

.....

Twelve Virtues of Rationality

<http://yudkowsky.net/virtues/>
©2006 by Eliezer S. Yudkowsky

*This material may be freely reproduced in unaltered form
(including the copyright and this notice).*

.....

Twelve Virtues of Rationality

<http://yudkowsky.net/virtues/>
©2006 by Eliezer S. Yudkowsky

*This material may be freely reproduced in unaltered form
(including the copyright and this notice).*

The first virtue is curiosity. A burning itch to know is higher than a solemn vow to pursue truth. To feel the burning itch of curiosity requires both that you be ignorant, and that you desire to relinquish your ignorance. If in your heart you believe you already know, or if in your heart you do not wish to know, then your questioning will be purposeless and your skills without direction. Curiosity seeks to annihilate itself; there is no curiosity that does not want an answer. The glory of glorious mystery is to be solved, after which it ceases to be mystery. Be wary of those who speak of being open-minded and modestly confess their ignorance. There is a time to confess your ignorance and a time to relinquish your ignorance.

The first virtue is curiosity. A burning itch to know is higher than a solemn vow to pursue truth. To feel the burning itch of curiosity requires both that you be ignorant, and that you desire to relinquish your ignorance. If in your heart you believe you already know, or if in your heart you do not wish to know, then your questioning will be purposeless and your skills without direction. Curiosity seeks to annihilate itself; there is no curiosity that does not want an answer. The glory of glorious mystery is to be solved, after which it ceases to be mystery. Be wary of those who speak of being open-minded and modestly confess their ignorance. There is a time to confess your ignorance and a time to relinquish your ignorance.

territory" or "experience of success and failure" or "Bayesian decision theory". But perhaps you describe incorrectly the nameless virtue. How will you discover your mistake? Not by comparing your description to itself, but by comparing it to that which you did not name.

If for many years you practice the techniques and submit yourself to strict constraints, it may be that you will glimpse the center. Then you will see how all techniques are one technique, and you will move correctly without feeling constrained. Musashi wrote: "When you appreciate the power of nature, knowing the rhythm of any situation, you will be able to hit the enemy naturally and strike naturally. All this is the Way of the Void."

These then are twelve virtues of rationality:

Curiosity, relinquishment, lightness, evenness, argument, empiricism, simplicity, humility, perfectionism, precision, scholarship, and the void.

territory" or "experience of success and failure" or "Bayesian decision theory". But perhaps you describe incorrectly the nameless virtue. How will you discover your mistake? Not by comparing your description to itself, but by comparing it to that which you did not name.

If for many years you practice the techniques and submit yourself to strict constraints, it may be that you will glimpse the center. Then you will see how all techniques are one technique, and you will move correctly without feeling constrained. Musashi wrote: "When you appreciate the power of nature, knowing the rhythm of any situation, you will be able to hit the enemy naturally and strike naturally. All this is the Way of the Void."

These then are twelve virtues of rationality:

Curiosity, relinquishment, lightness, evenness, argument, empiricism, simplicity, humility, perfectionism, precision, scholarship, and the void.

If you fail to achieve a correct answer, it is futile to protest that you acted with propriety.

How can you improve your conception of rationality? Not by saying to yourself, "It is my duty to be rational." By this you only enshrine your mistaken conception. Perhaps your conception of rationality is that it is rational to believe the words of the Great Teacher, and the Great Teacher says, "The sky is green," and you look up at the sky and see blue. If you think: "It may look like the sky is blue, but rationality is to believe the words of the Great Teacher," you lose a chance to discover your mistake.

Do not ask whether it is "the Way" to do this or that. Ask whether the sky is blue or green. If you speak overmuch of the Way you will not attain it.

You may try to name the highest principle with names such as "the map that reflects the

If you fail to achieve a correct answer, it is futile to protest that you acted with propriety.

