Theory of instrumental rationality

Decision theory = theory about motivated and rational choices.
It answers the following question: given one's preferences, the risks and the choice set of a situation, what decision is it rational to take?

Rational decision involves evaluating the stakes: the payoffs and the risks.
Examples: Darwin considering marrying; C. Columbus and the risk of sailing west; undergoing or not heart surgery.

A decision problem includes:
- states
- outcomes
- acts (or possible choices)

It can be expressed in a table of the type:

Deciding about heart surgery

	Group I: 7.1%	Group II: 7.8%	Group III: 13.9%	Group IV: 71.2%
Operation	0.05 years	2.1 years	3.9 years	14.8 years
No operation	1.5 years	1.5 years	1.5 years	1.5 years

Each outcome is more or less desirable. This is expressed by ordering outcomes.
The ordering can be done on an ordinal scale or on a cardinal scale.

Some rules of rational decision making when risks are unknown:
- Maximin and leximin rule: consider the worst outcomes for each state, pick the action that correspond to the best of these worst outcomes.
- Optimism-pessimism rule: weight the best and the worst outcomes in view of your optimistic or pessimistic assumptions. Multiply the worst and the best outcomes by p (a measure of your pessimism that is <1) and 1-p respectively. Pick the action with the highest number.
- Minimax regret: calculate the regret for each outcome: it is equal to the payoff of the given outcome minus the payoff of the best possible outcome in the same state (i.e. check across possible choices; it is the highest number in the column). Apply the maximin rule to the result.

- Maximise expected utility: pick the action with the maximum expected utility.

Expected utility of an action: consider all the possible outcomes that an action could lead to. Evaluate their respective utilities and multiply each of them by the probability that the outcome actually obtains (it is the probability of the corresponding state). The sum of these is the expected utility of the action.
Example considering only two possible states: S and not S
The expected utility of an action A given uncertainty about a state S
= Probability(S)*Utility(S|A) + Probability(not S)Utility(not S|A)

Axiomatisation:
=> What assumptions are made when assuming rational decision-making?

- Preferences are supposed to apply to all possible outcomes (completeness: between any two outcomes, one is either better, worst or equivalent to the other). The consequent ordering is asymetric (if x>y, then not y>x) and transitive.
- The axioms for cardinal scales for preferences, required for rational choice under risk, are more demanding. They include relations of preferences among **lotteries**.

Most of the time, utility functions for given items (esp. money) have decreasing marginal utility. People are consequently risk averse. Numerous experiments show that people are risk seeking when actions might lead to losses: it is called loss aversion (the disutility of loosing something is bigger than the utility of winning this same thing).

Rational choice theory in the social sciences (c.f. Levitt and Dubner): understanding human behaviour goes through an analysis of people’s incentives.

Can decision theory be a useful tool for cognitive psychology?

We cannot say that decision theory provides an adequate description of human cognition. Evidence includes all the work in behavioural economics specifying the bounds of rationality, esp. the research on ‘heuristics and biases’ (Twersky and Kanheman).
Several attitudes are possible:

- The theory of rational decision-making is nonetheless a good ‘as if’ model. Most of the time, social agents behave as if they were rationally maximising their material pay-off. This non-psychological assumption is good enough for economics (adopted by many economists and by social scientists of the ‘methodological individualism’ school).

- The theory of rational decision-making is still the best description of human decision-making, but needs to be amended on several issues. In particular, information is costly to acquire and process (Gintis; to some extent Kahneman with dual process theory).
- A fruitful zero-hypothesis: theories of rational decision-making provide very precise and predictive zero-hypotheses. Furthermore, both common sense and evolutionary arguments justify assuming by default that cognitive processes comply with some rationality criteria.
- A functional description of cognitive mechanisms or strategies: what is it that people maximise?
- A fair assumption in controlled experiments that are meant to reveal preferences. We assume the following causal chain from thoughts to behaviour:

  motives
  + instrumental rationality
  + (adequate beliefs of) stakes/choice set
  => actual choice

  Thus, if you know the actual choice, the choice set and assume rationality, then you can infer what the motives are.

Predictable irrationality

Three classic cases of preference reversal:

- Twersky and Kahneman (1981): framing an option in terms of lives saved or risk of death leads to preference reversal.

  600 people affected by a deadly disease:

  - option A ~~saves 200 people's lives~~ 400 will die
  - option B has a 33% chance ~~of saving all 600 people~~ that no people will die and a 66% possibility ~~of saving no one~~ that all 600 will die

  ~~72%~~ 22% of participants chose option A
  ~~28%~~ 78% of participants chose option B.

- Allais paradox: the insertion of an independent ‘gamble’ leads to preference reversal. In Allais’ paradox gambles 1 and 2 include a .88 probability of winning 1 million while gambles 3 and 4 do not include this probability. But apart from this independent gamble, 1 is similar to 3 and 2 is similar to 4. People prefering 1 should therefore prefer 3. But this is not what happens.

	Ticket no. 1	Ticket no. 2–11	Ticket no. 12–100
Gamble 1	$1M	$1M	$1M
Gamble 2	$0	$5M	$1M
Gamble 3	$1M	$1M	$0
Gamble 4	$0	$5M	$0

- Ellsberg paradox: people show ambiguity aversion, i.e. they prefer a bet where they know the probabilities of winning. Thus they prefer gamble 1 to gamble 2, because they don't know the number of black balls and they prefer gamble 4 to gamble 3. However, if we assume that people' choice are formed after forming proabilistic beliefs, then choosing gamble 1 express the belief that there are most probably less than 30 black balls. But with this belief, they should choose gamble 3 rather than gamble 4. Which is not what happens.

	30	60	
	Red	Black	Yellow
Gamble 1	$100	$0	$0
Gamble 2	$0	$100	$0
Gamble 3	$100	$0	$100
Gamble 4	$0	$100	$100