judgements of coherence of a chronological explanation of events. The difference between the argumentation-based account proposed here and the story building accounts lies in the explicit use of reasons to argue for a conclusion. This distinction is similar to the contrast of relational and temporal structures proposed by Schum.

However, at a more fundamental level, there are similarities between these decision processes. Both theories seek to develop coherent representations. That is, representations are developed that minimise the number of contradictions and inconsistencies between elements in the representations. For example, in a coherent account all the facts are explained or all the contradictory arguments are resolved. Coherence has been proposed as a basis for naturalistic decision making (Cropp, Banks, Eghlai, 2011; McAndrew, Gore & Banks, 2009) based on the work of Thagard (1989) but these models do not distinguish between different forms of coherent representation. We suggest that both story building and argumentation-based decision making are based on different forms of representation but the aim of both is to be coherent.

The nature of the decision made may determine which form of representation is used. A chronological story represents what has happened. It may be that this form of representation is best suited to deciding what happens next, i.e. continuing the chronological account into the future. This inference is a mental simulation, a decision process that is common in NDM domains where decisions are made about the consequences of an action in concrete, physical environment such as those faced by fireground commanders. Argumentation is better suited to decisions where there is a need to persuade others (Mercier & Sperber, 2011). Reasons for the conclusion are presented to demonstrate explicitly that the decision is correct. Argumentation is also better suited to explicitly resolving uncertainties or contrasting alternative possibilities (e.g. Cohen et al. 1996; Schum, 1993). Therefore we propose that argumentation-based decision making will be more common in domains with these characteristics, generalising these findings beyond the domain of resolving labour disputes.

### **Argumentation-based Decision Aid**

The practical application of the argument model was demonstrated through the development of the decision aid. This increased the accuracy of decision making by both novices and experts. It was developed through an open-ended focus group discussion in which a consensus was reached by SMEs about the most important questions and information to assess when resolving a labour dispute. These questions were subsequently mapped onto Toulmin's argument model to assess if they prompted components of an argument model. There was a reasonable correspondence between them with the decision aid prompting the decision maker to assess data, warrants, backing, and rebuttals. This suggests that SMEs advocate argumentation as part of the decision making process.

The most effective aspect of the decision aid was the requirement to consider warrants and backing. The frequency of these statements increased, as did the accuracy of the decisions. As warrants are associated with accurate decisions, this implies that encouraging decision makers to

develop explicit warrants emphasises the need for sound arguments and this leads to better decisions. Laboratory studies have shown requiring people to make reasons explicit emphasises analytic reasoning and improves its accuracy (Platt & Griggs, 1993) and the decision aid appears to have a similar effect. More surprising was the reduction in data statements given that the aid encourages decision makers to attend to them. It may be that decision aid encourages a more efficient processing of data so that each piece of evidence is not mentioned more frequently than is required. The lack of rebuttals is also surprising given the question encouraging decision makers to consider challenges to the claim. This suggests that the SMEs recognise the importance of this, but participants were not able implement the advice perhaps due to the cognitive complexity of holding competing arguments in mind simultaneously.

Whilst we did not conduct a formal evaluation of the subjective experience of using the decision aid, anecdotal evidence suggests that it was well received by the users. Several novice participants copied the decision aid for later use in their work and several expert participants suggested that it would be a valuable tool for training. Further research on this or similar decision aids would be necessary to assess their usability. Further research would also be valuable to discover how best novices could be trained using a decision aid. Indeed, given the improved performance of both experts and novices when using the aid, it is possible that the aid should be used by all employees.

Given the simplicity of our method for developing the decision aid, it is striking how effective it was. Although the questions in the decision were specific to this context, the simple method could readily be applied to other domains through a focus group with SMEs that elicits the key questions and information that they use in making analytic decisions. These findings suggest that the key component of the argument model is the elaboration of warrants, that is, the reasoned justifications or rationale behind the decision. Therefore we expect that checklists in other domains would be comprised of questions which encourage the exploration and development of



reasons to support a decision. However, it is an empirical question which components of Toulmin's argument model, if any, are found in decision aids in different domains.