How can you improve your conception of rationality? Not by saying to yourself, "It is my duty to be rational." By this you only enshrine your mistaken conception. Perhaps your conception of rationality is that it is rational to believe the words of the Great Teacher, and the Great Teacher says, "The sky is green," and you look up at the sky and see blue. If you think: "It may look like the sky is blue, but rationality is to believe the words of the Great Teacher," you lose a chance to discover your mistake.

Do not ask whether it is "the Way" to do this or that. Ask whether the sky is blue or green. If you speak overmuch of the Way you will not attain it.

You may try to name the highest principle with names such as "the map that reflects the

The second virtue is relinquishment. P. C. Hodgell said: "That which can be destroyed by the truth should be." Do not flinch from experiences that might destroy your beliefs. The thought you cannot think controls you more than thoughts you speak aloud. Submit yourself to ordeals and test yourself in fire. Relinquish the emotion which rests upon a mistaken belief, and seek to feel fully that emotion which fits the facts. If the iron approaches your face, and you believe it is hot, and it is cool, the Way opposes your fear. If the iron approaches your face, and you believe it is cool, and it is hot, the Way opposes your calm. Evaluate your beliefs first and then arrive at your emotions. Let yourself say: "If the iron is hot, I desire to believe it is hot, and if it is cool, I desire to believe it is cool." Beware lest you become attached to beliefs you may not want.

The second virtue is relinquishment. P. C. Hodgell said: "That which can be destroyed by the truth should be." Do not flinch from experiences that might destroy your beliefs. The thought you cannot think controls you more than thoughts you speak aloud. Submit yourself to ordeals and test yourself in fire. Relinquish the emotion which rests upon a mistaken belief, and seek to feel fully that emotion which fits the facts. If the iron approaches your face, and you believe it is hot, and it is cool, the Way opposes your fear. If the iron approaches your face, and you believe it is cool, and it is hot, the Way opposes your calm. Evaluate your beliefs first and then arrive at your emotions. Let yourself say: "If the iron is hot, I desire to believe it is hot, and if it is cool, I desire to believe it is cool." Beware lest you become attached to beliefs you may not want.

The third virtue is lightness. Let the winds of evidence blow you about as though you are a leaf, with no direction of your own. Beware lest you fight a rearguard retreat against the evidence, grudgingly conceding each foot of ground only when forced, feeling cheated. Surrender to the truth as quickly as you can. Do this the instant you realize what you are resisting; the instant you can see from which quarter the winds of evidence are blowing against you. Be faithless to your cause and betray it to a stronger enemy. If you regard evidence as a constraint and seek to free yourself, you sell yourself into the chains of your whims. For you cannot make a true map of a city by sitting in your bedroom with your eyes shut and drawing lines upon paper according to impulse. You must walk through the city and draw lines on paper that correspond to what you see. If, seeing the city unclearly, you think that you can shift a line just a little to the right, just a little to the left, according to your caprice, this is just the same mistake.

The third virtue is lightness. Let the winds of evidence blow you about as though you are a leaf, with no direction of your own. Beware lest you fight a rearguard retreat against the evidence, grudgingly conceding each foot of ground only when forced, feeling cheated. Surrender to the truth as quickly as you can. Do this the instant you realize what you are resisting; the instant you can see from which quarter the winds of evidence are blowing against you. Be faithless to your cause and betray it to a stronger enemy. If you regard evidence as a constraint and seek to free yourself, you sell yourself into the chains of your whims. For you cannot make a true map of a city by sitting in your bedroom with your eyes shut and drawing lines upon paper according to impulse. You must walk through the city and draw lines on paper that correspond to what you see. If, seeing the city unclearly, you think that you can shift a line just a little to the right, just a little to the left, according to your caprice, this is just the same mistake.

Before these eleven virtues is a virtue which is nameless.

Miyamoto Musashi wrote, in *The Book of Five Rings*:

"The primary thing when you take a sword in your hands is your intention to cut the enemy, whatever the means. Whenever you parry, hit, spring, strike or touch the enemy's cutting sword, you must cut the enemy in the same movement. It is essential to attain this. If you think only of hitting, springing, striking or touching the enemy, you will not be able actually to cut him. More than anything, you must be thinking of carrying your movement through to cutting him."

