

# Deep Q-Life Happiness:

## Formalising Joy, Fulfilment, and the Way

Dr. Yves J. Hilpisch<sup>1</sup>

December 10, 2025 (work in progress)

### Abstract

This paper applies the Deep Q-Life framework to the question of happiness. We combine three lenses—empirical findings from happiness science, long-running debates in philosophy, and practical wisdom from contemplative traditions—and cast them into a shared reinforcement learning language. By treating happiness-related decisions as Markov decision processes (MDPs) with explicit states, actions, rewards, and discount factors, we clarify what different traditions optimise, how they weigh present versus future joy, and where common misunderstandings arise (for example between pleasure, excitement, contentment, and fulfilment). The result is a conceptual map and a set of simple tools that help readers design their own “way to happiness” as a learnable policy rather than a fixed destination.

---

<sup>1</sup>Get in touch: <https://linktr.ee/dyjh>. Web page <https://hilpisch.com>. Research, structuring, drafting, and visualizations were assisted by GPT 5.1 as a co-writing tool under human direction. This paper is for educational purposes only and does not constitute medical, psychological, financial, or legal advice.

## Contents

|           |   |           |
|-----------|---|-----------|
| <b>1</b>  | <b>Why Happiness, Why Deep Q-Life</b>                                 | <b>2</b>  |
| <b>2</b>  | <b>Language for Happiness: Joy, Pleasure, Contentment, Fulfilment</b> | <b>2</b>  |
| 2.1       | Working Definitions and Everyday Examples . . . . .                   | 3         |
| 2.2       | Hedonic and Eudaimonic Happiness . . . . .                            | 3         |
| 2.3       | A Simple Contentment Ratio . . . . .                                  | 4         |
| <b>3</b>  | <b>Evidence from Happiness Science</b>                                | <b>5</b>  |
| 3.1       | Measurement: From Life Satisfaction to Momentary Affect . . . . .     | 5         |
| 3.2       | Dynamics, Adaptation, and Set Points . . . . .                        | 6         |
| <b>4</b>  | <b>Perspectives from Philosophy and Religion</b>                      | <b>7</b>  |
| 4.1       | Classical and Modern Philosophy . . . . .                             | 7         |
| 4.2       | Buddhist and Contemplative Perspectives . . . . .                     | 7         |
| <b>5</b>  | <b>Deep Q-Life Happiness as an MDP</b>                                | <b>8</b>  |
| 5.1       | States: What Matters for Happiness . . . . .                          | 8         |
| 5.2       | Actions and Habits . . . . .  | 8         |
| 5.3       | Rewards, Value Functions, and Discounting . . . . .                   | 8         |
| <b>6</b>  | <b>Everyday Frames as Formal Models</b>                               | <b>9</b>  |
| 6.1       | Contrasting Frames: Achievement, Comfort, and Presence . . . . .      | 9         |
| <b>7</b>  | <b>Designing Your Own Way to Happiness</b>                            | <b>10</b> |
| 7.1       | Clarifying Signals and Language . . . . .                             | 10        |
| 7.2       | Building a Small Happiness Lab . . . . .                              | 10        |
| 7.3       | Guardrails and When to Seek Help . . . . .                            | 11        |
| <b>8</b>  | <b>Case Sketches</b>  | <b>11</b> |
| 8.1       | Alex: From Burnout to Sustainable Fulfilment . . . . .                | 11        |
| 8.2       | Sam: Navigating Social Anxiety and Connection . . . . .               | 12        |
| <b>9</b>  | <b>Limitations, Culture, and Ethics</b>                               | <b>12</b> |
| <b>10</b> | <b>Outlook and Connections to Other Deep Q-Life Papers</b>            | <b>12</b> |

# 1 Why Happiness, Why Deep Q-Life

Modern readers stand in the middle of several overlapping stories about happiness: consumer promises of “more”, scientific claims about measurable well-being, and philosophical or religious teachings that sometimes point away from external rewards altogether. This paper uses the Deep Q-Life (DQL) framework to align these stories in a single mathematical and narrative language.

Deep Q-Life views everyday life as a reinforcement learning (RL) problem. At each step you inhabit a state, choose an action, receive a reward, and move into a new state. The main Deep Q-Life paper on habits [1] introduced this structure with a focus on behaviour change; subsequent companions on sobriety and ownership [2, 3] extended it to alcohol and possessions. The present paper keeps the same backbone but reorients the reward function around different flavours of happiness.