## **Conclusion**

To conclude, this study has investigated NDM in a domain with low time pressure and high novelty where there is no typical course of action. These circumstances are well suited to analytic decision making and are typical of many decision domains in business and organisational settings where analytic decision making is possible and potentially valuable. The study has shown that expert and accurate analytic decision making in this domain occurs through a process of argumentation rather than constructing a story of the situation. A key component of this is the explicit use of warrants or reasoned justifications for a conclusion rather than a story building process. A simple checklist based on the questions asked by SMEs mapped onto argument model, increased the frequency of warrants, and increased decision accuracy. This provides further support for the role of argumentation-based decision making in this domain. We suggest that argumentation-based decision making is more common because of the benefits of using explicit arguments to persuade others and demonstrate the accuracy of the decision and also the need to explicitly resolve inconsistent and uncertain information when making the decision.

## References

- Armenis, D. C., & Neal, A. (2008). Recognizing potential: A naturalistic investigation of employee promotion decisions. *Journal of Cognitive Engineering and Decision Making, 2*, 63-87. doi: 10.1518/155534308X284372
- Bacon, D. L., Fulton, B. J., & Malott, R. W. (1983). Improving staff performance through the use of task checklists. *Journal of Organizational Behavior Management, 4*, 17-25. doi: 10.1300/J075v04n03\_03
- Baron, J. (1995). Myside bias in thinking about abortion. *Thinking & Reasoning, 1*, 221-235. doi: 10.1080/13546789508256909
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum: Hillsdale, NJ.
- Cohen, M. S., Freeman, J. T., & Wolf, S. (1996). Metarecognition in time-stressed decision making: Recognizing, critiquing, and correcting. *Human Factors, 38*, 206-219. doi: 10.1177/001872089606380203
- Crandall, B., & Getchell-Reiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Advances in Nursing Science, 16*, 42-51.
- Cropp, N., Banks, A., & Elghali, L. (2011). Expert Decision Making in a Complex Engineering Environment: A Comparison of the Lens Model, Explanatory Coherence, and Matching

Heuristics. *Journal of Cognitive Engineering and Decision Making*, 5, 255-276. doi:  
10.1177/1555343411415795

Doll, J., Livesey, J., McHaffie, E., & Ludwig, T. D. (2007). Keeping an uphill edge: Managing cleaning behaviors at a ski shop. *Journal of Organizational Behavior Management*, 27, 41-60. doi:  
10.1300/J075v27n03\_04

Edwards, W. (1961). Behavioral decision theory. *Annual Review of Psychology*, 12, 473-498.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88, 915-933. doi: 10.1002/sce.20012

Ericsson, K. A. (2002). Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. *Applied Cognitive Psychology*, 16, 981-987. doi:  
10.1002/acp.925

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251. doi: 10.1037/0033-295X.87.3.215

Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data (2<sup>nd</sup> rev. ed.)  
MIT Press: Cambridge, MA.

Fisher, A. (2004). *The logic of real arguments*. Cambridge: Cambridge University Press.

Gawande, A. (2010). *The checklist manifesto: How to get things right*. New York: Metropolitan Books.

Gore, J., Banks, A., Millward, L.J., & Kyriakidou, O. (2006). Naturalistic Decision Making and organisations: Reviewing pragmatic science. *Organization Studies*, 27, 925-942. doi: 10.1177/0170840606065701

Green, D. W., & McCloy, R. (2003). Reaching a verdict. *Thinking and Reasoning*, 9, 307-333. doi:10.1080/1354678034000268

Hitchcock, D., & Verheij, B. (Eds.). (2007). *Arguing on the Toulmin model: New essays in argument analysis and evaluation (Vol. 10)*. Dordrecht, The Netherlands: Springer.

