Standard usability (in person or remote)
This is a typical user test. You show a user a product, give them specific scenarios, and ask them to try to perform tasks.

It is good for:
•Learning whether people can perform specific tasks
•Understanding why particular metrics (ie. purchase flow, sign up flow) look the way they do
•Finding confusing bits in specific user flows

It sucks for:
•Learning how people actually use your product
•Understanding whether people will enjoy using your product
•Understanding whether people will return or continue using your product
•Testing over all messaging or visual design
•Understanding if you product idea is any good

Perform it on:
•Prototypes
•Your real product
•Your competitors' products

Purely Observational usability (in person or remote)
This is a user test where you simply watch the user do whatever they would normally do with the product. For brand new users, you might tell them how they found the product, but then you basically let them do whatever they would normally do while you observe and ask a few questions about why people do particular things.

It is good for:
•Understanding how your users are actually interacting with the product
•Getting a better understanding of the real new user experience
•Finding your biggest customer pain points
•Getting inspiration for new features
•Getting an idea of whether people might enjoy your product (not perfect though)
•Idea validation (in that it allows you to see if people have the problem you want to solve)

It sucks for:
•Understanding whether people will return or continue using your product
•Testing over all messaging or visual design
•Figuring out whether people can perform particular tasks with your product (unless they happen to perform that task while you're watching)

Perform it on:
•Your real product
•Your competitors' products

Contextual Inquiry (typically in person)
This is a very open ended, generative type of research that allows you to better understand your users and the context in which they might use your product. This is similar to purely observational usability, but I often perform it before there's a product at all. This is where you're really starting to understand the users' problems.

It is good for:
•Understanding user problems
•Learning how users are currently solving their problems
•Seeing the environment in which people are performing certain tasks and learning about their workflows and constraints
•Generating ideas for features or products
•Idea validation (in that it allows you to see if people have the problem you want to solve)
•Getting an idea of whether people might enjoy your product (not perfect though)

It sucks for:
•Understanding whether people will return or continue using your product
•Testing over all messaging or visual design
•Figuring out whether people can perform particular tasks with your product (unless they happen to perform that task while you're watching)

Perform it on:
•humans in their natural environment

Five Second Testing (remote micro-usability)
Allows people to see your landing page for 5 seconds and then asks a few questions about it.

It's good for:
•Seeing whether your CTAs and messaging are conveying the real use of a particular screen

It sucks for:
•Everything else

Perform it on:
•Landing page mockups
•Any page where you need to understand if the main message of the screen is being conveyed in a super obvious manner

Click Tests (remote micro-usability)
These show mockups to people and have them click through a particular scenario.

It is good for:
•Testing navigation on main user flows
•Testing to make sure CTAs are findable

It sucks for:
•Testing any sort of navigation on secondary user flows (ie. error states, wrong turns, etc.)
•Everything else

Perform it on:
•Early mockups

A/B Testing
Quantitative testing, single or multi-variant.
It is good for:
•Testing new designs against old to see if there was an improvement
•Spotting bugs or problems in new designs (if they fail against old designs)
•Testing single variable changes to see what converts best (ie. CTA text or Header messaging on a page)
•Idea validation (in that you can put out a “stub” and see if people try to interact with it before dedicating yourself to building a whole product)

It sucks for:
•Understanding WHY that single variable change or new design was better or worse.

Perform it on:
•Your actual product and every change you ever release

Other
There are several other things you can use like card sorting (good for figuring out big taxonomies or hard navigation problems), focus groups (good if you like getting one opinion from a lot of people - I hate them), statistical quant (good if you hate iterating), journal studies (good if you want to follow patterns over a long period of usage), and some others I’m probably forgetting, but I don’t really recommend any of these for startups, since they tend to be expensive and hard to run.

Plain Customer Interviews (remote or in person)
A lot of people run these before they have any products at all. They’re part of customer validation. They’re harder to do than you think though, because you’re just talking abstractly about things, and it’s often hard to get good information from them. Watching people in a contextual inquiry or observational study will typically give you better information because you can see what people do rather than just believe what they say.