This guide serves three main audiences in the happiness domain:

- primary readers: people who want a grounded, science-informed way to improve everyday well-being without getting lost in self-help jargon;
- mentors and supporters: coaches, therapists, educators, and friends who help others reflect on happiness and habits in a structured, compassionate way;
- edge cases: readers facing intense transitions or constraints (burnout, illness, caregiving, migration) who need to redesign what “happiness” can realistically mean in their circumstances.

Alongside these groups there are curious readers from reinforcement learning, philosophy, or contemplative practice who seek a common language between their domains. The central promise of Deep Q-Life Happiness is that diverse perspectives—from subjective well-being research to Stoicism and Buddhism—can be expressed as specific Markov decision processes (MDPs) with distinct reward functions, discount factors, and policies.

## Scope and Disclaimers

This paper offers conceptual tools, examples, and simple experiments for reflecting on happiness; it does *not* offer medical, psychological, financial, or legal advice. Readers experiencing severe or persistent distress, suicidal thoughts, substance dependence, or major psychiatric symptoms should seek support from licensed professionals (physicians, psychotherapists, psychiatrists, or emergency services). The Deep Q-Life framework is intended as a complement to, not a replacement for, evidence-based treatment and community or spiritual support.

We begin with language. Before writing down any equations, Section 2 disentangles common terms like joy, pleasure, excitement, contentment, and fulfilment. This shared vocabulary then anchors the scientific overview in Section 3, the philosophical and religious survey in Section 4, and the formal Deep Q-Life happiness model in Section 5. Later sections move from theory to practice, offering examples, checklists, and gentle experiments to help readers articulate and train their own “way” to happiness.

# 2 Language for Happiness: Joy, Pleasure, Contentment, Fulfilment

Everyday talk about happiness compresses several distinct experiences into a single word. For modelling and for honest conversation, it helps to separate them. This section sketches a working vocabulary that we will use throughout the paper; readers are invited to refine the labels so they fit their own cultures and lives.

## 2.1 Working Definitions and Everyday Examples

The following cluster of terms will appear repeatedly; they are not universal truths, but they provide a crisp starting point:

- **Pleasure** refers to short-lived, often sensory pleasantness: the taste of good food, warmth after a cold walk, the feeling of sliding into clean sheets.
- **Excitement** adds arousal and novelty: the rush before a performance, a rollercoaster, a risky deal, or a first date.
- **Joy** is a lighter, often social or aesthetic uplift: laughing with friends, seeing a child discover something, hearing music that moves you.
- **Contentment** points to a quiet “this is enough” feeling: sitting after a day of meaningful work, a simple meal when not rushed, a tidy room with no urgent tasks.
- **Fulfilment** connects actions to values and identity: working on a cause you care about, raising children, creating art, mentoring others.
- **Happiness** in this paper refers to an overall pattern across time that combines these ingredients: the long-run quality of life, not a single spike of feeling.

In the language of Markov decision processes (MDPs), pleasure and excitement are often associated with immediate rewards; contentment and fulfilment have slower dynamics, reflecting long-horizon returns from relationships, contribution, and growth. One purpose of Deep Q-Life Happiness is to help you see which part of this bundle you are optimising in a given decision.

### A Small Glossary for Conversations

When talking with partners, friends, or teams, it is helpful to know whether you are seeking the same thing when you say “I want to be happier”. Try questions such as:

- “Right now, are you longing more for comfort, excitement, connection, or progress?”
- “If we call tonight a success, what will we have maximised: fun, calm, depth, or relief?”
- “In this season of life, which matters more: fewer bad days, or more peak moments?”

This sort of micro-clarity prevents many silent mismatches in goals: one person may be chasing excitement while the other craves contentment.

## 2.2 Hedonic and Eudaimonic Happiness

Positive psychology distinguishes between *hedonic* and *eudaimonic* happiness. Hedonic perspectives emphasise pleasure, enjoyment, and the presence of positive emotions plus the absence of negative emotions. Eudaimonic perspectives, following Aristotle and later virtue ethics, emphasise living in accordance with one’s values, exercising excellence, and contributing to something larger than oneself.

