

TRADING KNOWLEDGE LIBRARY

The Plan That Can Be Tested

Trading Strategies



Trading Papers

The Plan That Can Be Tested

Writing rules precise enough that a stranger could follow them

Published by	Trading Papers
Series	The trading knowledge library
Edition	First edition · 2026-08-26
Format	A4 · digital

Educational content only. Nothing in this document is financial advice, no part of it is a recommendation to buy or sell anything, and it knows nothing about your circumstances. Trading carries risk, including the loss of the amount invested.

© 2026 Trading Papers. All rights reserved.



Trading Papers

Contents

01	Two documents called the same thing	4
02	The stranger test	5
03	The seven fields	6
04	Rules and preferences	7
05	Making a condition testable	8
06	Size, which is not optional	9
07	The record the plan is judged on	10
08	One plan, written out	11
09	Testing it honestly	13
10	When a rule may change	14
11	What automation actually is	16
12	What an unattended plan needs	17
13	What the plan says about growth	18

01

Two documents called the same thing

Almost every trader has a plan, and almost none of them has the kind being described here. The word covers two entirely different documents: one is a statement of what you intend to do, and the other is a specification of what will be done. They look similar on paper and behave nothing alike.

	A statement of intent	A specification
Says	Trade with the trend. Cut losses. Be disciplined.	Long only while the daily close is above the 50-day average.
Can be	Agreed with, and then forgotten by Thursday.	Applied by a stranger, and shown to be wrong.

The two documents that both get called a trading plan. The left one is what most people write and it cannot be wrong about anything, which is the same as saying it cannot be tested.

The left column is what most plans contain: trade with the trend, cut losses quickly, be disciplined, do not overtrade. Every line is true, none can be violated in any way you could prove afterwards, and by Thursday the document has no bearing on what happens. A plan that cannot be broken is not being followed; it is being agreed with.

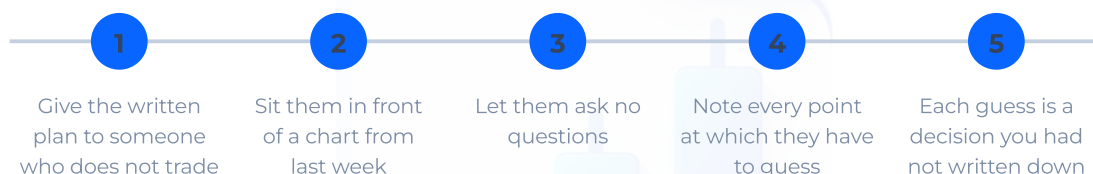
The right column is duller and it is the only version that does anything. Its property is not strictness but decidability: another person could read it, look at a chart, and reach the same conclusion you would. That property is what makes a record meaningful, what makes testing possible, and what eventually makes automation possible — which is where this paper ends.

This paper assumes the first two in the series: that you have a description of the job, and that you apply the same standard of evidence to your own results as you would to a stranger's. The plan is where that standard becomes daily practice.

02

The stranger test

There is a single test that settles whether what you have written is a plan, and it takes an hour.



The test that decides whether you have a plan. It costs nothing, takes an hour, and almost every plan fails it the first time — which is the useful outcome, because it shows exactly which decisions are still being made in the moment.

Hand the document to someone who does not trade, sit them in front of last week's chart, and let them ask nothing. Every point at which they have to guess is a decision you believed you had written down and had not. Most plans produce six or seven such points on the first attempt, and every one of them is a place where, in live trading, the decision is made by whatever mood is at the keyboard.

The test is uncomfortable for a specific reason: the gaps are never in the parts you find difficult. They are in the parts so obvious to you that writing them seemed unnecessary — which instrument, at what time, how much, and what counts as the condition being met. Those are exactly the decisions that drift, because nothing in the document holds them still.

03

The seven fields

A plan that survives the stranger test contains seven things. There is nothing sacred about the number; it is what remains when you remove everything that does not change a decision.