Every step of your reasoning must cut through to the correct answer in the same movement. More than anything, you must think of carrying your map through to reflecting the territory.

Before these eleven virtues is a virtue which is nameless.

Miyamoto Musashi wrote, in *The Book of Five Rings*:

"The primary thing when you take a sword in your hands is your intention to cut the enemy, whatever the means. Whenever you parry, hit, spring, strike or touch the enemy's cutting sword, you must cut the enemy in the same movement. It is essential to attain this. If you think only of hitting, springing, striking or touching the enemy, you will not be able actually to cut him. More than anything, you must be thinking of carrying your movement through to cutting him."

Every step of your reasoning must cut through to the correct answer in the same movement. More than anything, you must think of carrying your map through to reflecting the territory.

The eleventh virtue is scholarship. Study many sciences and absorb their power as your own. Each field that you consume makes you larger. If you swallow enough sciences the gaps between them will diminish and your knowledge will become a unified whole. If you are gluttonous you will become vaster than mountains. It is especially important to eat math and science which impinges upon rationality: Evolutionary psychology, heuristics and biases, social psychology, probability theory, decision theory. But these cannot be the only fields you study. The Art must have a purpose other than itself, or it collapses into infinite recursion.

The eleventh virtue is scholarship. Study many sciences and absorb their power as your own. Each field that you consume makes you larger. If you swallow enough sciences the gaps between them will diminish and your knowledge will become a unified whole. If you are gluttonous you will become vaster than mountains. It is especially important to eat math and science which impinges upon rationality: Evolutionary psychology, heuristics and biases, social psychology, probability theory, decision theory. But these cannot be the only fields you study. The Art must have a purpose other than itself, or it collapses into infinite recursion.

The fourth virtue is evenness. One who wishes to believe says, "Does the evidence permit me to believe?" One who wishes to disbelieve asks, "Does the evidence force me to believe?" Beware lest you place huge burdens of proof only on propositions you dislike, and then defend yourself by saying: "But it is good to be skeptical." If you attend only to favorable evidence, picking and choosing from your gathered data, then the more data you gather, the less you know. If you are selective about which arguments you inspect for flaws, or how hard you inspect for flaws, then every flaw you learn how to detect makes you that much stupider. If you first write at the bottom of a sheet of paper, "And therefore, the sky is green!", it does not matter what arguments you write above it afterward; the conclusion is already written, and it is already correct or already wrong. To be clever in argument is not rationality but rationalization. Intelligence, to be useful, must be used for something other than defeating itself. Listen to hypotheses as they plead their cases before you, but remember that you are not a hypothesis, you are the judge. Therefore do not seek to argue for one side or another, for if you knew your destination, you would already be there.

The fourth virtue is evenness. One who wishes to believe says, "Does the evidence permit me to believe?" One who wishes to disbelieve asks, "Does the evidence force me to believe?" Beware lest you place huge burdens of proof only on propositions you dislike, and then defend yourself by saying: "But it is good to be skeptical." If you attend only to favorable evidence, picking and choosing from your gathered data, then the more data you gather, the less you know. If you are selective about which arguments you inspect for flaws, or how hard you inspect for flaws, then every flaw you learn how to detect makes you that much stupider. If you first write at the bottom of a sheet of paper, "And therefore, the sky is green!", it does not matter what arguments you write above it afterward; the conclusion is already written, and it is already correct or already wrong. To be clever in argument is not rationality but rationalization. Intelligence, to be useful, must be used for something other than defeating itself. Listen to hypotheses as they plead their cases before you, but remember that you are not a hypothesis, you are the judge. Therefore do not seek to argue for one side or another, for if you knew your destination, you would already be there.

The fifth virtue is argument. Those who wish to fail must first prevent their friends from helping them. Those who smile wisely and say: "I will not argue," remove themselves from help, and withdraw from the communal effort. In argument strive for exact honesty, for the sake of others and also yourself: The part of yourself that distorts what you say to others also distorts your own thoughts. Do not believe you do others a favor if you accept their arguments; the favor is to you. Do not think that fairness to all sides means balancing yourself evenly between positions; truth is not handed out in equal portions before the start of a debate. You cannot move forward on factual questions by fighting with fists or insults. Seek a test that lets reality judge between you.