Kaempf, G. L., Klein, G., Thordsen, M. L., & Wolf, S. (1996). Decision making in complex naval command-and-control environments. *Human Factors*, 38, 220-231. doi: 10.1177/001872089606380204

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *American Psychologist*, 64, 515-526. doi: 10.1037/a0016755

Klein, G. (2008). Naturalistic decision making. *Human Factors*, 50, 456-460. doi: 10.1518/001872008X288385

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986, September). Rapid decision making on the fire ground. In *Proceedings of the Human Factors and Ergonomics Society annual meeting*, 30, 576-580. doi: 10.1177/154193128603000616

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, 4, 186-209. doi: 10.1518/155534310X12844000801203

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

Lipshitz, R. (1993). Decision making as argument-driven action. In G.A. Klein, J. Orasanu, R. Calderwood, & C.E. Zsombok (Eds.). *Decision making in action: Models and methods* (pp. 172-181). Norwood: Ablex.

Lipshitz, R. (1994). Decision making in three modes. *Journal for the Theory of Social Behaviour*, 24, 47-65. doi: 10.1111/j.1468-5914.1994.tb00246

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, 14, 331-352. doi: 10.1002/bdm.381

Lipshitz, R., & Shulimovitz, N. (2007). Intuition and emotion in bank loan officers' credit decisions. *Journal of Cognitive Engineering and Decision Making*, 1, 212-233. doi: 10.1518/155534307X232857

McAndrew, C., & Gore, J. (2013). Understanding Preferences in Experience-Based Choice: A Study of Cognition in the “Wild”. *Journal of Cognitive Engineering and Decision Making*, 7, 179-197. doi: 10.1177/1555343412463922

- McAndrew, C., Gore, J., & Banks, A. (2009). Convince Me: Modeling Naturalistic Decision Making. *Journal of Cognitive Engineering and Decision Making*, *3*, 156-175. doi: 10.1518/155534309X441862
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, *34*, 57-74. doi: 10.1017/S0140525X10000968
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259. doi: 10.1037/0033-295X.84.3.231
- Orasnu, J., & Connolly, T. (1993). The reinvention of decision making. In G.A. Klein, J. Orasnu, R. Calderwood, & C.E. Zsombok (Eds.). *Decision making in action: Models and methods* (pp. 3-20). Norwood: Ablex.
- Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, *51*, 242-258. doi: 10.1037/0022-3514.51.2.242
- Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 521-533. doi: 10.1037/0278-7393.14.3.521
- Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology*, *62*, 189-206. doi: 10.1037/0022-3514.62.2.189

Platt, R. D., & Griggs, R. A. (1993). Facilitation in the abstract selection task: The effects of attentional and instructional factors. *The Quarterly Journal of Experimental Psychology*, *46*, 591-613. doi: 10.1080/14640749308401029

Prakken, H. (2005). AI & Law, logic and argument schemes. *Argumentation*, *19*, 303-320. doi: 10.1007/s10503-005-4418-7

Roth, E. M., Pfautz, J. D., Mahoney, S. M., Powell, G. M., Carlson, E. C., Guarino, S. L., Fichtl, T.C., & Potter, S. S. (2010). Framing and Contextualizing Information Requests: Problem Formulation as Part of the Intelligence Analysis Process. *Journal of Cognitive Engineering and Decision Making*, *4*, 210-239. doi: 10.1518/155534310X12844000801087

Salas, E., & Klein, G. A. (Eds.). (2001). *Linking expertise and naturalistic decision making*. Lawrence Erlbaum: Mahwah, NJ.