Field	The question it answers	Written as
Universe	what may be traded at all	a named list of instruments
Window	when the plan is active	hours, in one named clock
Condition	what has to be true first	a state of the market, testable
Trigger	what starts the trade	a price event, not an impression
Invalidation	what proves the idea wrong	a price, decided before entry
Size	how much is risked	a fixed fraction, computed from the stop
Exit	what ends the trade	a rule, and the same one every time

The seven fields a plan needs before it is testable. Anything not on this list is optional; anything missing from it is a decision that will be made under pressure by whoever is at the keyboard.

Read the middle column as a list of questions that will be answered one way or another. If the plan does not name the universe, you will trade whatever is moving; if it does not name the window, you will trade whenever you are free; if it does not fix the size, the size will be chosen by how the last trade went. The plan does not create these decisions — it decides them in advance, at a moment when nothing is at stake.

Two of the seven do the most work and are the most often vague. The invalidation has to be a price, decided before entry, corresponding to something that would actually mean the idea was wrong; and the size has to follow from that price by arithmetic rather than by feel. The rest of the plan can be mediocre and still function; these two cannot.

04

Rules and preferences

The most useful distinction in this paper is between a rule and a preference, because a document full of preferences reads exactly like a plan and does none of the work.

	A rule	A preference
Example	No entries after 16:00.	Prefer to trade in the morning.
Behaviour	Refuses a trade at 16:05 that would have worked.	Permits any trade, and is cited afterwards as a reason.

How to tell a rule from a preference in your own writing. The difference is not the wording but whether it can refuse you: a rule sometimes stops you doing what you want, and a preference never does.

The test is not the wording. It is whether the sentence can refuse you. A rule that has never once stopped you taking a trade you wanted is not a rule you are following; it is a description of what you were going to do anyway. Somewhere in a working plan there is a line that has cost you a winner, and its having done so is the evidence that it exists.

The word	What it means in practice	Replace with
strong	whatever felt strong at the time	a measured threshold
confirmation	one more candle than you first wanted	a named event
good setup	a setup you already liked	the checklist, ticked
usually	a permission to make exceptions	a percentage, and a count
if conditions allow	always, when you want to trade	the condition, stated

The words that quietly turn a rule into a preference. Each one is a place where the plan hands the decision back to you at the moment you are least able to make it well.

The words in that table are where rules turn back into preferences without anyone noticing. Strong, confirmation, good setup, usually — each hands the decision back to the person at the exact moment the plan was supposed to have already made it. Replacing them is tedious and it is most of the work of writing a plan that functions.

05

Making a condition testable

Conditions are where vagueness hides most comfortably, because a vague condition still feels like analysis. It is worth seeing the same idea written at four levels of precision.



Four ways to write the same condition, from unusable to testable. Each rung removes a place where a human has to decide, and the bottom rung is the only one a record can be kept against.

Every rung describes the same intuition, and only the bottom one can be counted. Note

what the descent costs: the top rung applies to almost every chart, and the bottom rung applies to a specific, countable subset. That narrowing feels like losing opportunity and is the entire point — a condition that admits everything has not conditioned anything.

It is also worth being honest that the bottom rung is not better because a moving average is wise. It is better because it is unambiguous. Any specific rule you can state and count will teach you more in fifty trades than an excellent instinct you cannot write down will teach you in five hundred.

06

Size, which is not optional

Of the seven fields, size is the one most often left as an understanding rather than a number, and it is the one that decides whether the plan survives being wrong.

Stop distance	Risk budget	Position	What changes
10 points	1% of the account	large	the position, not the risk
25 points	1% of the account	medium	the position, not the risk
60 points	1% of the account	small	the position, not the risk
Any	'a comfortable amount'	unknown	the risk, every single time

Size is the field most often left implicit, and it is the one that decides whether a plan survives being wrong. Every row risks the same fraction of the account; only the position changes, because only the distance changed.