The fifth virtue is argument. Those who wish to fail must first prevent their friends from helping them. Those who smile wisely and say: "I will not argue," remove themselves from help, and withdraw from the communal effort. In argument strive for exact honesty, for the sake of others and also yourself: The part of yourself that distorts what you say to others also distorts your own thoughts. Do not believe you do others a favor if you accept their arguments; the favor is to you. Do not think that fairness to all sides means balancing yourself evenly between positions; truth is not handed out in equal portions before the start of a debate. You cannot move forward on factual questions by fighting with fists or insults. Seek a test that lets reality judge between you.

The tenth virtue is precision. One comes and says: The quantity is between 1 and 100. Another says: the quantity is between 40 and 50. If the quantity is 42 they are both correct, but the second prediction was more useful and exposed itself to a stricter test. What is true of one apple may not be true of another apple; thus more can be said about a single apple than about all the apples in the world. The narrowest statements slice deepest, the cutting edge of the blade. As with the map, so too with the art of mapmaking: The Way is a precise Art. Do not walk to the truth, but dance. On each and every step of that dance your foot comes down in exactly the right spot. Each piece of evidence shifts your beliefs by exactly the right amount, neither more nor less. What is exactly the right amount? To calculate this you must study probability theory. Even if you cannot do the math, knowing that the math exists tells you that the dance step is precise and has no room in it for your whims.

The tenth virtue is precision. One comes and says: The quantity is between 1 and 100. Another says: the quantity is between 40 and 50. If the quantity is 42 they are both correct, but the second prediction was more useful and exposed itself to a stricter test. What is true of one apple may not be true of another apple; thus more can be said about a single apple than about all the apples in the world. The narrowest statements slice deepest, the cutting edge of the blade. As with the map, so too with the art of mapmaking: The Way is a precise Art. Do not walk to the truth, but dance. On each and every step of that dance your foot comes down in exactly the right spot. Each piece of evidence shifts your beliefs by exactly the right amount, neither more nor less. What is exactly the right amount? To calculate this you must study probability theory. Even if you cannot do the math, knowing that the math exists tells you that the dance step is precise and has no room in it for your whims.

The ninth virtue is perfectionism. The more errors you correct in yourself, the more you notice. As your mind becomes more silent, you hear more noise. When you notice an error in yourself, this signals your readiness to seek advancement to the next level. If you tolerate the error rather than correcting it, you will not advance to the next level and you will not gain the skill to notice new errors. In every art, if you do not seek perfection you will halt before taking your first steps. If perfection is impossible that is no excuse for not trying. Hold yourself to the highest standard you can imagine, and look for one still higher. Do not be content with the answer that is almost right; seek one that is exactly right.

The ninth virtue is perfectionism. The more errors you correct in yourself, the more you notice. As your mind becomes more silent, you hear more noise. When you notice an error in yourself, this signals your readiness to seek advancement to the next level. If you tolerate the error rather than correcting it, you will not advance to the next level and you will not gain the skill to notice new errors. In every art, if you do not seek perfection you will halt before taking your first steps. If perfection is impossible that is no excuse for not trying. Hold yourself to the highest standard you can imagine, and look for one still higher. Do not be content with the answer that is almost right; seek one that is exactly right.

The sixth virtue is empiricism. The roots of knowledge are in observation and its fruit is prediction. What tree grows without roots? What tree nourishes us without fruit? If a tree falls in a forest and no one hears it, does it make a sound? One says, "Yes it does, for it makes vibrations in the air." Another says, "No it does not, for there is no auditory processing in any brain." Though they argue, one saying "Yes", and one saying "No", the two do not anticipate any different experience of the forest. Do not ask which beliefs to profess, but which experiences to anticipate. Always know which difference of experience you argue about. Do not let the argument wander and become about something else, such as someone's virtue as a rationalist. Jerry Cleaver said: "What does you in is not failure to apply some high-level, intricate, complicated technique. It's overlooking the basics. Not keeping your eye on the ball." Do not be blinded by words. When words are subtracted, anticipation remains.