Formally, hedonic happiness fits a reward function that weighs short-term mood heavily. Eudaimonic happiness looks more like an objective that discounts the present less, giving substantial weight to long-term coherence and contribution. A Deep Q-Life agent can blend these views by combining multiple reward channels: one for immediate affect, one for relationships, one for meaning. Figure 1 sketches this idea.

## Visualization: Reward Channels for Happiness

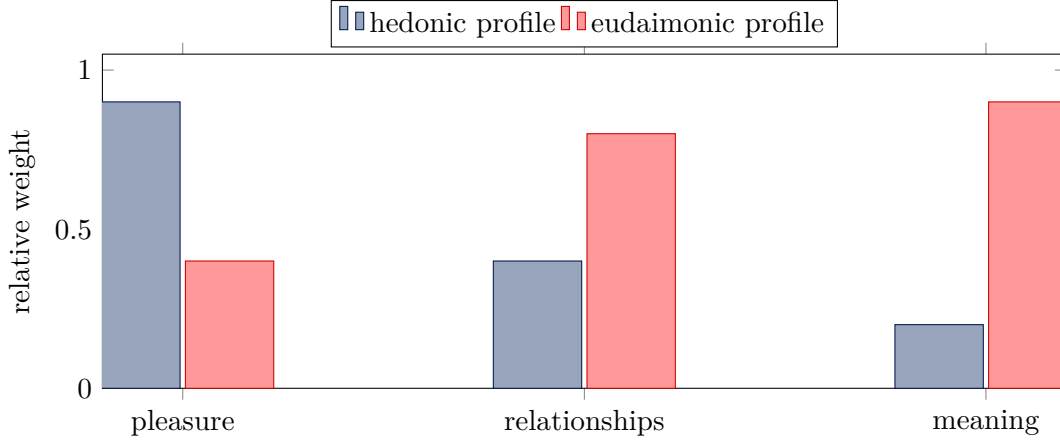


Figure 1: Illustrative reward weights for hedonic and eudaimonic happiness. In practice, individuals choose their own mixture; Deep Q-Life treats these weights as part of the design of the reward function.

### 2.3 A Simple Contentment Ratio

In the Deep Q-Life Ownership companion, we introduced a deliberately rough “contentment ratio”. Let  $X$  denote what you effectively *have and can use*: your current portfolio of time, money, skills, relationships, freedoms, and options. Let  $Y$  denote what you *want but do not yet have*: desired possessions, opportunities, status, and experiences. A qualitative contentment score is then

$$C = \frac{X}{Y}.$$

You do not need numbers; it is enough to sense whether  $C$  feels well below 1, roughly around 1, or comfortably above 1. When  $C > 1$ , you feel that what you have more than covers what you want. When  $C < 1$ , the sense of want dominates.

Equivalently, one can look at the inverse fraction

$$R = \frac{Y}{X},$$

which captures how large the “want list” looms compared with the “have list”. A value  $R > 1$  signals that desires currently outgrow resources;  $R < 1$  corresponds to a quieter, more spacious relationship between wanting and having. Both views describe the same tension: whether life feels like catching up or like having room to breathe.

## Contentment Ratio Revisited

Quick uses of the ratio in daily life:

- **Naming the numerator and denominator.** Instead of “I am just unhappy”, try “Right now my want-list  $Y$  is huge (promotion, new flat, different city), while my have-list  $X$  is small in the domains that matter.” This already suggests where change might help.
- **Two levers, not one.** You can improve  $C = X/Y$  by increasing  $X$  (building skills, savings, friendships, health) *or* by gently shrinking  $Y$  (questioning inherited desires, lowering unnecessary comparison, simplifying goals).
- **Explaining surprising contentment.** Some people with modest possessions or income feel deeply content because they cultivate a small, coherent  $Y$ . Others with high income and many options feel chronically behind because  $Y$  grows faster than  $X$ .

The contentment ratio is not a full theory of happiness, but it provides a compact way to talk about “enough”, and it meshes naturally with the state and reward structures we use throughout Deep Q-Life.

## 3 Evidence from Happiness Science

Over the past decades, researchers in psychology, economics, and sociology have assembled a large empirical literature on happiness and related constructs such as subjective well-being (SWB), life satisfaction, and affect balance. This section sketches how happiness is measured, which predictors are robust across studies, and how time and adaptation complicate the picture. It is not a full review; rather, it highlights findings that are especially useful for Deep Q-Life modelling.