The first three rows are the same decision applied to different distances: risk a fixed fraction of the account, and let the position be whatever makes the stop cost that fraction. Nothing about your confidence enters the arithmetic. The last row is what happens without the rule, and it has a predictable shape — larger on the trades you like, which are not the trades that win more often.

There is one addition worth making from the start, and the previous paper's standard of evidence is the reason. Cap the position independently of the formula, and take the smaller

of the two. When the stop is unusually tight the formula asks for a size that will slip badly, and the cap is the only part of the plan that notices.

07

The record the plan is judged on

A plan and a log are the same project. The plan says what will be done; the log says what was done; the difference between them is the only thing that can be improved.

Column	Why it is there
Date and instrument	the ordinary index
Which condition was met	ties the trade to the rule that produced it
Trigger price and fill	measures slippage without a separate exercise
Stop distance and size	shows whether risk really was constant
Exit rule that fired	separates the plan from the improvisation
Result in R	makes trades of different sizes comparable
Rule broken, if any	the single most informative column

The record the plan is tested against. The middle column is the one that makes the log useful later: without it you can see what happened, and never why the plan said to do it.

Most logs record entry, exit and profit, which is enough to compute a result and not enough to learn anything. The columns that matter are the ones tying each trade to the rule that produced it and recording whether the rule was actually followed. Without them, a losing month cannot be told apart from a month of ignoring the plan, and those two have opposite remedies.

Followed the plan [] the only trades the plan is responsible f

Rule bent [] entered early, or without the condition

No rule at all [] these belong in a separate account

Schematic, and the measurement almost nobody takes: what share of trades actually followed the plan. Until this number is high, the record is measuring a person rather than a method, and changing the method cannot help.

That last figure is the measurement almost nobody takes, and it should be taken before any other. If a third of your trades bent a rule, the record is measuring you rather than the method — and changing the method in response is the most common expensive mistake in this paper. Compliance first, expectancy second.

08

One plan, written out

The seven fields are easier to judge as a document than as a list, so here is one written in full. It is deliberately ordinary: nothing in it is recommended, and the only property being demonstrated is that every line can be acted on without asking what was meant.

Trading Papers

Field	This plan says
Universe	EUR/USD and GBP/USD only
Window	07:00–11:00 UTC, weekdays
Condition	daily close above the 50-day average
Trigger	a 5-minute close above the previous session's high
Invalidation	the low of the candle that triggered it
Size	0.5% of the account, computed from that distance
Exit	half at 2R, the rest on a close below the 20-period average

One plan, written out in full, so the seven fields stop being an idea. Nothing here is recommended — the point is the form: every row is a sentence a stranger could act on without asking what was meant.

Read it as the stranger would. There is no instrument to choose, no hour to judge, no assessment of whether the market looks right, and no decision about size — the distance from the trigger to the invalidation determines it. The only thing left to a human is noticing that the trigger occurred, which is precisely the part a machine takes over later.

Two lines are doing quiet work. The condition uses the daily chart while everything else happens on the five-minute, which means the plan is silent for weeks at a time and nothing about that is a malfunction. And the exit is a rule rather than a target, so it does the same thing on a day when you feel certain as on a day when you do not.

Month	What happens	What it is
One	the plan is followed exactly	the sample everyone wishes they had
Two	one trade taken just outside the window	the window becomes a preference
Three	size raised after three winners	the fraction becomes a feeling
Four	the log is a week behind	the evidence stops being collected
Five	the plan is rewritten from memory	a new plan, with no sample at all

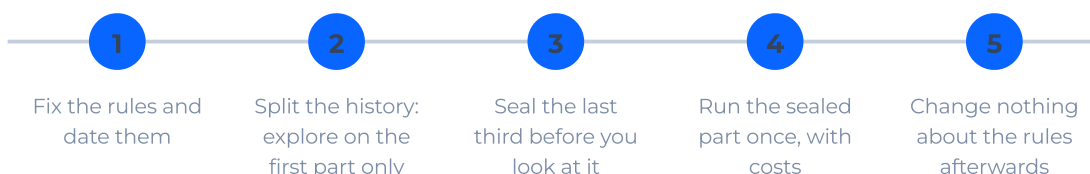
How a written plan stops being followed, in the order it happens. None of these is a decision to abandon it; each is a small local amendment that felt reasonable at the time.