The sixth virtue is empiricism. The roots of knowledge are in observation and its fruit is prediction. What tree grows without roots? What tree nourishes us without fruit? If a tree falls in a forest and no one hears it, does it make a sound? One says, "Yes it does, for it makes vibrations in the air." Another says, "No it does not, for there is no auditory processing in any brain." Though they argue, one saying "Yes", and one saying "No", the two do not anticipate any different experience of the forest. Do not ask which beliefs to profess, but which experiences to anticipate. Always know which difference of experience you argue about. Do not let the argument wander and become about something else, such as someone's virtue as a rationalist. Jerry Cleaver said: "What does you in is not failure to apply some high-level, intricate, complicated technique. It's overlooking the basics. Not keeping your eye on the ball." Do not be blinded by words. When words are subtracted, anticipation remains.

The seventh virtue is simplicity. Antoine de Saint-Exupéry said: "Perfection is achieved not when there is nothing left to add, but when there is nothing left to take away." Simplicity is virtuous in belief, design, planning, and justification. When you profess a huge belief with many details, each additional detail is another chance for the belief to be wrong. Each specification adds to your burden; if you can lighten your burden you must do so. There is no straw that lacks the power to break your back. Of artifacts it is said: The most reliable gear is the one that is designed out of the machine. Of plans: A tangled web breaks. A chain of a thousand links will arrive at a correct conclusion if every step is correct, but if one step is wrong it may carry you anywhere. In mathematics a mountain of good deeds cannot atone for a single sin. Therefore, be careful on every step.

The seventh virtue is simplicity. Antoine de Saint-Exupéry said: "Perfection is achieved not when there is nothing left to add, but when there is nothing left to take away." Simplicity is virtuous in belief, design, planning, and justification. When you profess a huge belief with many details, each additional detail is another chance for the belief to be wrong. Each specification adds to your burden; if you can lighten your burden you must do so. There is no straw that lacks the power to break your back. Of artifacts it is said: The most reliable gear is the one that is designed out of the machine. Of plans: A tangled web breaks. A chain of a thousand links will arrive at a correct conclusion if every step is correct, but if one step is wrong it may carry you anywhere. In mathematics a mountain of good deeds cannot atone for a single sin. Therefore, be careful on every step.

The eighth virtue is humility. To be humble is to take specific actions in anticipation of your own errors. To confess your fallibility and then do nothing about it is not humble; it is boasting of your modesty. Who are most humble? Those who most skillfully prepare for the deepest and most catastrophic errors in their own beliefs and plans. Because this world contains many whose grasp of rationality is abysmal, beginning students of rationality win arguments and acquire an exaggerated view of their own abilities. But it is useless to be superior: Life is not graded on a curve. The best physicist in ancient Greece could not calculate the path of a falling apple. There is no guarantee that adequacy is possible given your hardest effort; therefore spare no thought for whether others are doing worse. If you compare yourself to others you will not see the biases that all humans share. To be human is to make ten thousand errors. No one in this world achieves perfection.

The eighth virtue is humility. To be humble is to take specific actions in anticipation of your own errors. To confess your fallibility and then do nothing about it is not humble; it is boasting of your modesty. Who are most humble? Those who most skillfully prepare for the deepest and most catastrophic errors in their own beliefs and plans. Because this world contains many whose grasp of rationality is abysmal, beginning students of rationality win arguments and acquire an exaggerated view of their own abilities. But it is useless to be superior: Life is not graded on a curve. The best physicist in ancient Greece could not calculate the path of a falling apple. There is no guarantee that adequacy is possible given your hardest effort; therefore spare no thought for whether others are doing worse. If you compare yourself to others you will not see the biases that all humans share. To be human is to make ten thousand errors. No one in this world achieves perfection.