### 3.1 Measurement: From Life Satisfaction to Momentary Affect

Happiness research uses several complementary measurement strategies:

- **Global evaluations** ask people to rate their life as a whole, for example via satisfaction scales [4].
- **Experience sampling** and day reconstruction methods capture moment-to-moment affect across many small episodes [5].
- **Domain-specific measures** track satisfaction with work, relationships, health, or finances.
- **Psychological need fulfilment** focuses on autonomy, competence, and relatedness, as in self-determination theory [6].

From a Deep Q-Life perspective, these correspond to different ways of aggregating rewards across time and contexts. Global evaluations approximate long-run value functions; experience sampling resembles a high-resolution stream of immediate rewards; need-based measures capture structural properties of the state space (for example, whether autonomy is chronically constrained).

## Quick Data-Driven Intuitions

Empirical work suggests several broad patterns:

- Life satisfaction and momentary affect are correlated but not identical; some people report “good lives” with many stressful days, others report pleasant days without strong meaning or progress.
- Income matters up to a point but exhibits diminishing returns, and its relationship with daily affect is more nuanced than simple “more is better” [7].
- High-quality relationships, sleep, physical activity, and reduced rumination show robust associations with well-being [8, 9].

These patterns are compatible with many philosophical positions; Deep Q-Life uses them to calibrate which state variables and actions deserve attention when designing a happiness-focused policy.

### 3.2 Dynamics, Adaptation, and Set Points

Hedonic adaptation describes the tendency for emotional responses to new circumstances—positive or negative—to fade over time. Lottery winners and accident survivors often drift back toward previous levels of reported well-being, although not always fully and not on all dimensions. Longitudinal studies suggest a combination of trait-like “set points” and environment- or habit-driven deviations [10, 9].

In Deep Q-Life terms, adaptation shows up as a slowly changing reward function and as shifts in baseline state values. A new car produces a spike in immediate reward that decays as the novelty becomes part of the default state. Practices such as gratitude journaling, mindfulness, or investing in relationships may not change external states quickly but can reshape how rewards are perceived and integrated over time.

#### Visualization: Adaptation Curves

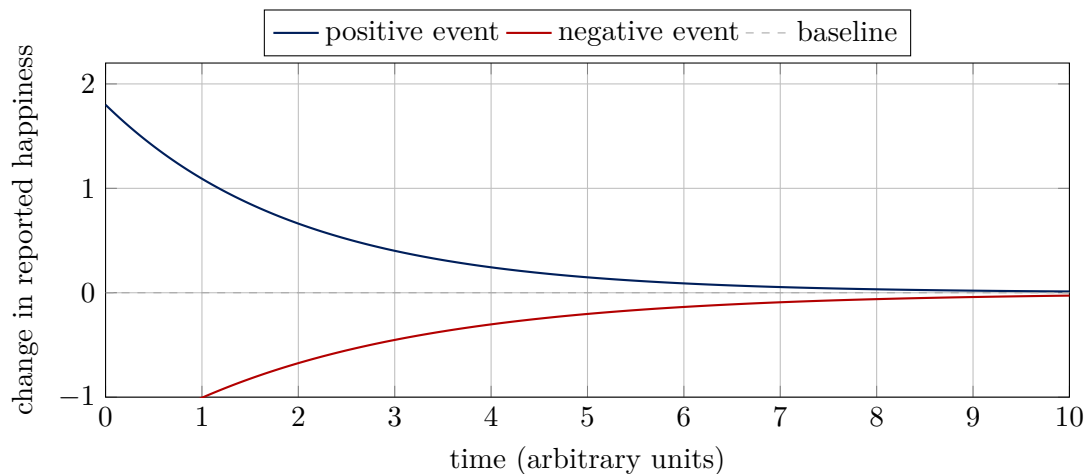


Figure 2: Illustrative adaptation curves after positive and negative events. Real trajectories are more complex, but the tendency to drift back toward a baseline is robust in many datasets.

## 4 Perspectives from Philosophy and Religion

Long before formal surveys and experience sampling, philosophical schools and religious traditions wrestled with questions about the good life. Their answers differ in language and emphasis, but many can be translated into Deep Q-Life terms: each tradition embodies a particular choice of reward function, discounting, and allowable actions.