Schum, D.A. (1993). Argument structuring and evidence evaluation. In R. Hastie (Ed.). *Inside the juror: The psychology of juror decision making*. (pp. 175-191). Cambridge: Cambridge University Press.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11-36. doi: 10.1016/0010-0277(93)90034-S

Shanteau, J. (1992a). Competence in experts: The role of task characteristics. *Organizational Behavior & Human Decision Processes*, *53*, 252-266. doi: 10.1016/0749-5978(92)90064-E

Shanteau, J. (1992b). How much information does an expert use? Is it relevant? *Acta Psychologica*, 81, 75-86. doi: 10.1016/0001-6918(92)90012-3

Termination and Layoff Benefits (1980). In *Employment Act 1955 (ACT 265) & Regulations*, International Law Book Services (2004).

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, 12, 435-467. doi: 10.1017/S0140525X00057046

Toulmin, S. E. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Weinstock, M. (2011). Knowledge-telling and knowledge-transforming arguments in mock jurors' verdict justifications. *Thinking & Reasoning*, 17, 282-314. doi: 10.1080/13546783.2011.575191



Table 1

Mean Frequency of the Argument Model Codes in each Condition (standard deviations in parentheses)

	Experienced Decision Makers		Novice Decision Makers	
	No Decision Aid	Decision Aid	No Decision Aid	Decision Aid
Data	25.27 (14.04)	16.50 (9.89)	30.09 (13.17)	17.29 (14.79)
Warrant	3.09 (1.87)	13.29 (3.45)	0.55 (1.04)	10.07 (4.01)
Backing	11.45(5.59)	8.86 (5.32)	10.00 (5.10)	6.71 (4.43)
Proportion	0.83 (0.13)	0.99 (0.03)	0.72 (0.16)	0.96 (0.07)
Correct Claims				

Table 2

Correlations between the Frequencies of Argument Codes

	Data	Warrant	Backing
Data	-		
Warrants	-.15	-	
Backing	.29*	-.08	-
Proportion Correct	-.20	.66**	-.03
Claims			

\* $p < .05$ , \*\* $p < .001$

Figure 1: A basic story for a labour dispute illustrating the data component of Toulmin's argument model

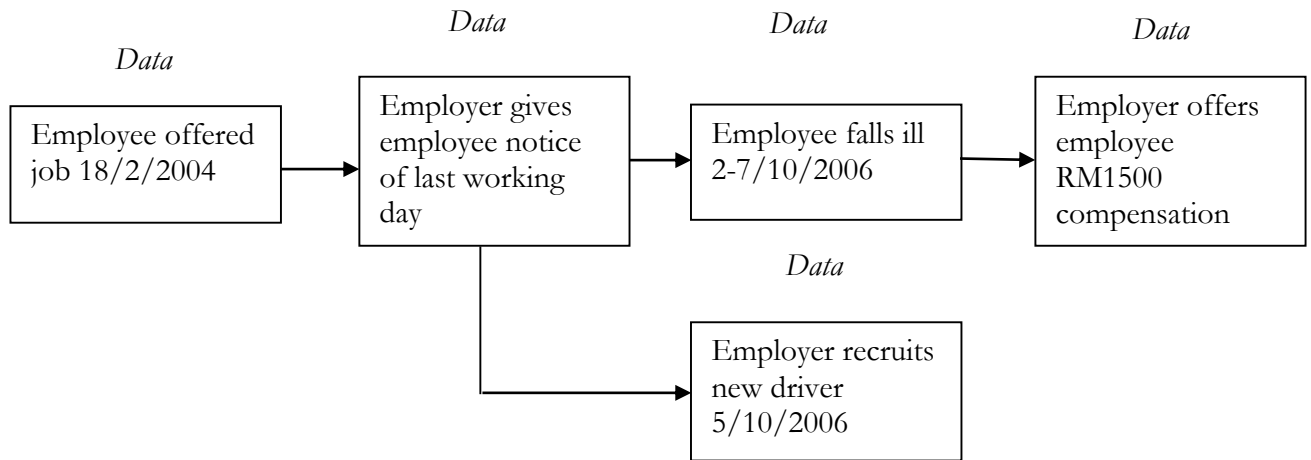


Figure 2: A basic pattern of argument analysis for a labour dispute illustrating all components of Toulmin's argument model