The table beneath it is the ordinary life of a written plan, and it is worth reading before rather than after. No step in it is a decision to abandon anything; each is a small amendment made for a reason that sounded good on the day. The result by month five is a document with no sample behind it, which is where most traders are when they conclude that plans do not work.

09

Testing it honestly

Once the plan is written and the log exists, the plan can be tested. The procedure is short and the discipline in it is the whole difficulty.



How to test a written plan without fooling yourself. The third step is the one that gets skipped, and skipping it turns the exercise into a search for the version of the plan that would have worked.

The step that gets skipped is the third. Sealing part of the history before looking at it is what separates a test from a search: if every rule was chosen while looking at all the data, the resulting curve describes that data and nothing else. A plan tested only on the period it was invented in has been fitted, however sincerely.

Test	What it shows	What it cannot show
On history	the rules are coherent and were profitable	that you chose them before seeing the data
On sealed data	the rules survive data they were not fitted to	that you can execute them
Live, smallest size	whether you can actually follow them	much about expectancy yet

The three tests a plan passes in order, and what each one can and cannot establish. Skipping to the third is the usual route, and it is why so many plans are abandoned in month two.

The three tests answer different questions and the order matters. History shows the rules are coherent; sealed data shows they survive something they were not built on; live trading at the smallest size shows whether you can execute them at all — which is a question about you, arrives fastest, and is the one most people skip past on their way to size.

10

When a rule may change

A plan that never changes is not disciplined, it is fossilised. The question is what licenses a change, and there is a clean answer.

	Legitimate change	Not a change, a reaction
Trigger	A hundred trades, and the record disagrees with the rule.	Three losses, and the rule feels wrong.
What follows	One rule altered, dated, new sample begun.	Several rules altered, no date, sample restarted invisibly.

When a rule may be changed, and when the urge to change it is the thing being tested. The lower-left cell is where most plans die, quietly and with good intentions.

A legitimate change is one the record asked for: a hundred trades in which a rule can be shown to have cost more than it saved. That change is made one rule at a time, dated, and followed by a new sample beginning on that date — because a plan altered mid-sample produces a record of neither version.



The monthly review, which is where a plan improves. It is a meeting with a document, and the order matters: compliance is read before results, because results interpreted before compliance teach the wrong lesson.

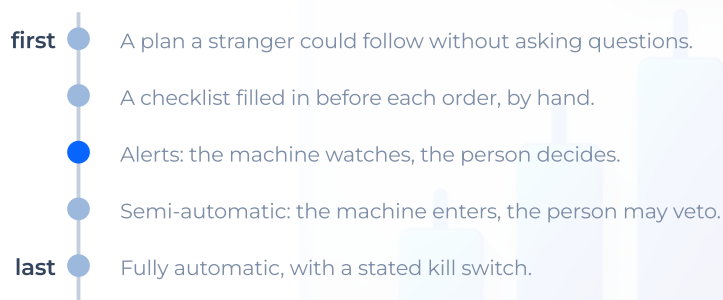
The monthly review is where that happens, and its order is deliberate. Compliance is read first, then the broken-rule column, then results. Reading results first means every subsequent judgement is made by someone who already knows whether the month was good, and that person cannot assess a rule fairly.

One useful constraint while the habit is forming: no rule may be changed in the same week it was broken. It removes nothing legitimate — a rule the record genuinely condemns will still be condemned next week — and it removes almost everything illegitimate.

11

What automation actually is

A written plan precise enough for a stranger is already most of the way to a plan precise enough for a machine, and this is the honest reason automation appears at the end of this paper rather than in a paper of its own.



The path from a written plan to code, in the order the steps actually happen. Nothing here starts with programming: automation is the last step of specification, not a substitute for it.