### 4.1 Classical and Modern Philosophy

At least three philosophical strands are especially relevant:

- **Aristotelian eudaimonia** frames happiness as flourishing through virtuous activity over a lifetime. The focus lies on character, practical wisdom, and a life that “hangs together” across roles and seasons.
- **Stoicism** emphasises aligning with what is under one’s control and accepting the rest. Happiness is linked to inner freedom and the practice of virtue regardless of external fortune.
- **Utilitarianism and consequentialism** connect happiness to aggregate pleasure or preference satisfaction, often at the level of societies.

From a Deep Q-Life standpoint, Aristotelian and Stoic views resemble agents with high discount factors and reward functions that heavily weight integrity, courage, and justice. Utilitarianism pushes the reward function to include other people’s states, turning happiness into a multi-agent problem.

### 4.2 Buddhist and Contemplative Perspectives

Many Buddhist teachers highlight a paradoxical insight: “There is no way to happiness, happiness is the way,” a line popularised by Thich Nhat Hanh. Here happiness is not a distant state achieved once certain external conditions are met; it is the quality of awareness and compassion brought to each step.

In Deep Q-Life language, these teachings question the assumption that happiness is the maximisation of a scalar sum of rewards. They encourage reducing craving and attachment, loosening identification with transient states, and cultivating present-moment awareness. One way to express this formally is to treat the reward function not just as “how good things feel” but as “how skillfully one relates to experience”, including discomfort.

#### Present-Moment Practice as a Policy

A contemplative approach to happiness can be seen as a small set of policies for common states:

- when discomfort or craving arises, pause, breathe, and acknowledge it rather than reacting automatically;
- when joy arises, notice and appreciate it without clinging or fear of loss;
- when boredom or restlessness appears, experiment with small, wholesome actions (helping someone, learning something, moving the body).

These moves change transitions and rewards even when external circumstances stay the same.



## 5 Deep Q-Life Happiness as an MDP

With language and empirical context in place, we can now describe happiness-oriented life design as a Markov decision process. The point is not to reduce humans to numbers, but to make assumptions explicit so that they can be inspected, debated, and gently improved.

### 5.1 States: What Matters for Happiness

For happiness, useful state descriptions typically include:

- **Physiology:** sleep, energy, pain, hunger, and other bodily signals;
- **Emotion and thought:** current mood, level of worry or rumination, sense of purpose;
- **Relationships:** perceived support, conflict, loneliness, and opportunities for connection;
- **Autonomy and competence:** room to choose, perceived efficacy, progress on personally meaningful projects;
- **Context:** time of day, location, workload, obligations in the next 24 hours.

Self-determination theory [6] suggests that autonomy, competence, and relatedness are key psychological nutrients; we treat them as central components of the happiness-relevant state space.

### 5.2 Actions and Habits

Actions are the options available from each state. Many evidence-backed happiness interventions can be seen as actions in this MDP:

- connecting with others (sending a kind message, calling a friend, joining a shared activity);
- moving the body (a short walk, stretching, dancing, light exercise);
- shaping attention (gratitude practice, mindfulness, journaling, limiting rumination triggers);
- investing in progress (small steps on meaningful projects, learning, creative work);
- managing load (saying no, delegating, scheduling rest).

Deep Q-Life treats recurring versions of these moves as habits: compact policies that trigger the action in specific states with little extra deliberation.

### 5.3 Rewards, Value Functions, and Discounting

Let  $S_t$  denote the state at time  $t$ ,  $A_t$  the action, and  $R_t$  the scalar reward. A standard reinforcement learning objective is

$$J(\pi \mid s_0) = \mathbb{E}_\pi \left[ \sum_{t=0}^{\infty} \gamma^t R_t \mid S_0 = s_0 \right],$$

where  $\gamma \in [0, 1)$  is a discount factor and  $\pi$  is the policy. In happiness applications, it is often helpful to imagine  $R_t$  as a weighted sum of channels:

$$R_t = w_{\text{pleasure}} r_t^{\text{pleasure}} + w_{\text{connection}} r_t^{\text{connection}} + w_{\text{meaning}} r_t^{\text{meaning}} + \dots$$

Weights and discount factors encode philosophical commitments: a highly hedonic life places large weight on immediate  $r_t^{\text{pleasure}}$  and may use a lower  $\gamma$ , whereas a eudaimonic approach tilts weight toward  $r_t^{\text{meaning}}$  and uses a higher  $\gamma$ .

## Design Questions for Your Happiness MDP

When making or revisiting life decisions, it can help to ask:

- “Which states do I repeatedly experience that strongly affect my well-being? How can I describe them compactly?”
- “Which small, realistic actions from those states tend to generate better days or weeks?”
- “If I had to choose three reward components to track—for example, mood, connection, and progress—which would I pick, and how would I weight them?”
- “Does my behaviour suggest a low or high discount factor—do I often trade tomorrow’s happiness for today’s comfort, or the reverse?”

Writing down rough answers already clarifies which “version of happiness” you have been optimising and whether it matches what you consciously endorse.

## 6 Everyday Frames as Formal Models

People carry implicit models of happiness: “success and status bring happiness”, “family is everything”, “freedom matters most”, “service is joy”, or “happiness is overrated; meaning is what counts”. Each slogan corresponds to a particular choice of state variables, reward weights, and allowed actions.

### 6.1 Contrasting Frames: Achievement, Comfort, and Presence

Three stylised frames illustrate the idea:

- **Achievement frame:** rewards promotions, recognition, income, and visible milestones; costs are effort, stress, and opportunity cost on relationships or health.
- **Comfort frame:** rewards low effort, predictability, and avoiding discomfort; costs are stagnation and missed growth.
- **Presence frame:** rewards quality of awareness, alignment with values, and compassionate action, even in difficulty.

In Deep Q-Life notation, the achievement frame gives large positive reward to states with high status markers, often with a relatively high discount factor but narrow state description. The comfort frame uses a smaller discount factor and focuses on short-term relief. The presence frame shifts attention to how skilfully one navigates each state, echoing contemplative teachings.

## Visualization: Three Reward Landscapes

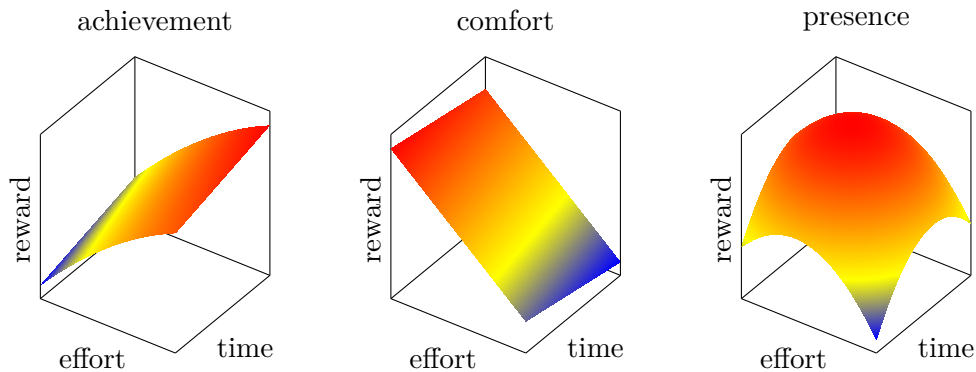


Figure 3: Illustrative reward surfaces for three frames. The achievement frame peaks at high effort and later times; the comfort frame peaks at low effort now; the presence frame rewards moderate, sustainable effort with stable awareness.

## 7 Designing Your Own Way to Happiness

This section turns the formalism and evidence into simple, humane practices. The goal is not to “maximise” happiness in a mechanical way but to help readers align actions with their own definition of a good life, while respecting constraints and uncertainty.

### 7.1 Clarifying Signals and Language

One practical move is to define a small handful of signals you care about and can observe regularly. These need not be perfect; they simply anchor reflection. Common choices include:

- daily mood on a 0–10 scale, recorded at a fixed time;
- sense of connection (for example, number of meaningful interactions per day);
- sense of progress on one or two meaningful projects;
- sleep quality or energy level;
- moments of joy, gratitude, or beauty noticed.

These signals serve as partial observations of the underlying state and reward processes. Over weeks, simple plots or notes reveal patterns that can inform policy updates.