The ladder is the order the steps actually happen in, and none of the early ones involve programming. A checklist filled in by hand before each order removes more errors than most people expect. Alerts move the watching to the machine and leave the deciding with the person, which is where most of the benefit is and most of the risk is not.

	What it fixes	What it cannot fix
In execution	Hesitation, fatigue, and the trade taken out of boredom.	An edge that was never there.
In the method	Sizing arithmetic, done identically every time.	Rules fitted to the data they were found in.

What automation changes and what it does not. The right column is the reason an automated version of a bad plan loses money faster rather than differently.

What automation fixes is execution: hesitation, fatigue, the trade taken because nothing had happened for two hours, and sizing arithmetic done identically every time. What it cannot fix is the method. An automated version of a plan with no edge does not lose differently; it loses faster and with better documentation, and the documentation is genuinely the consolation prize.

It is worth adding the part that vendors of automated systems tend not to: building one properly is a long, unglamorous engineering job. The strategy is a small part of it. Data handling, order errors, partial fills, reconnections, broker quirks and the arithmetic of position rounding are most of the work, and they are the parts that fail at three in the morning.

12

What an unattended plan needs

Anything that runs without you needs a set of decisions that have nothing to do with entries, and every one of them should be made while nothing is running.

Control	Set to	Why
Daily loss limit	a figure, in R or in cash	stops a broken day becoming a broken month
Maximum open positions	a number	correlated trades are one trade
Kill switch	a condition that stops everything	a way out that needs no judgement
Disconnection behaviour	what happens if it drops	the market does not pause with you
Review cadence	a date, in the calendar	unattended is not the same as unwatched

What an automated plan needs beyond its rules. Each row is a decision made while nothing is running, and each exists because the alternative is deciding it during the first bad afternoon.

The kill switch deserves particular emphasis because it is the control that exists for the moment when your judgement is least available. It should be a condition — a loss figure, a number of consecutive errors, a data feed gap — that halts everything without requiring anyone to assess the situation, because assessing the situation is exactly what will not happen well at that moment.

The last row is the one people misunderstand. Unattended does not mean unwatched. A strategy left running without a scheduled review is not automated, it is abandoned, and the difference shows up as a slow drift no one notices until a statement is opened for a different reason.

13

What the plan says about growth

One last field is missing from most plans, and the final paper in this series is about what happens when it is: what the plan says as the account gets larger.

As the account grows	What changes	What the plan should say
Position sizes rise	slippage grows with size	a size cap, independent of the formula
The same risk feels larger	1% is now a real sum	the fraction, fixed in advance
Withdrawals begin	compounding is interrupted	when and how much, written down
Confidence rises	the temptation is to loosen rules	no rule changes without a hundred trades

What a plan needs to say about growth, because an account that doubles is not the same account twice. Each row is a limit that arrives quietly and is best decided before it does.

Every row arrives quietly. Positions grow with the account, and slippage grows with positions; one per cent of a larger account is a sum that feels different even though the arithmetic is unchanged; withdrawals interrupt compounding and should be scheduled rather than improvised; and confidence rises after a good run, which is precisely when a plan is most likely to be loosened and least likely to deserve it.

Deciding these in advance costs nothing while the account is small. Deciding them after a doubling is deciding them in the state the next paper describes — which is the state in which almost every gain of the preceding months is handed back.

Trading Papers

Term	As used in this paper
Plan	A written specification of every decision, fixed in advance.
Rule	A statement that can refuse you a trade you want.
Condition	A testable state of the market, required before a trigger counts.
Trigger	The price event that starts the trade.
Compliance	The share of trades that actually followed the plan.
Out-of-sample	Data withheld while the rules were being chosen.
Kill switch	A pre-set condition that halts everything without judgement.

The terms used in a fixed sense throughout. They carry forward into the last paper of the series, which follows one account through a good run and what came after it.

Educational content only. Nothing in this document is financial advice, and no part of it is a recommendation to buy or sell anything.

Trading Papers