### 7.2 Building a Small Happiness Lab

Deep Q-Learning (DQL) in machines relies on exploration, replay, and gradual policy improvement. A human-scale analogue can be implemented with light rituals:

- choose one or two states you want to improve (for example, “late evening alone at home” or “Monday mornings at work”);
- brainstorm a small set of candidate actions that are realistic in those states;
- experiment with these actions for a week, noting how they affect your chosen signals;
- run a brief weekly review to keep, tweak, or drop candidate habits, updating your “policy”.

Readers familiar with the DQL habits paper can treat this as a specialised variant of the weekly review loop, with happiness-specific rewards and constraints.

#### Seven-Day Micro-Experiment

A concrete template for a week-long experiment:

- **Day 1–2:** track baseline states and signals without changing anything.
- **Day 3–5:** introduce one small action in a chosen state (for example, a five-minute walk after work, sending one kind message, or a two-minute breathing pause before opening social media).
- **Day 6–7:** review notes, rate days on a simple happiness scale, and decide whether the action is worth keeping, adjusting, or discarding.

This process frames happiness practices as experiments rather than verdicts on your character.

### 7.3 Guardrails and When to Seek Help

Not all dips in happiness are amenable to self-experimentation. Warning signs include persistent anhedonia (loss of interest or pleasure), thoughts of self-harm, severe sleep disruption, substance dependence, or inability to function in daily roles. In such cases, self-directed Deep Q-Life exercises should be paused and professional support sought, using the formal framework mainly as a language to communicate patterns and goals to clinicians and trusted supporters.

## 8 Case Sketches

The following short case sketches illustrate how different people might use a Deep Q-Life happiness lens. The aim is to show diversity in circumstances and in what “happiness” means, while keeping the underlying mathematics similar.

### 8.1 Alex: From Burnout to Sustainable Fulfilment

Alex is a mid-career professional who has checked many external boxes: promotions, a good salary, an impressive LinkedIn profile. From the outside, Alex looks successful; on the inside, most days feel like running on fumes. Mornings start with email in bed, evenings end with exhausted scrolling. There are flashes of excitement—a big presentation, a product launch—but contentment and fulfilment are rare visitors.

In Deep Q-Life terms, Alex has been living inside the achievement frame from Section 6.1. The state description is narrow: calendar, inbox, revenue, and status markers dominate; energy, health, and relationships appear only as vague background noise. The implicit reward function pays generously for visible output and recognition, discounts sleep and recovery, and treats time with family or friends as something to fit in “once real work is done”.

A happiness-focused Deep Q-Life intervention for Alex starts by rewriting the MDP:

- **State expansion.** Daily states are enriched with sleep hours, stress 0–10, physical symptoms, and a simple relationship score (for example, “Did I have at least one unhurried conversation today?”).
- **Reward rebalancing.** Immediate mood and status still matter, but the reward function gains explicit terms for medium-horizon energy and for progress on personally meaningful, non-career projects.

- **Policy experiments.** In one concrete week, Alex commits to leaving the office on time twice, schedules a 20-minute walk after work, and protects two evenings for friends or hobbies, treating these as deliberate actions in frequently visited states.

On paper, these are tiny moves. In experience, they generate new transitions: slightly better sleep, a different Monday-morning mood, moments of joy that are not tied to work. Weekly reviews make the shift visible: as the policy gradually favours recovery and connection, the value of high-pressure, late-night work sessions drops. Alex’s version of “fulfilment” becomes less about a single heroic outcome and more about a sustainable pattern of days.

## 8.2 Sam: Navigating Social Anxiety and Connection

Sam is a student who often dreads group situations. Before a seminar or a party, Sam’s heart rate climbs; the safest option seems to be staying in the room with headphones on. Afterwards, there is a double ache: loneliness from not having gone, and self-criticism for “failing” again. Intellectually, Sam knows that friendships and shared projects are important for long-run happiness. Emotionally, the comfort frame dominates: avoiding anxiety today feels more rewarding than facing it for the sake of tomorrow.

With a Deep Q-Life lens, Sam and a trusted supporter sit down to write a compact MDP. States include context (class, cafeteria, online), anxiety 0–10, fatigue, and whether at least one safe person is nearby. Actions are small and graded: saying hello without starting a full conversation, sitting closer to the group instead of at the back, sending one kind message, or planning a short walk with a friend after class. Rewards include not only immediate relief or discomfort, but also a simple connection score and explicit shaping for self-compassion (for example, giving extra credit for trying a hard thing, regardless of outcome).

Over a few weeks, Sam runs micro-experiments: on three days, when anxiety is below a chosen threshold, Sam chooses a “reach out by one step” action instead of full avoidance. Each evening, Sam notes how this affected mood and the sense of belonging. In RL language, Sam is gathering transitions  $(s, a, r, s')$  and updating a policy. In human language, Sam is discovering that some feared states are survivable, that connection can be built in tiny increments, and that being kind to oneself after awkward moments is part of the happiness reward function rather than a luxury.

## 9 Limitations, Culture, and Ethics

The Deep Q-Life happiness model is deliberately simple. It does not capture all aspects of culture, history, power, or injustice, and it cannot, by itself, resolve structural barriers to well-being. Optimising an individual reward function in an unhealthy environment can even backfire, leading to maladaptive accommodation rather than change.

Readers are therefore invited to hold the framework lightly: as a tool for clarity and gentle experimentation, not as a totalising theory. Ethical questions arise about whose happiness is counted in a given reward function, how intergenerational effects are weighed, and how to avoid blaming individuals for suffering rooted in structural conditions. These remain active areas for philosophical and policy debate.

## 10 Outlook and Connections to Other Deep Q-Life Papers

Deep Q-Life Happiness sits alongside the habit, sobriety, and ownership companions [1, 2, 3] as part of a broader project: treating everyday life design as a series of learnable policies grounded in both mathematics and human stories. Happiness here is not a static target but a pattern emerging from many decisions about attention, relationships, work, and rest.

For readers who wish to go deeper into the formal reinforcement learning machinery, the main Deep Q-Life paper on habits [1] provides more detailed derivations and visualisations of value functions, Bellman equations, and Deep Q-Learning updates. For concrete domains where happiness is tightly coupled to specific behaviours, the sobriety and ownership guides [2, 3] offer domain-specific policies and experiments.

### Practical Next Steps

A simple sequence for integrating the ideas in this paper:

- choose one happiness-related signal to track for the next two weeks;
- pick one small, evidence-backed action to test in a recurring state;
- schedule a weekly review to update your informal policy, informed by what you notice and by your own philosophical or spiritual commitments.

Over months, these small loops can accumulate into a personal “way to happiness” that is both scientifically informed and deeply your own.

## References

- [1] Y. J. Hilpisch. Deep Q-Life: Markov, Bellman, and Habits. Technical report, 2025. Available at [https://hilpisch.com/dql\\_habits.pdf](https://hilpisch.com/dql_habits.pdf).
- [2] Y. J. Hilpisch. Deep Q-Life Sobriety: Applying RL Habit Loops to Alcohol-Free Living. Technical report, 2025. Available at [https://hilpisch.com/dql\\_sobriety.pdf](https://hilpisch.com/dql_sobriety.pdf).
- [3] Y. J. Hilpisch. Deep Q-Life Ownership: Applying RL to Stuff, Spending, and Space. Technical report, 2025. Available at [https://hilpisch.com/dql\\_ownership.pdf](https://hilpisch.com/dql_ownership.pdf).
- [4] E. Diener. Subjective well-being. *Psychological Bulletin*, 95(3):542–575, 1984.
- [5] D. Kahneman, A. B. Krueger, D. A. Schkade, N. Schwarz, and A. A. Stone. A survey method for characterizing daily life experience: The Day Reconstruction Method. *Science*, 306(5702):1776–1780, 2004.
- [6] E. L. Deci and R. M. Ryan. The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4):227–268, 2000.
- [7] D. Kahneman and A. Deaton. High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38):16489–16493, 2010.
- [8] M. E. P. Seligman. *Flourish: A Visionary New Understanding of Happiness and Well-being*. Free Press, 2011.
- [9] S. Lyubomirsky. *The How of Happiness: A New Approach to Getting the Life You Want*. Penguin Press, 2007.
- [10] P. Brickman and D. T. Campbell. Hedonic relativism and planning the good society. In M. H. Appley, editor, *Adaptation-Level Theory*, pages 287–302. Academic Press, 1971.
- [11] T. N. Hanh. *Peace Is Every Step: The Path of Mindfulness in Everyday Life*. Bantam, 1991.